

**Natural Product-Derived Lipopeptides are Selective and Effective Modulators of
the Coactivator Med25**

by

Olivia N. Pattelli

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in The University of Michigan
2022

Doctoral Committee:

Professor Anna K. Mapp, Chair
Associate Professor Tomasz Cierpicki
Professor David H. Sherman
Associate Professor Gregory G. Tall

Olivia N. Pattelli

patteoli@umich.edu

ORCID iD: [0000-0002-6849-601X](https://orcid.org/0000-0002-6849-601X)

© Olivia N. Pattelli 2022

Acknowledgements

The last year and a half of graduate school has been filled with the strangest times of my life as a result of a never-ending pandemic. Yet, somehow this time has also been some of the best and that is truly because of the people I have spent it with. The end of this journey in Michigan is filled with bittersweet emotions, but I feel so much gratitude for the people in my life who have surrounded me with constant support and warmth through it all.

The first person I would like to thank is my advisor, Professor Anna Mapp. You always have been there for me when I need you, both with science and with life. You have taught me how to be an independent scientist and pushed me out of my comfort zone in such a gentle way. You have been the most supportive mentor and have always seen the potential in me, even if I didn't. You approach everything you do with such kindness and understanding and it has been a joy to work for you the past five years.

Next, I would like to thank my dissertation committee members: Professor David Sherman, Professor Greg Tall, and Professor Tomasz Cierpicki. I am so appreciative of the wonderful feedback and support I have gotten from you all over the past few years. I have learned so much from you and have enjoyed the times we have met to discuss my research. Thank you so much for all the ideas and discussions; I am a better scientist because of them.

My first research experience was at the University of Illinois where I worked with a graduate student, Betsy Parkinson, in Professor Paul Hergenrother's lab doing synthesis. To Betsy, thank you for being the most wonderful mentor an undergraduate could ask for. You were so kind and so patient and you made my first time in a research lab feel so comfortable. Thank you to Paul for taking me in as a biology major who knew very little about organic chemistry. You gave me the opportunity to explore my interests in your lab and encouraged me to further pursue those interests by applying to chemical biology

graduate programs. It is safe to say, that I would not be here, earning my Ph.D. without guidance from both of you.

To my wonderful labmates, each and every one of you holds a very special place in my heart. It isn't very often that you find a group of people who make coming into work so easy and enjoyable. You are not just my coworkers, but you have also become my friends and I am so grateful for all the memories we have made throughout the years. I need to give a very special shout out to Estefania. During the pandemic, we were isolated to a bay together without anyone else around. Though it was lonely for us being away from the rest of the lab, I am thankful for the time I got to spend with you and the friendship we have developed as a result. There is truly no one else I would rather talk about Marvel movies and TikTok with.

The next person I would like to thank is my roommate Jessi. We have been living together for five years now and you have been such a wonderful support to have throughout the years. There are so many memories memories that we have made along the way and I have truly gone through it all with you. It has been an honor to live with you. Oh also, thank you for always helping me capture whatever types of creatures are trying to inhabit our apartment in any given week. I literally couldn't do any of it without you.

Next I want to give a shout out to Sam and Jorge, who are not only my labmates but also very close friends. You were two of the first people I met when I moved to Michigan and you have been by my side through every up and down that's happened since. I am so happy to have you both in my life. I appreciate both of you and our friendships and also our group chat. I hope it never dies.

Lastly, I would like to thank my family. First I need to give a shout out to my siblings Alex, Lisa, and Amanda, who I constantly admire as the youngest sibling. I don't get to see you all very much, but when I do I am always filled with so much happiness. I look up to each of you and am so happy to have you all in my life. I also need to give a shout out to my two perfect nieces who are just the sweetest, most lovely additions to our family. I love being your auntie and watching you both grow and explore the world.

Finally, I would like to thank my parents. You both have provided me with everything I could have ever asked for and you have never asked for anything in return. You have given me the space to grow and learn and carve my own path out of this life. You have

stood by my side, watching me make mistake after mistake, all while remaining supportive and patient. You offer constant stability and love and I sit here and hope that one day I can be at least half the people that you are. I am so immensely proud to be your daughter and I am constantly grateful to have you both as my parents.

Table of Contents

Acknowledgments	ii
List of Figures	viii
List of Tables	xi
List of Abbreviations	xii
Abstract	xv

Chapter 1. Targeting Transcriptional Protein-Protein Interactions

1.1 Abstract	1
1.2 Coactivator•activator PPIs are required for regulated transcription	2
What are transcriptional coactivators and activators?	2
Mediator is a coactivator complex	4
Med25 is Tail Mediator subunit	6
Med25 AcID has two distinct binding faces	8
Specificity in coactivator•activator complexes	11
Coactivator-activator complexes and disease	12
1.3 The promise and challenge of synthetic regulators of coactivator-activator PPIs	15
Taking cues from nature: PPI modulators	15
‘Druglike’ inhibitors are hard to find	18

Natural products are attractive molecules to target transcriptional PPIs	21
Med25 AcID is difficult to target	24
1.4 Dissertation Summary	27
1.5 References	30
Chapter 2. An Amphipathic Lipopeptide Modulates Med25 PPIs	
2.1 Abstract	35
2.2 Introduction	36
Lipopeptide natural products	36
Amphipathic peptides and lipids bind Med25	37
Identification of natural product lipopeptide to target Med25•activator PPIs	38
2.3 Results and Discussion	40
34913-1 analogs define critical components for inhibitory activity	40
Substitution in the lipid tail leads to improved activity	45
Lipopeptide analog 34913-8 demonstrates selectivity for Med25 PPIs	55
Lipopeptide 34913-8 engages with Med25 in a cellular context	56
2.4 Conclusions and Future Directions	58
2.5 Materials and Methods	60
2.6 References	73
Chapter 3. Lipopeptide 34913-8 is a Transcriptional Domain Mimic	
3.1 Abstract	77
3.2 Introduction	78
Activator•coactivator molecular recognition	78

Moving away from a non-specific binding model	79
3.3 Results and Discussion	82
Analysis of 34913-8 analogs with alanine substitutions	82
Importance of charged amino acids for 34913-8 binding	85
Determination of selectivity of the substitution analogs	87
34913-8 shows selectivity for ETV1/5	89
3.4 Conclusions and Future Directions	91
3.5 Materials and Methods	93
3.6 References	97
Chapter 4. Conclusions and Future Directions	
4.1 Summary	100
4.2 Conclusions	101
4.3 Future Directions	103
4.4 References	108
Appendix A. Peptide and Lipopeptide Characterization	111

List of Figures

Figure 1.1	Hub proteins interact with multiple binding proteins	3
Figure 1.2	Cryo-EM map of the Mediator tail module	5
Figure 1.3	Coactivator Med25 acts as a hub protein	6
Figure 1.4	Med25 and its binding domains	7
Figure 1.5	Structures of transcriptional coactivator ABDs	8
Figure 1.6	Distinct binding faces of Med25 AcID	11
Figure 1.7	Med25 upregulated in patient-derived, triple-negative breast cancer cell line.	13
Figure 1.8	Molecular intervention of Med25 AcID	15
Figure 1.9	Diversity of transcriptional protein-protein interactions	17
Figure 1.10	Binding faces of CBP GACKIX and its interaction pathways	18
Figure 1.11	Small molecule modulation of diverse classes of transcriptional protein-protein interactions	19
Figure 1.12	Examples of historically important natural products	21
Figure 1.13	Examples of natural product small molecules for PPIs	23
Figure 1.14	Theory behind fluorescence polarization assay	26
Figure 1.15	Fluorescence polarization assay to identify inhibitors of Med25•activator PPIs	27
Figure 2.1	Lipopeptide 34913-1 is a potent inhibitor of Med25•ATF6 α PPI	40
Figure 2.2	Lipopeptide 34913-1 contains three distinct components that can be modified	41

Figure 2.3	Peptides not sufficient for inhibitory activity against Med25•ATF6 α	43
Figure 2.4	Simplified lipid tail increases inhibitory activity for Med25•ATF6 α	44
Figure 2.5	NP40 detergent determines aggregation contributes to potency	45
Figure 2.6	Synthesis of (2S, 3R)-3-hydroxy-2-methylundecanoic acid	46
Figure 2.7	Analog 34913-8 is potent and selective inhibitor for Med25 PPIs	47
Figure 2.8	NP40 detergent determines aggregation doesn't contribute to potency	48
Figure 2.9	^1H , ^{13}C CSPs map induced by binding of 34913-8	50
Figure 2.10	^1H , ^{15}N CSPs map induced by binding of 34913-8	52
Figure 2.11	^1H , ^{13}C and ^1H , ^{15}N CSPs map induced by binding of 34913-9	54
Figure 2.12	Analog 34913-8 stabilizes full-length Med25 in VARI068 cell extracts	57
Figure 2.13	Analog 34913-8 shows inhibition of Med25 in a cellular context	58
Figure 3.1	Models of molecular recognition	79
Figure 3.2	ETV activator TAD sequence alignment	80
Figure 3.3	Variable residues mediate differences in Med25•ETV conformation behavior	81
Figure 3.4	Alanine residue substitutions of 34913-8	83
Figure 3.5	Alanine substitutions determine important residues for inhibitory activity	85
Figure 3.6	Charged residue substitutions of 34913-8	86
Figure 3.7	Relative inhibitory activity of substitution analogs	87
Figure 3.8	Determination of the selectivity of charge substitution analogs	89
Figure 3.9	34913-8 selectively inhibits Med25•ETV PPIs	91
Figure 4.1	Synthesis of lipid tail stereoisomers	105
Figure A.1	Structure of analog 34913-2	123
Figure A.2	Structure of analog 34913-3	126

Figure A.3	Structure of analog 34913-4	128
Figure A.4	Structure of analog 34913-5	130
Figure A.5	Structure of analog 34913-6	132
Figure A.6	Structure of analog 34913-7	135
Figure A.7	Structure of analog 34913-8	137
Figure A.8	Structure of analog 34913-9	139
Figure A.9	Structure of analog 34913-Ala1	142
Figure A.10	Structure of analog 34913-Ala2	145
Figure A.11	Structure of analog 34913-Ala3	148
Figure A.12	Structure of analog 34913-Ala4	151
Figure A.13	Structure of analog 34913-Ala5	154
Figure A.14	Structure of analog 34913-Ala6	157
Figure A.15	Structure of analog 34913-Ala7	160
Figure A.16	Structure of analog 34913-Gln1	163
Figure A.17	Structure of analog 34913-Asn2	166
Figure A.18	Structure of analog 34913-Gln1Asn2	169
Figure A.19	Structure of substitution analog impurity	172

List of Tables

Table 1.1	Tail subunits of Mediator that complex with TFs	6
Table 1.2	Med25 activator TADs and their binding faces	10
Table 2.1	Sequence of fluorescein labeled activator peptides	63
Table 2.2	Masses of fluorescein labeled activator peptides	63
Table 2.3	Structure of lipopeptide analogs and their masses	65
Table 3.1	Sequence of fluorescein labelled ETV peptides	94
Table 3.2	Masses of fluorescein labelled ETV peptides	94
Table 3.3	Structure of lipopeptide substitution analogs and their masses	95
Table A.1	Peak integrations of 34913-Ala1	143
Table A.2	Peak integrations of 34913-Ala2	146
Table A.3	Peak integrations of 34913-Ala3	149
Table A.4	Peak integrations of 34913-Ala4	152
Table A.5	Peak integrations of 34913-Ala5	155
Table A.6	Peak integrations of 34913-Ala6	158
Table A.7	Peak integrations of 34913-Ala7	161
Table A.8	Peak integrations of 34913-Gln1	164
Table A.9	Peak integrations of 34913-Asn2	167
Table A.10	Peak integrations of 34913-Gln1Asn2	170
Table A.11	Mass of impurity in substitution analogs	172

List of Abbreviations

ABD	activator binding domains
AcID	activator interacting domain
ATF6 α	activating transcription factor 6 α
BME	β -mercaptoethanol
CRISPR	clustered regularly interspaced short palindromic repeats
Cryo-EM	cryogenic electron microscopy
CSP	chemical shift perturbation
DBD	DNA binding domain
DCM	dichloromethane
DIC	diisopropylcarbodiimide
DIPEA	N,N-diisopropylethylamine
DMem	Dulbecco's Modified Eagle Medium
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNP	Dictionary of Natural Products
DTT	dithiothreitol
ER	endoplasmic reticulum
ETS	erythroblast transformation specific
ETV	Ets translocation variant
FA	fatty acyl
FBS	fetal bovine serum
FITC	fluorescein isothiocyanate
Fmoc	fluorenylmethyloxycarbonyl
FP	fluorescence polarization
FPLC	fast protein liquid chromatography
HBTU	hexafluorophosphate benzotriazole tetramethyl uronium

HER2	human epidermal growth factor receptor 2
HOBt	hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum coherence
IDR	intrinsically disordered region-1
IPTG	isopropyl β -D-1thiogalactopyranoside
K _d	dissociation constant
K _i	inhibitory constant
KIX	kinase-inducible domain interacting
LC-MS	liquid-chromatography mass spectrometry
MDM2	mouse double minute 2 homolog
mMed	mammalian Mediator
MMPs	matrix metalloproteinases
NA	norstictic acid
NMR	nuclear magnetic resonance
NR	nuclear receptor
PA	psoromic acid
PBS	phosphate buffered saline
PEA3	polyomavirus enhancer activator 3
PPI	protein-protein interactions
PTOV1	prostate Tumor Overexpressed 1
QTOF MS	quadropole time-of-flight mass spectrometry
RNA-seq	RNA sequencing
RT-qPCR	real-time quantitative polymerase chain reaction
S/N	signal-to-noise
SA	surface area
SDS-Page	sodium dodecyl sulphate–polyacrylamide gel electrophoresis
TAD	transcriptional activation domain
TAZ1	transcriptional adaptor zinc-binding 1
TBP	TATA box-binding protein
TF	transcription factor

TFA	trifluoroacetic acid
TEA	triethylamine
TIPS	triisopropylsilane
TNBC	triple-negative breast cancer
UPR	unfolded protein response
UV/vis	ultraviolet/visible spectroscopy
VP16	Herpes simplex virus protein 16
VWA	Von Willebrand factor type A

Abstract

Transcription plays a vital role in regulation of gene expression, with interference of this regulation often being prominent in disease. A key component of this process is the protein-protein interactions (PPIs) that occur between transcriptional coactivators and transcriptional activators, making them excellent targets for modulating gene expression for mechanistic insight and potential therapeutic development. Synthetically modulating these PPIs has inherent value for the further understanding of their mechanistic action, though targeting them is often considered undruggable due to their shallow binding surfaces and modest binding affinities. Natural products are an interesting class of compounds that have demonstrated success in inhibiting some of these difficult coactivator•activator PPIs, showcasing their potential as effective modulators.

Med25 is a coactivator subunit of the much larger coactivator complex, the Mediator. Med25 interacts with the transcriptional activation domain (TAD) of several different activators through its activator interaction domain (AcID). Several of its activator binding partners have been implicated in cancer and disease, including the ETV/PEA3 family and the oxidative stress response factor ATF6 α . The overarching goal of this dissertation was to identify and characterize small molecule inhibitors of the Med25 AcID•activator PPIs that will allow can be used for further study of other coactivator•activator PPIs.

This dissertation first describes the identification of a lipopeptide natural product, 34913-1, that was revealed to inhibit the Med25•ATF6 α PPI *in vitro* and in a cellular context. The structure of the lipopeptide allowed for the synthesis of several analogs that

tested via fluorescence polarization assays to determine the important components of the lipopeptide for inhibition of Med25 AcID. This resulted in the synthesis of the lead lipopeptide, 34913-8, that demonstrated inhibition of Med25 AcID *in vitro* and full length Med25 in mammalian cell extracts, while selectively inhibiting Med25 AcID PPIs over other coactivator•activator PPIs. The resemblance of the peptide sequence of 34913-8 to that of TADs suggested these lipopeptides interact with Med25 AcID similarly to the TAD of transcriptional activators. This led to the synthesis of several substitution analogs that were tested to determine the role that specific residues of the 34913-8 peptide sequence have on its inhibitory activity and selectivity for the Med25 PPIs. It was determined that several residues within the peptide sequence do serve an important purpose in selectively inhibiting these interactions. The overall results in this dissertation demonstrate the potential use of lipopeptide 34913-8 as a template for further development of analogs that can effectively and selectively target coactivator•activator PPIs.

Chapter 1

Targeting Transcriptional Protein-Protein Interactions

1.1 Abstract

Transcription is a central cellular process tightly regulated by many different protein-protein interactions (PPI). Dysregulation of the PPIs between transcriptional coactivators and activators is a key cause of transcriptional abnormalities underpinning several diseases, including cancer, developmental disorders, and diseases of protein homeostasis. Given their role in disease progression, these coactivator•activator PPIs are attractive targets for synthetic modulation, though they were once considered undruggable due to their dynamic nature, broad binding surfaces, and relatively weak binding affinity. While discovery of successful small molecules remains difficult, natural products have demonstrated their potential in inhibiting these difficult coactivator•activator PPIs.

This introductory chapter will define and detail the important features of transcriptional PPIs necessary for transcriptional activity and their roles in disease. The difficulties in challenging these interactions are explored, with several examples of the advancement of small molecule inhibitors discussed. Finally, a strategy is proposed for the development of small molecules to target the more difficult transcriptional PPIs and their downstream effects

1.2 Coactivator•activator PPIs are required for regulated transcription

What are transcriptional coactivators and activators?

Transcription is a complex process responsible for the expression of proteins essential to cellular function and homeostasis, and protein-protein interaction (PPIs) networks play a crucial role in tightly regulating this process. Some proteins in these networks, such as transcriptional coactivators, have been classified as hub proteins due to their ability to interact with several different proteins at the same or overlapping binding surfaces. Several studies have indicated that these hub proteins often contain intrinsically disordered regions (IDR), which convey varying degrees of structural plasticity.¹⁻⁴ This often results in distinct conformations that interact with a variety binding partners, often with moderate binding affinity (Figure 1.1). As a hub protein, transcriptional coactivators act as a bridge between DNA-bound transcriptional activators and the general components of the transcriptional machinery.^{5,6} The PPI between transcriptional coactivators and transcriptional activators occurs during the early stages of transcription and facilitate the appropriate nucleation of the pre-initiation complex.⁷ The focus of this dissertation work is on the coactivator Med25.

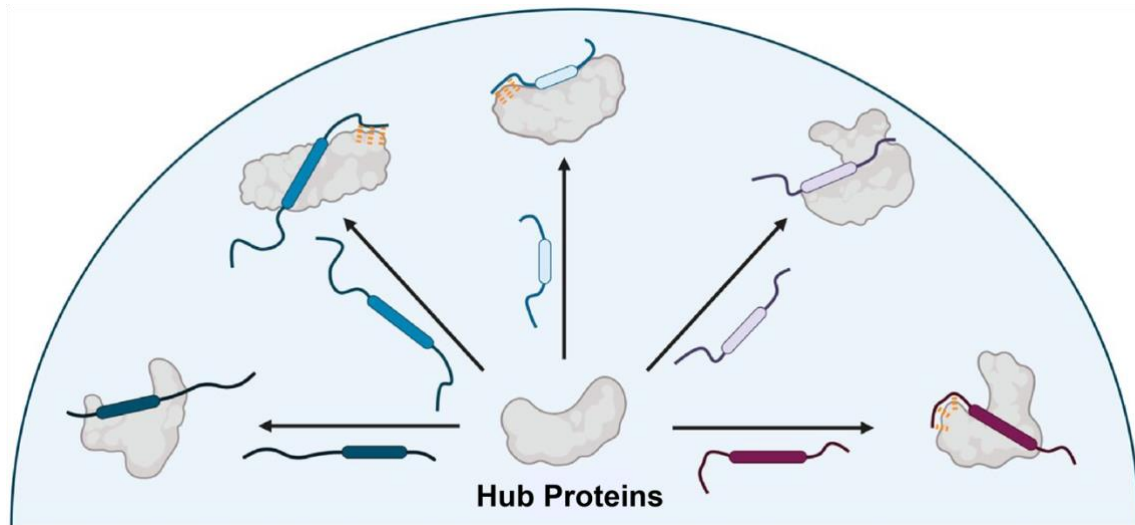


Figure 1.1. Hub proteins interact with multiple binding proteins. Hub proteins interact with several different binding partners using a common domain. This often results in the binding partners binding in unique conformations using the same binding face of the protein, conveying a degree of specific binding recognition. This figure is adapted from Henley, et al. (2020).⁸

Transcriptional coactivators contain an activator binding domain (ABD) that interacts with transcriptional activators. Activators contain two distinct domains responsible for binding, both which play a role in target gene specificity. The first is the DNA binding domain (DBD), which interacts with a DNA sequence near a promoter or enhancer region, thereby playing a critical role in regulating gene expression through the recruitment of gene-specific and general transcriptional machinery to the DNA. The second is the transcriptional activation domain (TAD), which makes crucial contact with coactivators and determines the duration of gene activation through the recruitment of additional transcriptional machinery.⁷ The DBD of activators are more well understood due to the discovery of a wide variety of well-defined structural motifs leading to a more thorough structural and functional characterization. As a result, this domain has been more successfully inhibited. TADs of activators, on the other hand, tend to adopt secondary structure upon interacting with their binding partners but generally lack

structure in their unbound state. Numerous studies have identified an amphipathic helix motif as the most observed motif among TADs when bound.^{9,10}

Mediator is a coactivator complex

One well-studied transcriptional coactivator hub protein complex is the Mediator, which serves as both a physical and functional bridge between transcription factors, including transcriptional activators and the general transcriptional machinery. Initially identified in yeast in the early 1990's, the evolutionarily conserved Mediator complex is comprised of over 30 subunits in humans.¹¹⁻¹³ Several structural advancements have determined the subunit composition of the Mediator complex. Most recently, El Khattabi et al. reported a 5.9 Å cryogenic electron microscopy (cryo-EM) structure mapping the mammalian Mediator (mMed) complex in its entirety.¹¹ Mediator is organized into 4 distinct modules: head, middle, tail, and CDK8 kinase. These modules are stabilized to one another by subunit Med14, which acts as the backbone connecting the modules. The head and middle modules form the Core Mediator, composed of the minimal set of subunits active in transcription and are responsible for interacting with RNA Pol II. The tail module contains several subunits that primarily complex with transcription factors for transcriptional upregulation (Figure 1.2).^{11,12}

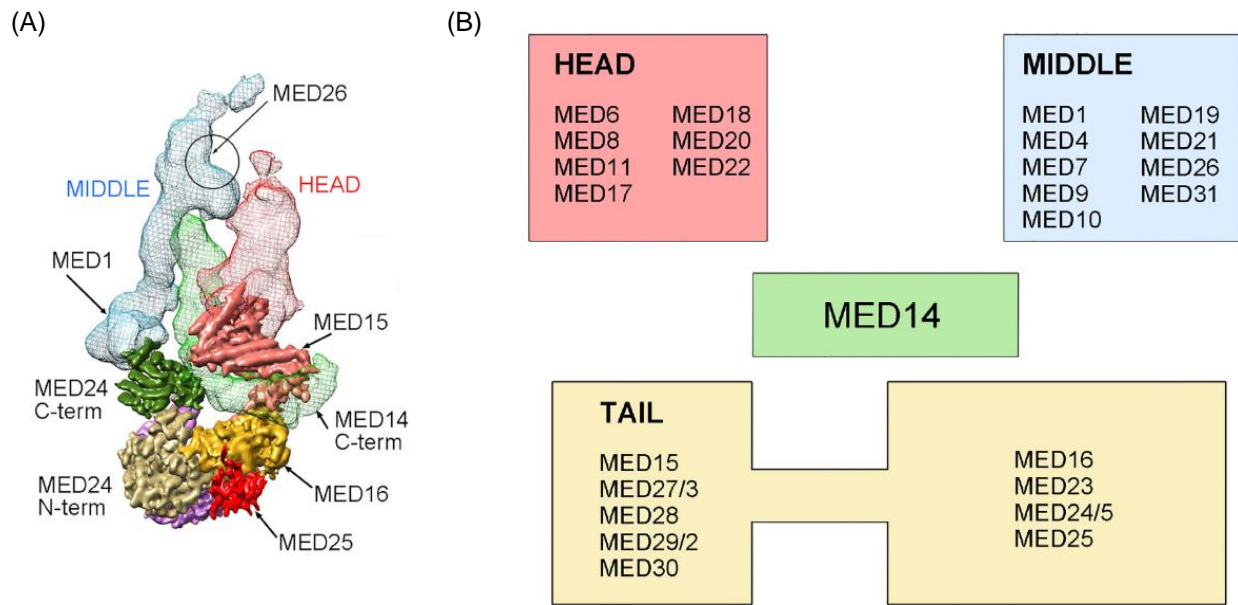


Figure 1.2. Cryo-EM map of Mediator tail module. (A) Subunit organization and interactions of the mMed tail module shown in relation to the other Mediator modules. Med25 is shown in red. (B) Assignment of modules for each Mediator subunit. Med14 was not assignment to a module due to its role as the Mediator backbone. This figure is adapted from El Khattabi et al. (2019) and reproduced with permission from Springer Nature.¹¹

Early studies of yeast and mammalian Mediator revealed that the complex has a well-defined but dynamic structure, suggesting that the presence of specific transcriptional activators can induce conformational changes in the complex. Several findings identified Mediator subunits that interact with these transcription factors tend to cluster in the variable tail module (Table 1.1). Structural analysis, primarily guided by nuclear magnetic resonance (NMR) spectroscopy data, uncovered detailed molecular interactions between TADs of transcriptional activators and subunits of the tail module.¹²⁻
¹⁹ Deletions of several of these nonessential tail subunits resulted in cells that were still viable, though their growth was slowed.¹¹ Further transcriptome analysis revealed that the number of affected genes increases as the number of deleted subunits increases, with cells composed of no tail subunits showing a 1.3x fold transcriptome downregulation.

This suggests that the presence of these Mediator tail subunits is important for expression of many genes.^{11,12}

Mediator Tail Subunit	Transcription Factor(s)
<i>Med15</i>	Gcn4 (yeast), SREBP
<i>Med23</i>	<i>C/EBPβ</i> , <i>ELK1</i> , <i>IRF7</i> , <i>ESX</i> , <i>E1a</i>
<i>Med25</i>	VP16, ETV1/ETV4/ETV5, ATF6α, p53

Table 1.1. Tail subunits of Mediator that complex with TFs. Subunits of the tail module and their known interacting transcriptional activators.

Med25 is a Tail Mediator subunit

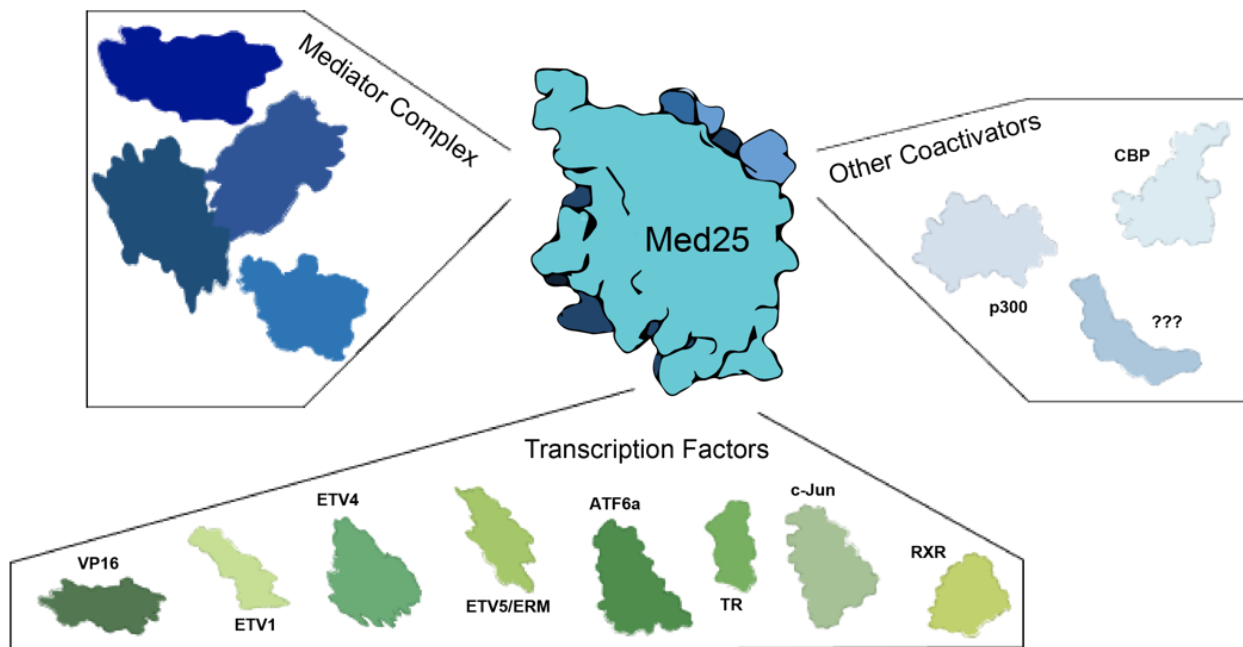


Figure 1.3. Coactivator Med25 acts as a hub protein. Med25 is a hub protein that can interact with multiple different proteins involved in the activation of transcription, including the rest of the Mediator complex, other coactivators, and several transcriptional activators. This is accomplished through only a couple binding domains on the protein, indicating there is significant overlap between the binding sites of all proteins known to interact with Med25.

Med25 is a nonessential coactivator subunit that acts as a transcriptional hub protein located in the tail module of the Mediator complex (Figure 1.3). Coactivator Med25 is composed of 747 residues and includes three distinct domains: a nuclear receptor (NR) box that is reported to interact with retinoic acid receptor α and the estrogen receptor α , a Von Willebrand factor type A (VWA) that anchors Med25 to the rest of the Mediator complex, and the activator interacting domain (AcID) which interacts with DNA bound transcriptional activators (Figure 1.4).¹⁶⁻¹⁸ Interest in understanding the mechanism for transcriptional activator recruitment resulted in the determination of a solution structure

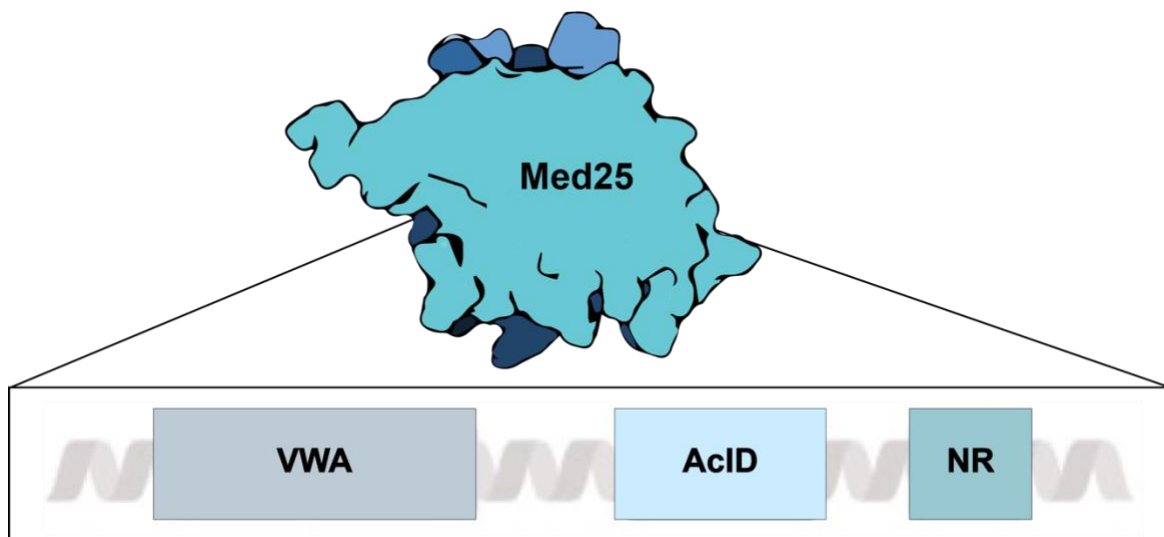


Figure 1.4. Med25 and its binding domains. The coactivator Med25 contains 3 domains: the VWA domain that bridges it to the rest of the Mediator complex, the ABD domain termed AcID that interactions with transcriptional activators, and a nuclear receptor (NR) that can bind to hormones such as retinoic acid.⁶²

of Med25 AcID using NMR, which showed it was composed of a central, closed seven-stranded beta-barrel (B1-B7) framed by 3 α -helices (α 1- α 3).

Previous sequence homology studies showed that prostate tumor overexpressed 1 (PTOV1), which contains 2 AcID domains in tandem, was the only other human protein with homology to Med25 AcID with sequence identify of 79.7% and 71.8% respectively,

suggesting these two PTOV1 AcIDs folded similarly to that of Med25 AcID.^{20,21} The β -barrel of Med25 AcID makes its structure distinct from typical coactivator motifs, such as the kinase-inducible domain interacting (KIX) and the transcriptional adaptor zinc-binding 1 (TAZ1) domains, which consist primarily of α -helices (Figure 1.5).

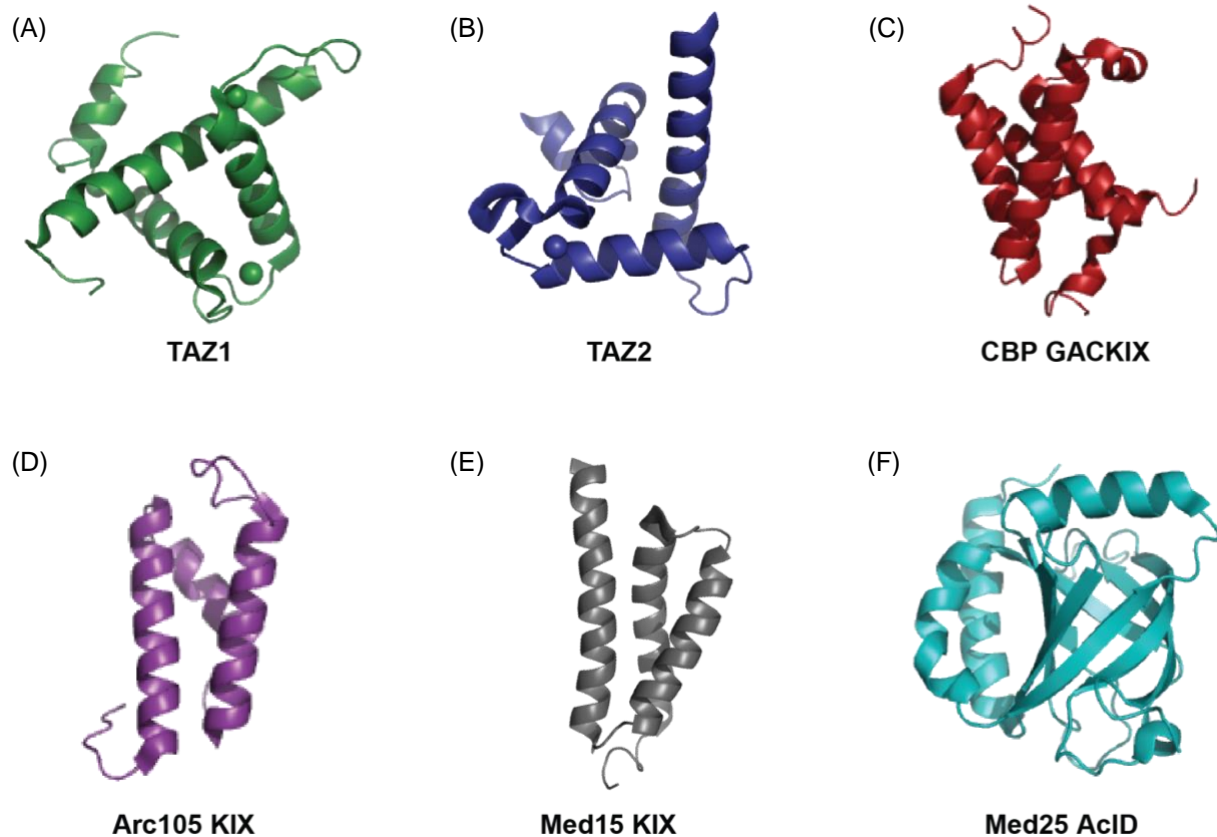


Figure 1.5. Structures of transcriptional coactivators ABDs. There are several examples of ABDs that are helical bundles connected by dynamic loop regions. (A) Taz1 (PDB ID 1U2N) and (B) TAZ2 (PDB ID 1F81) are two coactivators of the CBP complex. (C) CBP GACKIX (PDB ID 1SB0) (D) Arc105 KIX (PDB ID 2GUT), and (E) Med15 KIX (PDB ID 2K0N) are additional coactivators containing several helices. (F) The Med25 subunit of the human Mediator complex contains a seven-stranded β -barrel core, shown in turquoise, flanked by more common structural features: α -helices and loops (PDB ID 2XNF).

Med25 AcID has two distinct binding faces

Coactivator Med25 was first identified as a binding partner of the herpes simplex virus protein 16 (VP16) activator. VP16 contains two distinct functional TADs, H1 and H2,

located in the N-terminal region, which are individually able to activate transcription.¹⁷ Initial experiments demonstrated that the complete TAD of VP16 makes numerous contacts across the surface of the binding interface of Med25 AcID and adopts a partially folded conformation when bound.¹⁶ Additional NMR spectroscopy experiments monitoring distinct changes in the chemical shift of several amide resonances revealed evidence suggesting that the C-terminal portion of the VP16 TAD may bind to the opposite side of the AcID β -barrel than the N-terminal portion of the VP16 TAD.¹⁶⁻¹⁸

In addition to VP16, several other transcriptional activators have been reported to interact with Med25 AcID. This includes activating transcription factor 6 α (ATF6 α), a master regulator of endoplasmic reticulum (ER) stress response genes, the ETS translocation variant (ETV)/polyomavirus enhancer activator 3 (PEA3) family of erythroblast transformation specific (ETS) transcriptional activators that are involved in cancer, cell growth, and differentiation, and p53, a tumor suppressor that plays an important role in cell cycle arrest, DNA repair, metabolism, and apoptosis (Table 1.2).²²⁻
²⁵ Several studies have demonstrated that the TADs of VP16 and ERM directly interact with the H1 face of Med25 AcID despite sequence variabilities, suggesting that the structural plasticity of AcID plays an important role in the molecular recognition of the cognate activator binding partners. NMR experiments indicate ETV5 binding results in perturbations predominantly on the H1 surface of AcID.⁵⁰ It also shows that ATF6 α binding leads to significant chemical shift perturbations of the H2 binding surface, further supporting the notion that there are two distinct binding faces on Med25 AcID that can interact with transcriptional activators.²⁵ Furthermore, each of the transcriptional activator

binding partners has distinct but overlapping chemical shift patterns, suggesting several unique binding modes within AcID.¹⁶⁻¹⁸

TAD	Residues	Med25 AcID Binding Face
<i>Full-length VP16</i>	411–490	H1/H2
<i>VP16 H1</i>	411–452	H1
<i>VP16 H2</i>	453–490	H2
<i>ETV5/ERM</i>	38-68	H1
<i>ATF6α</i>	40-66	H2
<i>Ful-length p53TAD</i>	1-73	H1/H2
<i>p53 TAD1</i>	15-29	H1
<i>p53 TAD2</i>	39-57	H2

Table 1.2. Med25 activator TADs and their binding faces. TADs known to interact with Med25 AcID. Residues numbers of the TADs used for HSQC NMR experiments are provided, as well as the respective Med25 AcID faces they bind.

The presence of two binding sites on AcID suggests that there may be allosteric connection between the them, which was investigated by using two solvent exposed cysteines (C497 and C506) on the H1 face to screen a tethering library of disulfide containing point mutations of the H1-binding portion of VP16 (Figure 1.6). These experiments concluded that the tethering of the G450C mutation resulted in 100% tethering to C506 on the H1 site, resulting in significant perturbations on the H2 face, indicating that allostery does play a role in the binding of these activators to Med25 AcID. Molecular dynamics simulations revealed that while the β -barrel remained structurally rigid, the loops and flanking α -helices surrounding the β -barrel were more dynamic, suggesting a molecular recognition model that can be used for allosteric modulation.²⁵

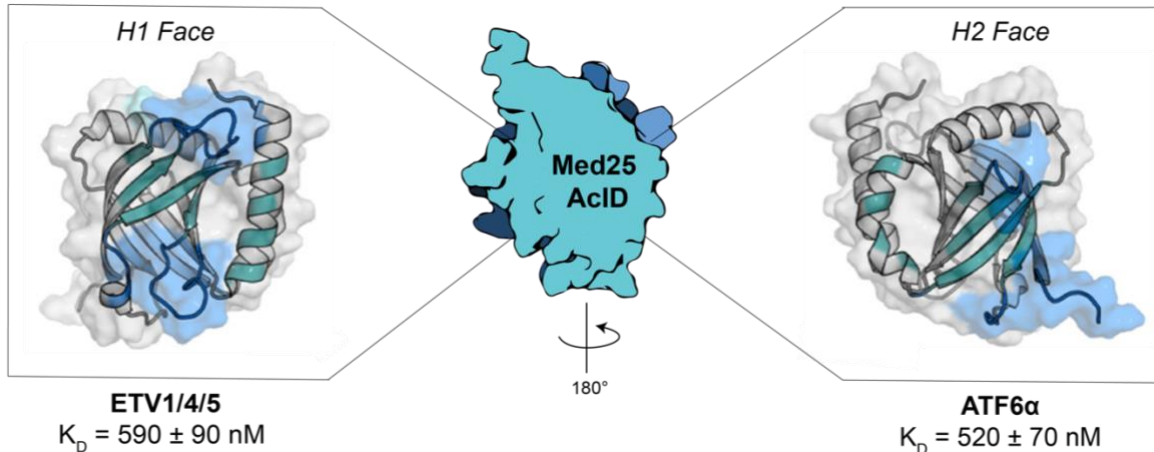


Figure 1.6. Distinct binding faces of Med25 AcID. It is proposed that Med25 AcID contains two distinct binding faces, H1 and H2, on opposite sides of the binding interface. The ETV activators are proposed to interact mostly with the H1 face while ATF6 α is proposed to interact with the H2 face. Both activators bind with modest binding affinities to Med25 AcID. Data that informs these conclusions are indicated as Med25 AcID residues that are significantly perturbed in HSQC NMR.

Specificity in coactivator•activator complexes

Transcriptional activators contain a TAD that is crucial for their interactions with the ABD of coactivators. The largest class of activators are TADs that are rich in acidic and hydrophobic residues and form amphipathic helices upon binding to transcriptional coactivators. For decades, coactivator•activator PPIs were thought to be largely nonspecific, with activators being deemed as ‘negative noodles’ or ‘acid blobs’.²⁶ Increased transcriptional activation is associated with an excess of acidic residues in the TAD of activators, giving the domain an overall negative charge. However, several lines of data indicate this is not the case and while acidic residues are important for transcriptional activation, they are not always sufficient for transcriptional activation.^{27,28}

One example of specific TAD binding is the yeast activator Gal4, which is required for the metabolism of galactose.⁶³ Gal4 is inhibited by Gal80 by blocking TATA box-binding protein (TBP) binding and activating transcription in the absence of galactose, with the Gal4 interacting with both Gal80 and TBP. Performance of cysteine scanning

mutagenesis identified amino acid residues that were key for these PPIs. Proline substitutions were made at key residues to evaluate the effect of secondary structure on transcriptional activation. These studies revealed that a 16-residue portion of the Gal4 TAD could participate in two different PPIs that have distinct sequence and structural requirements, with binding to Gal80 being highly sensitive to cysteine substitution compared to TBP. Taken together, this supports a model of molecular recognition for TAD-coactivator complexes that includes sequence-specific interactions.^{10,27-30}

An additional example focuses on Med25 and its interactions with the ETV/PEA3 family of Ets transcriptional activators including ETV1, ETV4, and ETV5. These three activators share >85% sequence identity in their TADs and contain similar arrangements of acidic and hydrophobic residues. It was determined that the small sequence differences in the TAD of these activators result in distinct conformational binding modes with Med25, despite a similar binding affinities. Furthermore, data revealed Med25 AcID went through significant remodeling upon complex formation, with dynamic substructures playing a particularly prominent role in the remodeling. In conclusion, a specific recognition model has emerged suggesting that the activator binding domains of coactivators recognize a diverse set of activators through specific intermolecular interactions.⁸

Coactivator-activator complexes and disease

Precise regulation of transcription is essential for cellular homeostasis, so it is unsurprising that dysregulation of transcription is associated with a wide variety of disease. Mutations or altered expression of transcriptionally-relevant proteins such as

coactivators and activators commonly contribute to diseases such as cancer, autoimmunity, diabetes, developmental/neurological disorders, obesity, and cardiovascular diseases.³¹⁻³² Using clustered regularly interspaced short palindromic repeats (CRISPR) knockouts of Med25, 750 genes were shown to have altered expression underscoring its importance in transcription.¹¹ Additionally, RNA sequencing (RNA-seq) experiments observed an upregulation in Med25 expression in a patient-derived triple negative breast cancer (TNBC) cell line, VARI068 (Figure 1.7).^{34,35}

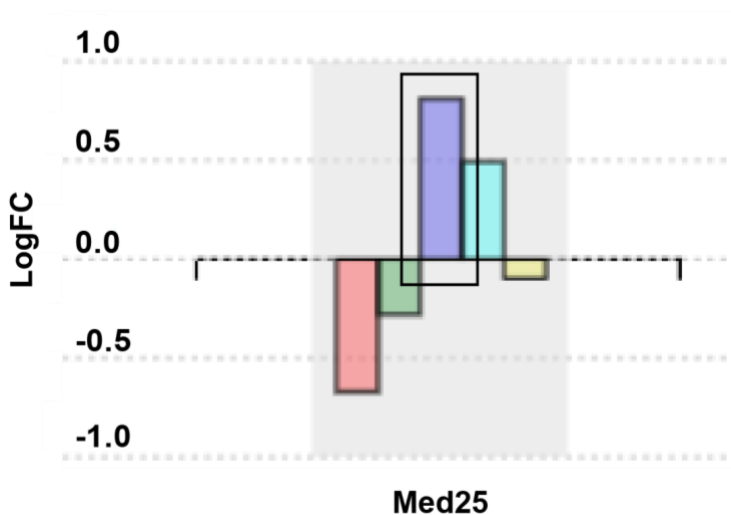


Figure 1.7. Med25 upregulated in patient-derived, triple-negative breast cancer cell line. The performance of RNA-seq experiments revealed the upregulation of Med25 in the patient-derived, TNBC cell line, VARI068 (shown in purple), as compared to the non-tumorigenic epithelial cell line, MCF10A. The additional bars correspond to patient-derived, TNBC cell line Sum149 (red), patient-derived, HER2-positive breast cancer cell line, Sum190 (green), and TNBC cell line MDA-MB-231 (cyan), and estrogen receptor-positive breast cancer cell line, MCF7 (yellow).

The binding partners of Med25 have also been implicated in the expression of several genes involved in disease. ATF6 α is a key player in the unfolded protein response (UPR), a cellular stress response triggered by the accumulation of unfolded or misfolded proteins in the ER. This accumulation results in the translocation of ATF6 α to the nucleus where it binds to Med25, leading to the upregulation of anti-apoptotic and pro-survival

genes. The most notable of these genes is HSPA5, or chaperone protein GRP78, that acts as the master regulator of the UPR. These cellular responses to ER stress are critical for cell survival and have been implicated in diseases such as neurodegeneration, diabetes, and atherosclerosis.¹⁴ It has also been reported that HSPA5 is overexpressed in a variety of cancer cell lines, solid tumors, and human cancer biopsies making Med25•ATF6 α an interesting therapeutic coactivator•activator PPI to target.³⁶⁻³⁹

ETV1 (Er81), ETV4 (PEA3), and ETV5 (ERM) are implicated in several diseases. The exact role and function of these activators is not entirely known, though the dysregulation of their expression is frequently associated with carcinogenesis, specifically metastasis and tumorigenesis. Studies indicate that these PEA3 transcription factors are activated in response to combinatorial activation of PI3-kinase and Ras signaling, a commonality in metastatic tumors. They also demonstrate high correlation to the expression of certain matrix metalloproteinases (MMPs), which play an important role in invasion by degradation of the extracellular matrix. Taken together, this further supports the argument that Med25 and its activator binding partners are important in a wide array of disease, making them particularly attractive PPIs for further investigation and potential molecular intervention (Figure 1.8).⁴⁰⁻⁴⁴

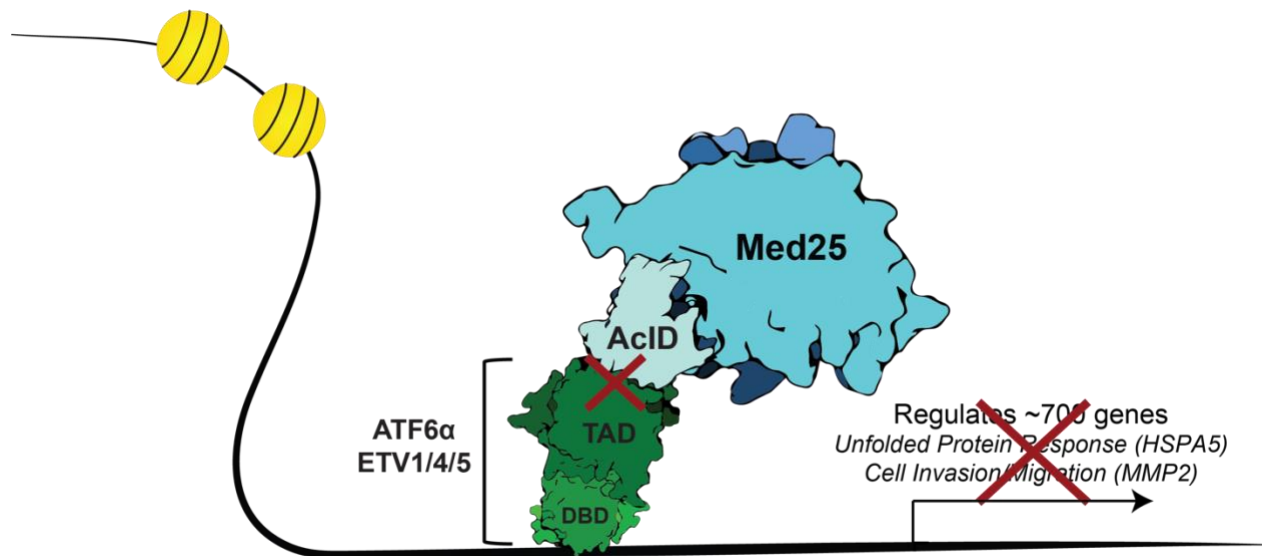


Figure 1.8. Molecular intervention of Med25 AcID. Med25 AcID and its PPIs with transcriptional activators have been implicated in a variety of human disease. For this reason, these specific interactions have high potential for molecular intervention using small molecules.

1.3 The promise and challenge of synthetic regulators of coactivator-activator PPIs

Taking cues from nature: PPI modulators

Traditionally, drug discovery has focused primarily on targeting the enzymatic components of multi-protein complexes, often by inhibiting the active site of signaling enzymes. There are advantages to targeting transcriptional PPIs involved in the signaling pathway as opposed to directly targeting the enzyme's catalytic activity. First, modulation of these specific PPIs disrupt certain stages of the signaling pathway involved in a disease state without completely turning it off, offering a level of selective biological tuning. Additionally, the interface of PPIs tend to be more unique and variable than enzyme active sites, which are typically highly conserved, possibly leading to greater target selectivity.⁴⁵ The frequent role of coactivator-activator complexes in human disease continues to make them an attractive target for molecular intervention. Despite the advantages of targeting

coactivator•activator PPIs, they have been challenging to inhibit with small molecules for a multitude of reasons and have largely been classified as ‘undruggable’.⁴⁵⁻⁴⁷

Difficulty in early small molecule discovery often stemmed from the lack of pockets that typically bind small molecules, with early high-resolution structures of coactivator•activator PPIs showing binding surfaces that are generally large and shallow (~1000-2000 Å² per side). Mutational analysis of some of these surfaces revealed localized regions of residues that were critical for interface binding, referred to as ‘hot spots’, narrowing the targeted PPI interface.^{45,46} It was determined that PPIs that have this smaller surface area containing ‘hot spot’ regions that bind with high affinity are more well-characterized and easier to inhibit, while PPIs with broad, flat surface areas that bind with relatively weak affinity tend to be more difficult (Figure 1.9).⁴⁸ It has been frequently observed that orthosteric inhibitors are ideal for targeting the PPIs characterized by small SA and tight affinity while allosteric inhibitors can be useful in targeting PPIs with more difficult features to inhibit.

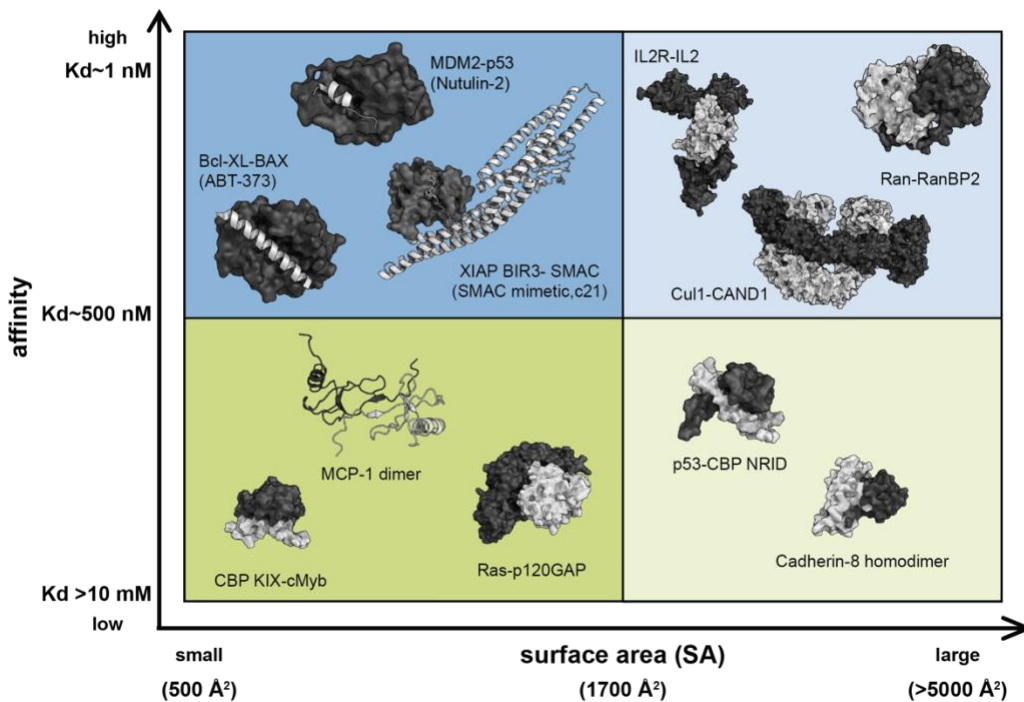


Figure 1.9. Diversity of transcriptional protein-protein interactions. Examples of transcriptional PPI classes according to the affinity and surface area of the specific interaction. Examples found at a transcriptional level within all four classes of PPIs are shown. This figure is adapted from Mapp et al. (2015) and reproduced with permission from Springer Nature.⁴⁶

Coactivator•activator PPIs containing large, complex binding surfaces are more difficult to modulate. Many of these transcriptional coactivators are conformationally dynamic, allowing them to interact with a variety of transcriptional activators. Overlapping regions of amino acids participate in these interactions, with each coactivator•activator complex requiring its own distinct structural conformation.⁴⁸ An example of this is the coactivator GACKIX motif of CREB binding protein (CBP) which can form both binary and ternary complexes with more than 15 transcriptional activators using only two binding surfaces, conveying a high level of selectivity between coactivators and their activator binding partners (Figure 1.10).^{49,50} These two binding surfaces are allosterically connected, meaning changes in one site will result in changes of the other site, allowing

for cooperative assembly of the binary and ternary complexes. Allosteric small molecules can be advantageous when looking at complex coactivator•activator PPIs as they often bind to more easily targetable sites that are away from the binding interface, restructuring the PPI surfaces and allowing for more tractable inhibition. Additionally, allosteric binding can induce subtle conformational changes that favor binding to specific transcriptional activators over others, conveying a level of selectivity that is not always possible with orthosteric inhibitors.

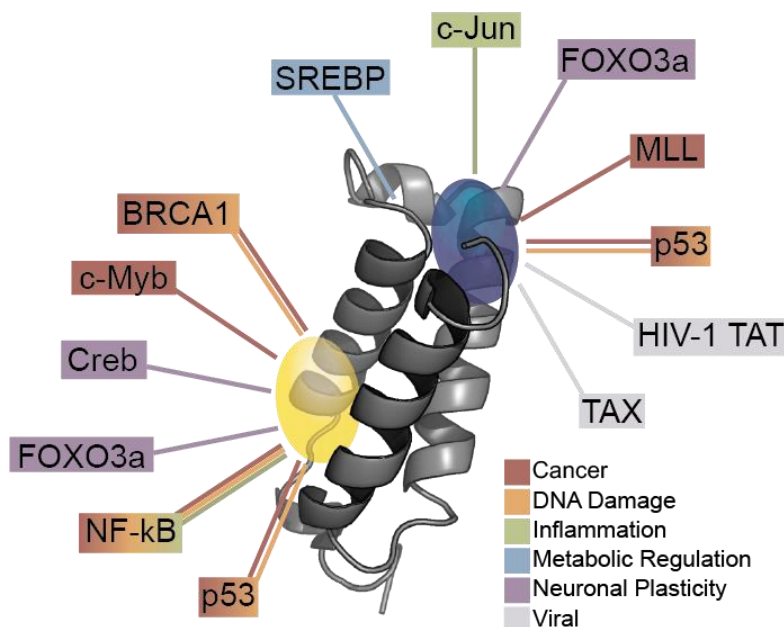


Figure 1.10. Binding faces of CBP GACKIX and its interaction partners. The two binding faces of CBP GACKIX and its interaction partners. GACKIX consists of two binding faces, termed the cMyb binding face and MLL binding face. As illustrated, GACKIX binds to a suite of different binding partners implicated in a variety of networks and diseases.^{49,50} This figure is adapted from Mapp et al. (2015) and reproduced with permission from Springer Nature.⁴⁶

'Druglike' inhibitors are hard to find

Among the estimated 650,000 PPIs in the human interactome, less than 0.01% have been successfully targeted.⁵¹ While PPIs have historically been challenging to target with small molecules, certain interactions with a specific set of features have been more

pliable to therapeutic intervention than others. Most of these known PPI modulators (>60%) can target PPIs that have a binding affinity of <math><1 \mu\text{M}</math> and a surface area less than 1800 A^2 , meaning targeting high affinity, small SA PPIs is more feasible than ones with weak affinity and high SA (Figure 1.11).⁴⁸

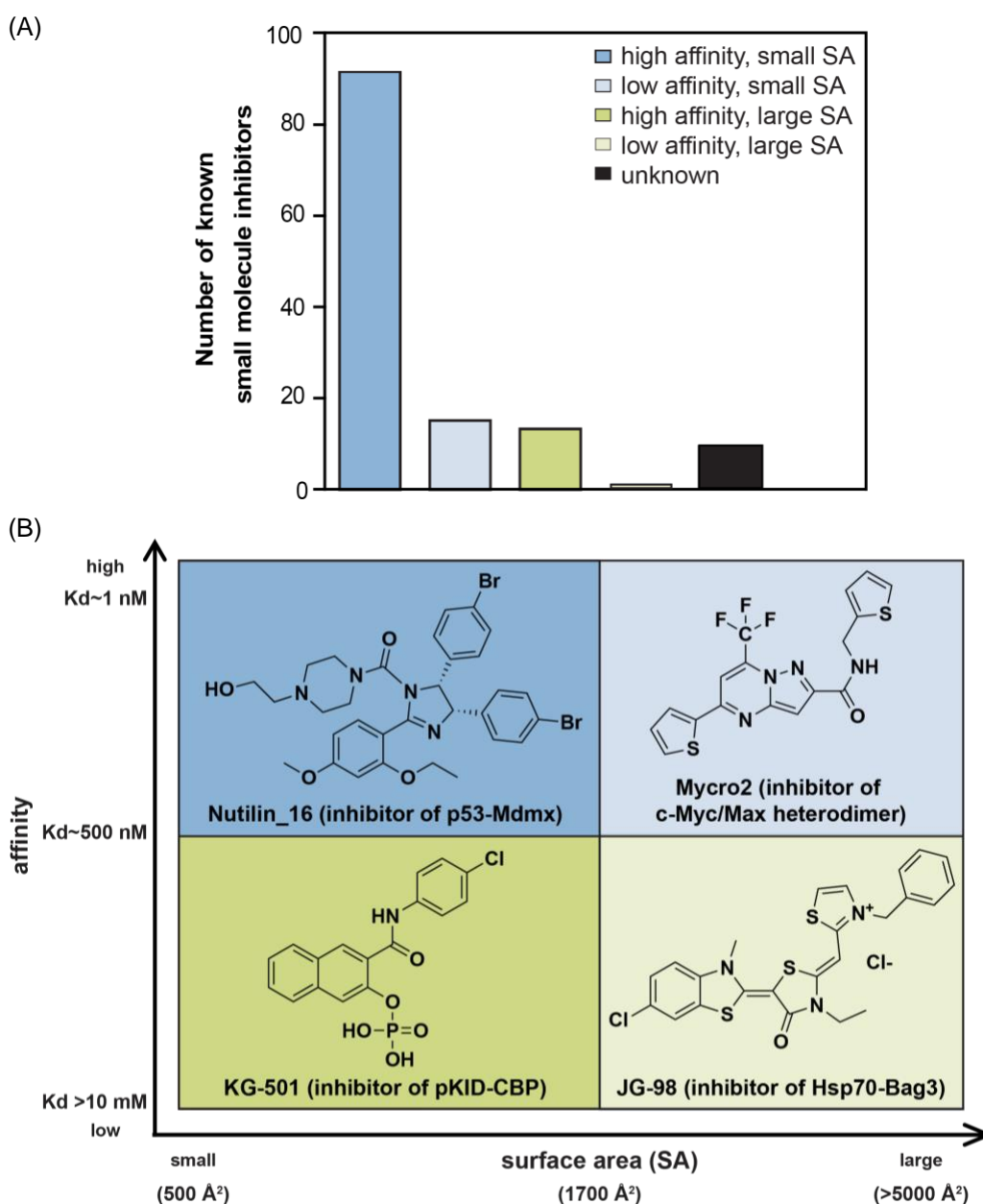


Figure 1.11. Small molecule modulation of diverse classes of transcriptional protein-protein interactions. (A) Modulators of PPIs were categorized based on diverse classes of PPIs. A majority of known PPI modulators successfully target PPIs with high affinity and small surface areas. (B) Examples of small molecules that target unique PPIs that represent each of the four PPI classes are provided. Figure adapted from Cesa et al. (2015).⁴⁸

The PPI between p53 and mouse double minute 2 homolog, MDM2, serves as a great example of the overall success in targeting high affinity, low SA PPI. The tumor suppressor, p53, is an activator that can form many complexes with different regulatory partners at varying ranges of affinity and SA binding. Many of these complexes are high affinity PPIs with a relatively small surface area that rely on a centralized region of residues that are crucial for the affinity of the interaction. Specifically, the interaction between p53 and MDM2 is a classic example of a successfully targeted transcriptional PPI. MDM2 acts as a masking protein when bound to p53, preventing p53 from functioning as a transcriptional activator. This PPI's relatively small SA (<1800 Å²) has made it easier to characterize structurally, leading to the complex being successfully modulated by small molecules, several of which are currently in clinical trials.^{46,51}

Transcriptional PPIs that have weaker binding affinities and contain broader interaction surfaces are difficult to modulate using small molecules due to their conformationally dynamic complexes. This means typical probe discovery methods are not well equipped for these interactions, resulting in the lack of small molecules to act as starting points for drug design.^{46,51,52} Additionally, the size and character of typical small molecule libraries have had limited success in the discovery of novel small molecules, resulting in the desire to develop and screen larger libraries with a wider array of compounds.^{51,52} Several methods have been developed that demonstrate the potential for the discovery of small molecules to target these more difficult-to-target transcriptional PPIs.

Natural products are attractive molecules to target transcriptional PPIs

Natural products are chemical compounds which are synthesized by biological organisms and act as excellent starting points for drug discovery.⁵³ Natural products rose in popularity after the widespread use of penicillin during World War II, resulting in a pharmaceutical expansion to screen microorganisms for potential disease-treating small molecules. By 1990, approximately 80% of drugs were natural products or derivatives. These compounds treated a wide array of diseases, including bacterial infections and cancer, revolutionizing modern medicine (Figure 1.12). Though there has been tremendous success in isolating natural products for drug discovery, there is still much to explore. It is thought that only a small portion of the world's biodiversity (<10%), especially in the marine environment, has been explored, suggesting that the majority of potential compounds have yet to be discovered.^{54,55}

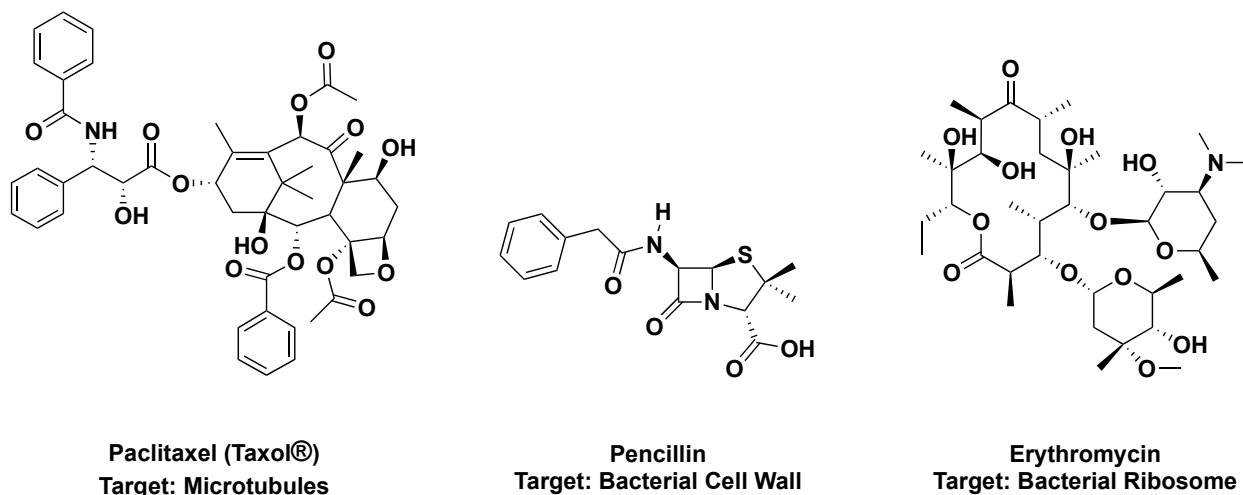


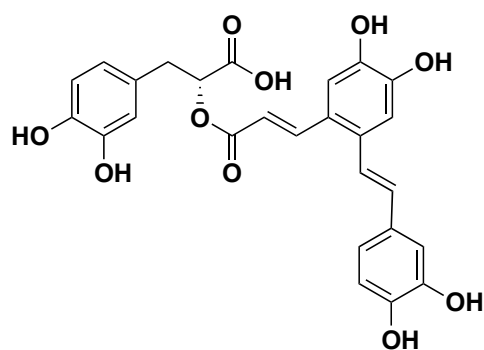
Figure 1.12. Examples of historically important natural products. Natural products cover a wide scope of structurally and functionally diverse small molecules, several of which have been highly successful in targeting disease.⁵⁵

Natural products have proven to be a rich source for biologically active compounds. These compounds tend to be attractive as drug leads due to their unique

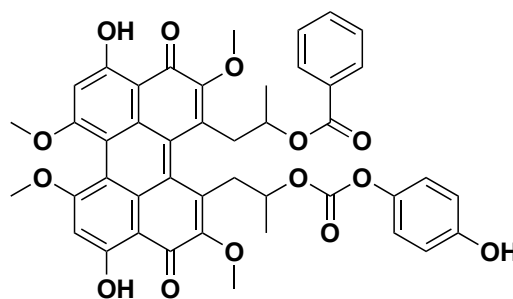
structural complexity and molecular diversity,⁵⁴⁻⁵⁶ especially in comparison to more easily accessible synthetic compounds. More specifically, of the over 70,000 natural products present in the Dictionary of Natural Products (DNP), about 40% contain structurally unique elements that are not present in the tested pool of synthetic compounds.⁵⁶ Overall, natural products tend to be higher in molecular weight and contain a lower number of nitrogen, halogen, and sulphur atoms when compared to synthetic molecules, though there is an increase in oxygen atoms present. Statistically, natural products contain more sp^3 hybridized atoms and have a larger number of rings and chiral centers per compound compared to synthetic compounds, attributing to their more sterically complex structures.⁵⁶ These unique functional elements can also be classified based off the natural source in which they are isolated, though biological activity in some sources have been investigated more thoroughly than others. This highlights the potential to explore certain natural sources that may contain specific structural elements of interest for a specific biologically target.⁵⁶ Overall, natural products offer fascinating structures that are not commonly found in synthetic substances, making them interesting compounds for therapeutic modulation of transcriptional PPIs.

Recently, natural product drug discovery has showcased its potential in targeting the more difficult transcriptional PPIs with the development of several probes. An excellent example from our lab identified two interesting natural products, sekikaic acid and lobaric acid, to target the coactivator CBP/p300 KIX and its activator binding partners. These compounds belong to a class of molecules known as depsides and depsidones, respectively, that are isolated from lichens.^{57,58} Two libraries were screened to identify potential compounds to modulate CBP/p300 KIX. The first was a 50,000 compound library

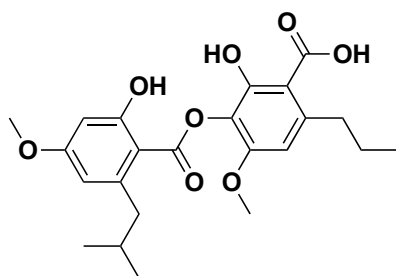
containing small drug-like molecules and the second was a 15,000 compound library containing natural product extracts isolated from marine-sediment derived microbes, cyanobacteria, lichens, and sponges. We identified sekikaic acid and lobaric acid as potent inhibitors of CBP KIX•MLL PPI (IC_{50} of 17 μ M and 25 μ M respectively), making them two of the most effective inhibitors of these conformationally dynamic PPIs.⁵⁷ This intimately demonstrates the significant potential that natural products have in targeting difficult transcriptional PPIs (Figure 1.13).



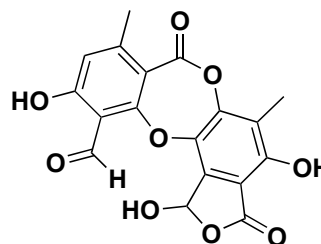
Salvionic Acid A
Target: SH2 domains of Src.Lck



PKF115-584
Target: beta-catenin/Tcf4



Sekikaic Acid
Target: CBP KIX



Norstictic Acid
Target: Med25 AcID

Figure 1.13. Examples of natural product small molecules for PPIs. Several natural products have been discovered to selectively inhibit transcriptional PPIs that have larger binding surface areas and more modest binding affinities.^{57,59,61,62}

Med25 AcID is difficult to target

Med25 PPIs are relevant in several human diseases, though it has proven difficult to modulate these complexes using small molecules due to their moderate binding affinities and transient interactions with multiple different binding partners. There have been a few methods and compounds from our lab that have been employed to probe the Med25•activator complexes to varying degrees. Firstly, the presence of two solvent exposed cysteine residues (C497 and C506) on the H1 face of AcID were utilized for a tethering screen composed of 1,600 compound fragments. Compound **22** was identified from this screen due to the high efficiency in labelling C506 on AcID. Transient kinetics experiments determined this fragment could recapitulate important binding features of the natural activator, despite **22** being significantly smaller than the activator binding partners. This showed promise in targeting Med25•activator PPIs using small molecules, despite the characteristics that make these interactions particularly challenging.²⁵

Additional interest in identifying probes for Med25 resulted in the screening of several commercially available libraries containing ~4000 compounds total against the Med25•VP16 PPI, with two depsidone family natural products, norstictic acid (NA) and psoromic acid (PA) emerging as hits with apparent IC₅₀ values of 2.3 ± 0.1 μM and 3.9 ± 0.3 μM, respectively. NA was shown to form covalent bonds with lysine residue side chains, primarily with K519 in the lysine rich dynamic loop on the border of the H2 face. Additional experiments indicated that NA acts as an orthosteric inhibitor of the H2 face of Med25 AcID and an allosteric inhibitor of the H1 face of Med25 AcID while also exhibiting high selectivity for Med25 PPIs when compared to other coactivators with similar binding characteristics. Finally, NA was determined to engage with endogenous Med25 and alter

its PPI network with downstream effects on tumor phenotype, further providing evidence that modulating Med25 with small molecules probes is a useful strategy to study its implication in human disease.⁵⁹ This also further demonstrates the potential in using a natural product or natural product-derived compound to modulate this specific transcriptional target.

A primary screening method used for the discovery of molecular probes is fluorescence polarization (FP). This assay is used to determine the apparent binding and dissociation between two molecules based on their molecular motion in solution. Generally, one of these molecules is small and fluorescently labeled, such as a peptide or ligand, while the other is larger in size (usually >10 kDa), such as a protein. These experiments are performed by titrating in the protein of interest to a fixed concentration of the fluorescently labeled probe. When excited by plane-polarized light, unbound fluorescently labeled probe will rotate quickly in solution, resulting in a low FP signal. As more protein of interest binds to the fluorescently labeled probe, the rotation will slow, resulting in an increase in FP signal compared to the unbound state (Figure 1.14).⁶⁰ To calculate the K_d of the two molecules, the protein concentration is plotted against the FP.

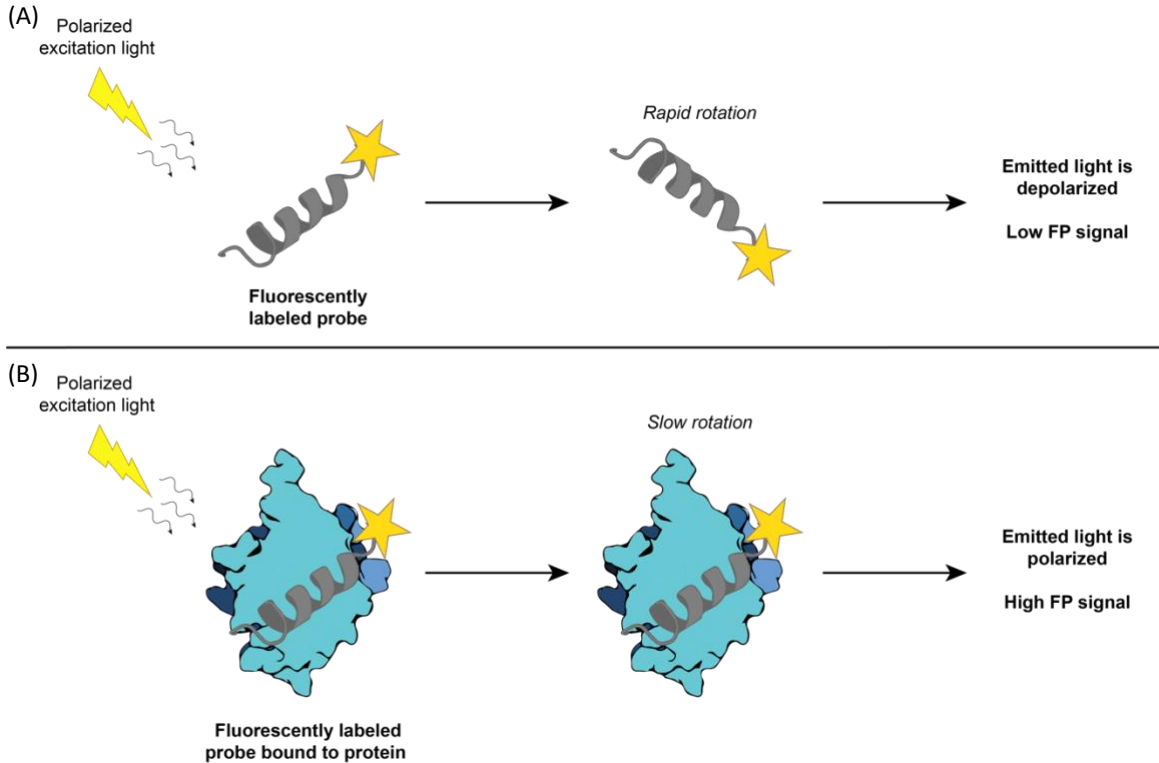


Figure 1.14. Theory behind fluorescence polarization assay. (A) Upon excitation with polarized light, a free fluorescently labeled probe rotates rapidly, resulting in the emission of depolarized light and a weak fluorescence polarization signal. (B) Upon excitation with polarized light, a fluorescently labeled probe bound to a protein will rotate slowly, resulting in the emission of polarized light and a high fluorescence polarization signal. Figure adapted from Du, Y (2015).⁶⁰

In our lab, the larger molecule is our protein of interest, Med25 AcID, while the fluorescently labeled probe is typically a fluorescein isothiocyanate (also referred to as FITC) labeled activator binding partner of Med25 AcID. To determine the inhibitory constant (K_i) of an inhibitor of interest, increasing concentrations of inhibitor are added to a fixed concentration of complexed protein and fluorescently labeled peptide based on their K_d . If the inhibitor is unable to compete off the larger fluorescently labeled peptide, rotation will remain slow, resulting in a high FP signal when excited by polarized light. If the inhibitor can compete off the fluorescently labeled peptide, rotation will increase, causing a decrease in FP signal when excited by plane-polarized light (Figure 1.15).⁶⁰ The K_i can be calculated by plotting the polarization against the inhibitor concentration to

determine if the compound of interest is an effective at inhibiting the Med25•activator PPI in solution.

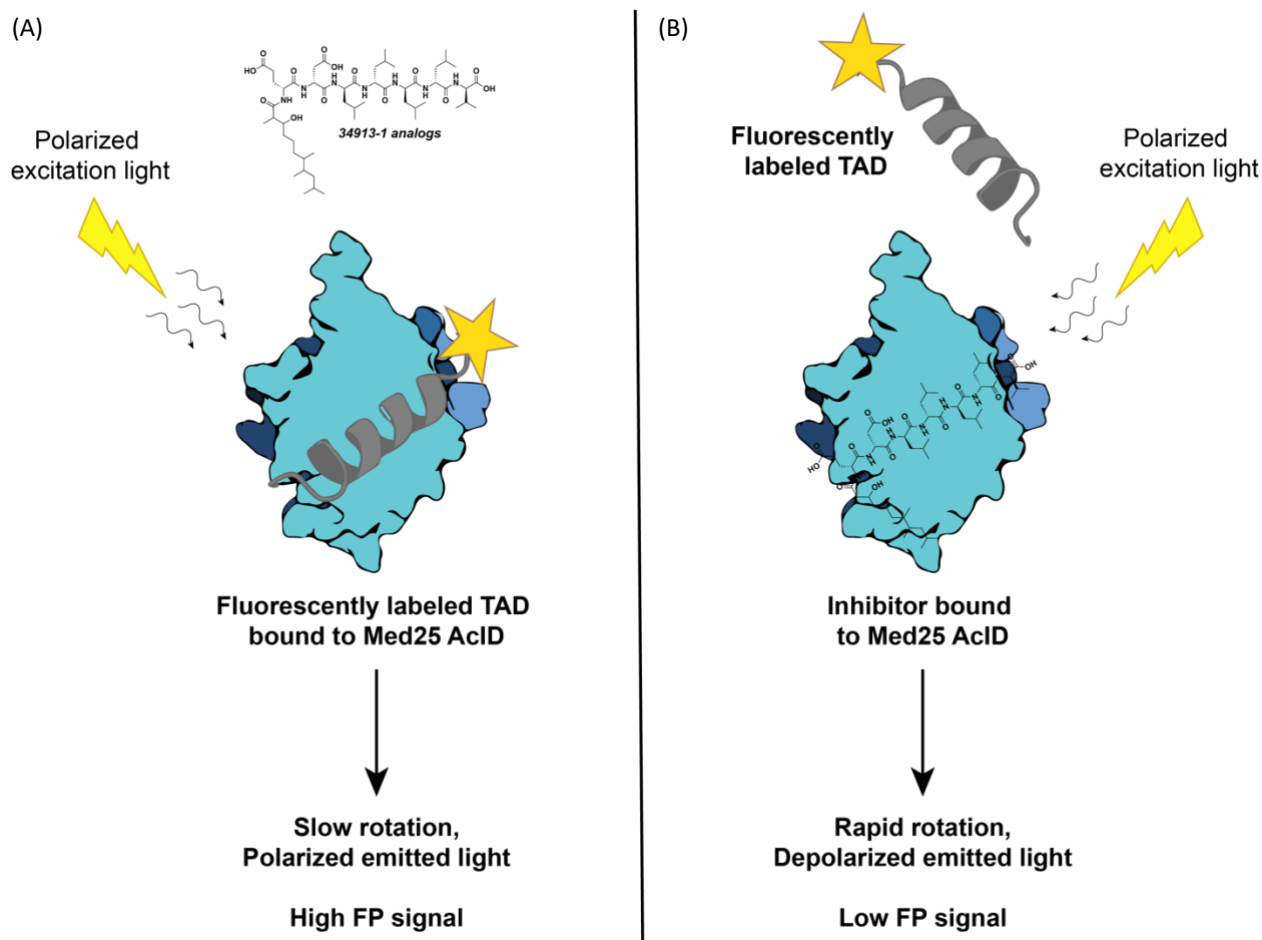


Figure 1.15. Fluorescence polarization assay to identify inhibitors of Med25•activator PPIs. (A) Upon excitation with polarized light, the fluorescently labeled activator peptide in complex with Med25 AcID rotates slowly, resulting in the emission of polarized light and a high fluorescence polarization signal. (B) Upon excitation with polarized light, an inhibitor binds to Med25 AcID by competing off the fluorescently labeled activator peptide, increasing the rotation of the complex. This results in the emission of depolarized light and a low fluorescence polarization signal.

1.4 Dissertation Summary

Due to the established importance of transcriptional coactivator•activator PPIs in various diseases, there is a necessity for the discovery of small molecule modulators that can successfully inhibit these interactions. The coactivator Med25 is a unique transcriptional target for the coactivator class of proteins due to its rare structural fold and

ability to interact with several activators implicated in cancer and disease. While several attempts have been made to target the Med25•activator PPIs, few have been successful. The subsequent chapters of this dissertation describe the development and characterization of a lipopeptide and its lead analogs that target Med25•activator PPIs.

In Chapter 2 of the dissertation, the discovery of a natural product denoted 34913-1 demonstrated activity for the Med25•ATF6 α PPI both *in vitro* and in a cellular context. Further simplified analogs of this natural product were synthesized to determine important components of its structure for Med25 AcID•ATF6 α PPI inhibition. This resulted in the development of two lead analogs, 34913-8 and 34913-9, that allowed for the investigation of a modified C-terminus for the activity and selectivity of these compounds both *in vitro* binding assays and using patient-derived breast cancer cell lines, informing the importance of molecular recognition for the successful inhibition of these coactivator•activator PPIs.

In Chapter 3, the peptide sequence of lead lipopeptide analog, 34913-8, was investigated due to its sequence similarities to the TADs of transcriptional activators, suggesting the lipopeptide may behave similarly. Several substitution analogs were synthesized to determine if amphipathicity or specific sequence identity is the driving force of inhibition for these lipopeptides. Further, these analogs will be tested against other coactivator•activator PPIs to determine the effects that slight sequence variation can have on the selectivity of these compounds for other coactivator proteins. Lastly, lead analog 34913-8 will be evaluated for its ability to selectively inhibit Med25 AcID•ETV PPIs, informing the possibility of using this lipopeptide and its analogs as probes to selectively inhibit one conformation of Med25 AcID over another. Taken together, the data in the

dissertation presents a class of molecular lipopeptide probes that can be designed and modified to successfully inhibit the more difficult to target coactivator•activator PPIs

1.5 References

1. Singh, G. P., Ganapathi, M., Dash, D. Role of intrinsic disorder in transient interactions of hub proteins. *Proteins*. 2007, 66(4), 761–765.
2. Schreiber, G., Keating, A. E. Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* 2011, 21(1), 50–61.
3. Higurashi, M., Ishida, T., Kinoshita, K. Identification of transient hub proteins and the possible structural basis for their multiple interactions. *Protein Sci.* 2008, 17(1), 72–78.
4. Krause, R., von Mering, C., Bork, P., Dandekar, T. Shared components of protein complexes--versatile building blocks or biochemical artefacts?. *BioEssays*. 2004, 26(12), 1333-1343.
5. Higurashi, M., Ishida, T., Kinoshita, K. Identification of transient hub proteins and the possible structural basis for their multiple interactions. *Protein Sci.* **2008**, 17(1), 72–78.
6. Krause, R., von Mering, C., Bork, P., Dandekar, T. Shared components of protein complexes--versatile building blocks or biochemical artefacts?. *BioEssays*. **2004**, 26(12), 1333-1343.
7. Ptashne, M. & Gann, A. *Genes and Signals (2002)*. (Cold Spring Harbor Lab Press, 2002).
8. Henley, M.J., Linhares, B.M., Morgan, B.S., et al. Unexpected specificity within dynamic transcriptional protein-protein complexes. *Proc. Natl. Acad. Sci.* **2020**, 117(44), 27346-27353.
9. Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature*. **1988**, 333(6170), 210-212.
10. Brzovic, P.S., Heikaus, C.C., Kisselev, L., et al. The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell*. **2011**, 44(6), 942-953.
11. El Khattabi, L., Zhao, H., Kalchschmidt, J., et al. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell*. **2019**, 178(5), 1145-1158.
12. Zhao, H., Young, N., Kalchschmidt, J., et al. Structure of mammalian Mediator complex reveals Tail module architecture and interaction with a conserved core. *Nat. Commun.* **2021**, 12(1), 1355.

13. Verger, A., Monté, D., Villeret, V. Twenty years of Mediator complex structural studies. *Biochem. Soc. Trans.* **2019**, 47(1), 399-410.
14. Sela, D., Conkright, J.J., Chen, L., et al. Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6 α . *J. Biol. Chem.* **2013**, 288(36), 26179-26187.
15. Landrieu, I., Verger, A., Baert, J.L., et al. Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* **2015**, 43(14), 7110-7121.
16. Vojnic, E., Mourão, A., Seizl, M., et al. Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **2011**, 18(4), 404-409.
17. Bontems, F., Verger, A., Dewitte F, et al. NMR structure of the human Mediator MED25 ACID domain. *J. Struct. Biol.* **2011**, 174(1), 245-251.
18. Milbradt, A.G., Kulkarni, M., Yi, T., et al. Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol.* **2011**, 18(4), 410-415.
19. Lee, M.S., Lim, K., Lee, M.K., Chi, S.W. Structural Basis for the Interaction between p53 Transactivation Domain and the Mediator Subunit MED25. *Molecules.* **2018**, 23(10), 2726.
20. Bénédict, P., Paciucci, R., Thomson, T.M., et al. PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene.* **2001**, 20(12), 1455-1464.
21. Yang, F., DeBeaumont, R., Zhou, S., Näär, A.M. The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci.* **2004**, 101(8), 2339–234.
22. Sela, D., Conkright, J.J., Chen, L., et al. Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6 α . *J. Biol. Chem.* **2013**, 288(36), 26179-26187.
23. Landrieu, I., Verger, A., Baert, J.L., et al. Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* **2015**, 43(14), 7110-7121.
24. Lee, M.S., Lim, K., Lee, M.K., Chi, S.W. Structural Basis for the Interaction between p53 Transactivation Domain and the Mediator Subunit MED25. *Molecules.* **2018**, 23(10), 2726.

25. Henderson, A.R., Henley, M.J., Foster, N.J., et al. Conservation of coactivator engagement mechanism enables small molecule allosteric modulators. *Proc. Natl. Acad. Sci.* **2018**, *115*(36), 8960-8965.
26. Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature.* **1988**, *333*(6170), 210-212.
27. Cress, W.D., Triezenberg, S.J. Critical structural elements of the VP16 transcriptional activation domain. *Science.* **1991**, *251*(4989), 87-90.
28. Regier, J.L., Shen, F., Triezenberg, S.J. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci.* **1993**, *90*(3), 883-887.
29. Warfield, L., Tuttle, L.M., Pacheco, D., et al. A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci.* **2014**, *111*(34), 3506-3513.
- 30 Tuttle, L.M., Pacheco, D., Warfield, L., et al. Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep.* **2018**, *22*(12), 3251-3264.
- 31 Bushwell, J.H. Targeting Transcription Factors in Cancer—from Undruggable to Reality. *Nat. Rev. Cancer.* **2019**, *19*(11), 611–624.
- 32 Brennan, P., Donev, R., Hewamana, S. Targeting Transcription Factors for Therapeutic Benefit. *Mol. Biosyst.* **2008**, *4*(9), 909–919.
- 33 Lambert, M., Jambon, S., Depaw, S., David-Cordonnier, M.H. Targeting Transcription Factors for Cancer Treatment. *Molecules*, **2018**, *23*(6), 1479.
- 34 Aw Yong, K.M., Ulintz, P.J., Caceres, S., et al. Heterogeneity at the Invasion Front of Triple Negative Breast Cancer Cells. *Sci. Rep.* **2020**, *10*(1), 5781.
35. Liu, M., Liu, Y., Deng, L., et al. Transcriptional Profiles of Different States of Cancer Stem Cells in Triple-Negative Breast Cancer. *Mol. Cancer.* **2018**, *17*(1), 65.
36. Haze, K., Yoshida, H., Yanagi, H., Yura, T., Mori, K. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol. Biol. Cell.* **1999**, *10*(11), 3787–3799.
37. Yoshida, H., Okada, T., Haze, K., et al. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol. Cell. Biol.* **2000**, *20*(18), 6755–6767.

38. Wang, Y., Shen, J., Arenzana, W., et al. Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. *J. Biol. Chem.* **2000**, 275(35), 27013–27020.
39. Fernandez, P. M., Tabbara, S.O., Jacobs, L.K., et al. Overexpression of the glucose-regulated stress gene GRP78 in malignant but not benign human breast lesions. *Breast Cancer Res. Treat.* **2000**, 59(1), 15–26.
40. Benz, C.C., O'Hagan, R.C., Richter, B., et al. HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene.* **1997**, 15(13), 1513–1525.
41. de Launoit, Y., Chotteau-Lelievre, A., Beaudoin, C., et al. The PEA3 group of ETS-related transcription factors. *Adv. Exp. Med. Biol.* **2000**, 480, 107–116.
42. Pellecchia, A., Pescucci, E., De Lorenzo, E., et al. Overexpression of ETV4 is oncogenic in prostate cells through promotion of both cell proliferation and epithelial to mesenchymal transition. *Oncogenesis.* **2012**. 1(7), e20.
43. Myers, E., Hill, A.D.K., Kelly, G., et al. A positive role for PEA3 in HER2-mediated breast tumour progression. *Br. J. Cancer.* **2006**, 95(10), 1404–1409.
44. Fauquette, V., Perrais, M., Cerulis, S., et al. The antagonistic regulation of human MUC4 and ErbB-2 genes by the Ets protein PEA3 in pancreatic cancer cells: implications for the proliferation/differentiation balance in the cells. *Biochem. J.* **2005**, 386(1), 35–45.
45. Pricer, R., Gestwicki, J.E., Mapp, A.K. From Fuzzy to Function: The New Frontier of Protein–Protein Interactions. *Acc. Chem. Res.* **2017**, 50(3), 584–589.
46. Mapp, A.K., Pricer, R., Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **2015**, 11(12), 891–894.
47. Thompson, A.D., Dugan, A., Gestwicki, J.E., Mapp, A.K. Fine-Tuning Multiprotein Complexes Using Small Molecules. *ACS Chem. Biol.* **2012**, 7(8), 1311–1320.
48. Cesa, L.C., Mapp, A.K., Gestwicki, J.E. Direct and Propagated Effects of Small Molecules on Protein–Protein Interaction Networks. *Front. Bioeng. Biotechnol.* **2015**, 3(119).
49. Dyson, H.J., Wright, P.E. Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. *J. Biol. Chem.* **2016**, 291(13), 6714–6722.
50. Thakur, J.K., Yadav, A., Yadav, G. Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Res.* **2014**, 42(4), 2112–2125.

51. Arkin, M., Tang, Y., Wells, J. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.* **2014**, 21(9), 1102-1114.
52. Arkin, M., Wells, J. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **2004**, 3(4), 301-317.
53. Breinbauer, R., Vetter, I.R., Waldmann, H. From protein domains to drug candidates-natural products as guiding principles in the design and synthesis of compound libraries. *Angew. Chem. Int. Ed.* **2002**, 41(16), 2879-2890.
54. Cragg, G.M., Newman, D.J. Natural products: a continuing source of novel drug leads. *Biochim. Biophys. Acta.* **2013**, 1830(6), 3670-3695.
55. Dias, D. A., Urban, S., Roessner, U. A historical overview of natural products in drug discovery. *Metabolites.* **2012**, 2(2), 303-336.
56. Henkel, T., Brunne, R.M., Müller, H., Reichel, F. Statistical Investigation into the Structural Complementarity of Natural Products and Synthetic Compounds. *Angew. Chem. Int. Ed.* **1999**, 38(5), 643-647.
57. Majmudar, C.Y., Højfedt, J.W., Arevang, C.J., et al. Sekikaic Acid and Lobaric Acid Target a Dynamic Interface of the Coactivator CBP/p300. *Angew. Chem. Int. Ed.* **2012**, 51(45), 11258-11262.
58. Fischer, E. SYNTHESIS OF DEPSIDES, LICHEN-SUBSTANCES AND TANNINS. *J. Am. Chem. Soc.* **1914**, 36(6), 1170-1201.
59. Garlick, J.M., Sturlis, S.M., Bruno, P.A., et al. Norstictic Acid Is a Selective Allosteric Transcriptional Regulator. *J. Am. Chem. Soc.* **2021**, 143(25), 9297-9302.
60. Du, Y. Fluorescence Polarization Assay to Quantify Protein-Protein Interactions in an HTS Format. *Protein-Protein Interactions.* **2015**, 1278, 529-544.
61. Lepourcelet, M., Chen, Y.-N.P., France, D.S., et al. Small-molecule antagonists of the oncogenic Tcf/ β -catenin protein complex. *Cancer Cell.* **2004**, 5(1), 91-102.
62. Sperl, B., Seifert, M.H.J., Berg, T. Natural product inhibitors of protein-protein interactions mediated by Src-family SH2 domains. *Bioorg. Med. Chem. Lett.* **2009**, 19(12), 3305-3309.

Chapter 2

An Amphipathic Lipopeptide Modulates Med25 PPIs

2.1 Abstract

The protein-protein interactions (PPI) between transcriptional coactivators and transcriptional activators are crucial for the proper expression of genes, and the dysregulation of these PPIs is a key driver of oncogenesis. The PPIs between the coactivator Med25, a subunit of the Mediator complex, and transcriptional activators such as ATF6 α and ETV5 are excellent examples. Binding of these cognate activators to Med25 results in the upregulation of disease-associated genes, making the PPIs particularly attractive targets for pharmacological modulation. However, the dynamic structures of the complexes and their modest activator binding affinities has hindered development of selective chemical probes.

Recently Beyersdorf and coworkers reported the discovery and characterization of a natural product lipopeptide, 34913-1, as an effective inhibitor of the Med25-ATF6 α PPI ($K_i = 8.8 \pm 0.6 \mu\text{M}$). In this Chapter, I examine the roles of the lipid and the peptide modules of the natural product through the chemical synthesis of analogs and functional assessment. This led to the identification of two lead analogs, 34913-8 and 34913-9, possessing comparable inhibitory activity ($K_i = 10 \pm 1 \mu\text{M}$ and $K_i = 6.6 \pm 0.3 \mu\text{M}$ respectively) to 34913-1. Data from *in vitro* binding experiments, HSQC NMR studies, and cell experiments demonstrate the importance that small changes in lipopeptide structure can have on the molecular recognition of these analogs for Med25•activator

PPIs. Taken together, these results highlight the utility of lipopeptides in the development of molecular probes to selectively inhibit Med25•activator PPIs.

2.2 Introduction

Lipopeptide natural products

Natural products have traditionally been a rich source of drug discovery due to their valuable chemical properties, including structural complexity, good permeability, and bioavailability. Lipopeptides are a diverse class of natural product secondary metabolites produced by a variety of fungi and bacteria. Structurally, lipopeptides are amphipathic molecules that are composed of a lipid moiety attached to a short linear or cyclic peptide chain.¹⁻⁴ Much of their structural diversity is due to the modifications in the length and composition of lipid tail moiety as well as the type, number, and configuration of the amino acids in the peptide sequence. Cyclic lipopeptides and lipopeptides composed of D-amino acids typically show increased resistance of the compounds to degradation by proteolytic enzymes.⁵

The lipopeptide class of natural products has received considerable attention due to their antibacterial, antiviral, and antifungal properties, often accompanied by minimal toxicity. Several lipopeptides, such as daptomycin and caspofungin, are FDA-approved as anti-microbial agents, further demonstrating the potential of these compounds as therapeutics.^{3,5} Lipopeptides have also demonstrated their significant potential for the treatment of various cancers. The amphiphilic lipopeptide Surfactin, for example, disrupts cell cycle progression, apoptosis, and cell proliferation in several cancer cell lines through the suppression of MMP-9 expression by inhibition of transcription factors, NF- κ B and

AP-1 in breast cancer cell lines.⁶ Additionally, cyanobacteria are a particularly rich source for anticancer agents, with lipopeptides representing over 40% of their isolated compounds.⁷

Amphipathic peptides and lipids bind Med25

The coactivator Med25, has several transcriptional activator binding partners, including the herpes simplex virus protein 16 (VP16), activating transcription factor 6 α (ATF6 α), and the Ets translocation variant/polyoma enhancer activator 3 (ETV/PEA3) family of transcription factors, that bind to Med25 through their amphipathic TADs.⁸⁻¹³ Several lines of evidence support the role of Med25•activator PPIs in oncogenesis and metastasis. For example, overexpression of the ETV activators has been correlated with invasive cancer phenotypes and knockdowns of all three ETV activators have revealed the downregulation of several genes associated with cancer metastasis.¹⁴⁻¹⁹ ATF6 α is implicated in the unfolded protein response which is involved in the activation of several anti-apoptotic and pro-survival genes, like HSPA5, that work to protect the cell from apoptosis during times of cellular stress, particularly in cancerous tumors.²⁰⁻²³ These activators are all known to interact specifically with the **Activator-Interaction Domain (AcID)** motif of Med25, and ETV5 binds to the H1 face with ATF6 α binding to the H2 face of the domain. Additionally, RNA-seq experiments performed in collaboration with the Merajver Lab at University of Michigan revealed an upregulation in Med25 expression in patient-derived triple negative breast cancer (TNBC) cell line VARI068.^{24,25}

The TAD of one of these activators, VP16, contains a short peptide sequence, often referred to as the VN8 domain, that has been identified in several other acidic TADs as a

motif containing several carbonyl-containing amino acids flanked by regions of bulky, hydrophobic residues.^{26,27} The activator ATF6 α contains a peptide sequence in its TAD that has sequence homology (75%) with VN8 and interacts with Med25. This suggests that a minimal amphipathic peptide sequence can serve as a starting point for inhibitor design or discovery.

Like the TADs of transcriptional activators, amphipathic lipids are another class of compounds that are driven primarily by electrostatic and hydrophobic interactions with their protein targets.²⁶⁻²⁹ Lipids are fundamental to several key processes in a cell and have been determined to bind and alter the function of proteins. In an effort to globally map the cellular protein targets of lipids in a native system, the Cravatt lab at Scripps Research Institute identified coactivator Med25 as lipid binding protein, although the functional implications of the lipid-Med25 complex are not yet known.^{30,31} This Med25-lipid interaction suggests the use of lipids as probes to further dissect the role of Med25 in transcription. Coupled with what we know about Med25 also binding amphipathic peptides, this suggests that natural product derived lipopeptides could be a valuable tool for this molecular intervention of Med25 and its activator binding partners.

Identification of natural product lipopeptide that targets Med25•activator PPIs

Natural products have been identified as promising molecules to target PPIs, with several having been developed recently as tools to study coactivator PPIs, such as Med25 AcID and CBPKIX. One such example is lipopeptide 34913-1, isolated from a cyanobacterial strain denoted 34913, and shown by Dr. Matthew Beyersdorf to be an inhibitor of Med25-activator PPIs *in vitro* and in cells.³² This lipopeptide is composed of a

heptapeptide consisting of D-amino acids attached to a substituted lipid with unknown stereochemistry via an amide bond at the N-terminus of the peptide (Figure 2.1A). Lipopeptide 34913-1 was characterized as an effective inhibitor for Med25 AcID with a K_i of $8.8 \pm 0.6 \mu\text{M}$ against Med25 AcID•ATF6 α PPI (Figure 2.1B). Thought to mimic the transcription activation domain of Med25 binding activators, this molecule was determined to interact with the H2 face of AcID in a binding mode analogous to that of the ATF6 α TAD.³² Finally, Beyersdorf reported that treatment of HeLa cells with 34913-1 lead to dose-dependent downregulation of the ATF6 α -regulated gene HSPA5 (Figure 2.1C).

A key question that emerged from the initial studies of Lipopeptide 34913-1 by Beyersdorf was if both the lipid and the peptide portion contribute to binding of Med25. I used chemical synthesis to develop simplified analogs that could answer these questions. From these studies I identified two analogs, 34913-8 and 34913-9, that show similar potency to the natural product despite a structurally simple lipid portion. Additionally, 34913-8 selectively inhibits Med25 PPIs *in vitro* and in cells engages with full-length Med25, with accompanying changes in Med25-dependent transcription. Thus, through these studies we have identified a synthetic analog that is readily accessible and with comparable activity to the starting natural product.

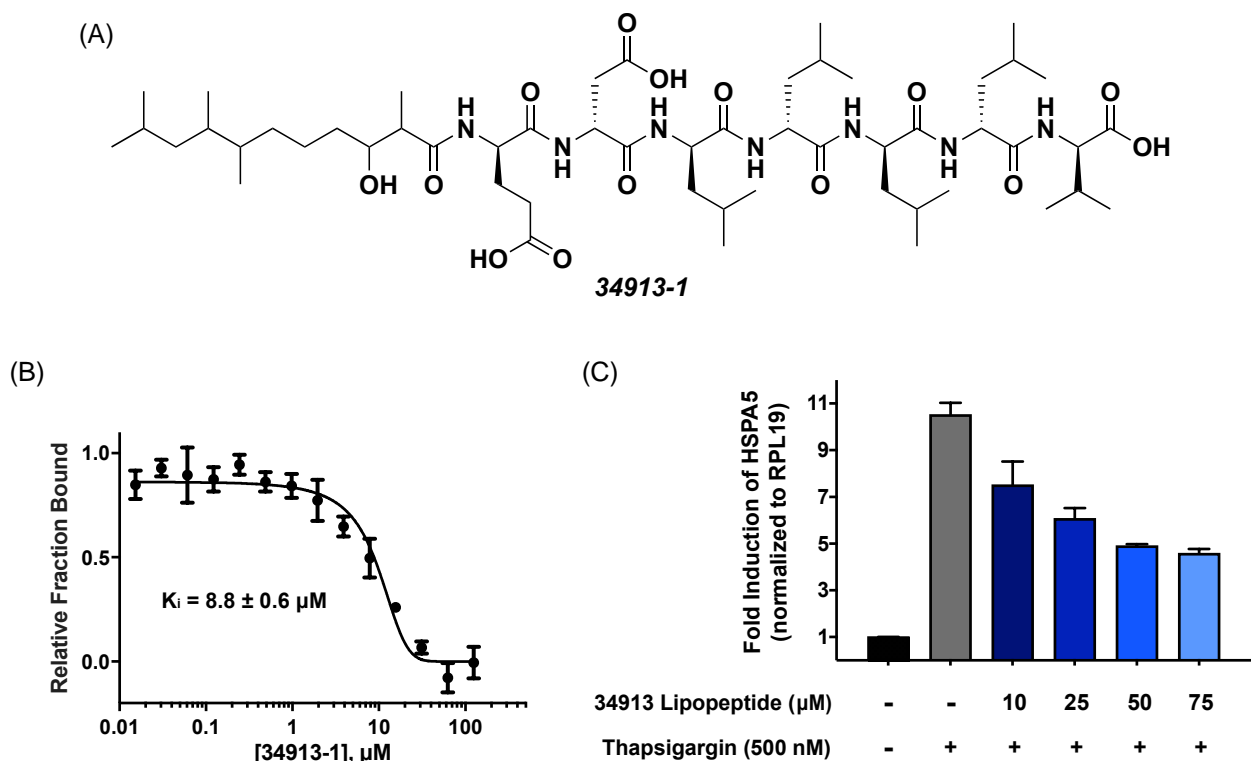


Figure 2.1. Lipopeptide 34913-1 is a potent inhibitor of Med25•ATF6 α PPI. (A) The novel lipopeptide, 34913-1, is composed of an acidic, D-amino acid heptapeptide that is thought to act as a transcriptional activation domain (TAD) mimetic. The peptide is connected to a lipophilic alkyl moiety via an amide bond at the N-terminus of the peptide that consists of an eleven-carbon chain containing five substituents of unknown stereochemistry. (B) The apparent EC_{50} values, which were determined through titrations of 34913-1 against Med25 AcID•ATF6 α were performed in triplicate with the indicated error (SDOM). (C) Analysis of HSPA5 transcript levels by qPCR indicates that when dosed with 500 nM thapsigargin, 34913-1 treatment decreases HSPA5 levels. HSPA5 transcript levels are normalized to the reference gene RPL19. Results shown are averages of technical triplicate experiments conducted in biological duplicate.

2.3 Results and Discussion

34913-1 analogs define critical components for inhibitory activity

While novel lipopeptide 34913-1 shows potent micromolar activity, with nearly 3-fold selectivity for Med25 AcID PPIs over CBP KIX PPIs, and good activity in a cellular context, additional production and isolation from cyanobacteria has not been possible. However, it was hypothesized that analogs could be chemically synthesized to assess the importance of the structurally complex lipid tail, the amino acid stereochemistry, and the carboxylic acid at the C-terminus (Figure 2.2). As described in subsequent sections,

all analogs were first assessed by in vitro binding inhibition experiments to benchmark their potency relative to the starting natural product.

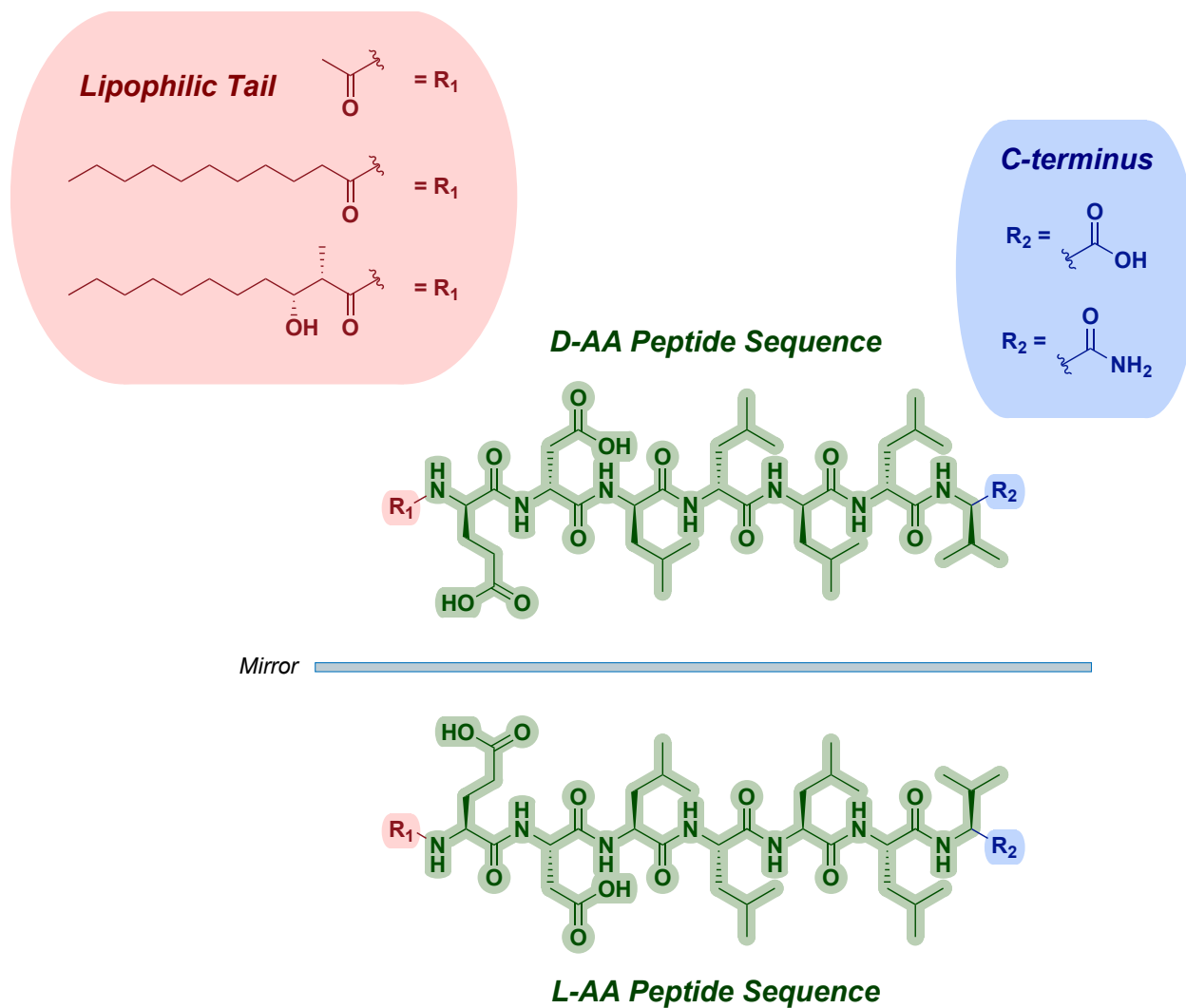


Figure 2.2. Lipopeptide 34913-1 contains three distinct components that can be modified. 34913-1 can be broken down into three more simpler components that can be dissected to determine the importance of each portions for its inhibitory activity against Med25 PPIs. The three modifiable components are: (1) the lipophilic tail (red), the peptide sequence (green), and the C-terminus (blue). The modifications that were tested are shown above.

In 34913-1, all seven amino acids in the peptide sequence are D-amino acids instead of the more commonly observed L-amino acids. To determine the importance of the amino acid stereochemistry, two N-terminally acetylated peptides with the same

sequence as the parent lipopeptide (N'-EDLLLLV) were synthesized, one composed solely of D-amino acids (34913-2) and the other composed of only L-amino acids (34913-3). Binding studies demonstrated that both the D-amino acid peptide sequence and the L-amino acid peptide sequence inhibited the Med25•ATF6 α PPI with comparable binding affinities. This indicates that stereochemistry of the amino acid sequence may not play a significant role in the binding specificity of these lipopeptides for Med25 AcID, though there is significant literature evidence suggesting that D-peptides are advantageous due to their resistance to proteolysis.³³ These results are consistent with studies of peptides derived from naturally occurring transcriptional activators, in which D- and L- sequences are functionally equivalent.

The TADs of transcriptional activators often have a net negative charge and this is true of those that interact with Med25. The parent lipopeptide 34913-1 has a net -3 charge stemming from two acidic residues and a carboxylic acid at the C-terminus. Based on what is known about acidic residues in TADs, it would be expected that the negative charge of the C-terminal carboxylic acid contributes to the inhibitory activity of these lipopeptides for the Med25•ATF6 α PPI. This impact on the recognition and binding of the lipopeptide to Med25 AcID was evaluated using two additional analogs, one containing the D-peptide sequence (N'-EDLLLLV) with a C-terminal amide (34913-4) and the other containing a L-peptide sequence (N'-EDLLLLV) with a C-terminal amide (34913-5) (Figure 2.3A). The inhibition assays determined that both 34913-4 and 34913-5 containing the C-terminal amide showed improved binding activity when compared to that of 34913-2 and 34913-3, initially indicating that the C-terminal negative charge does not contribute significantly. Additionally, the weakened inhibitory activity of these four

simplified analogs (34913-2, 34913-3, 34913-4, 34913-5) for the Med25•ATF6 α PPI suggest that the lipophilic tail portion of 34913-1 is critical for binding.

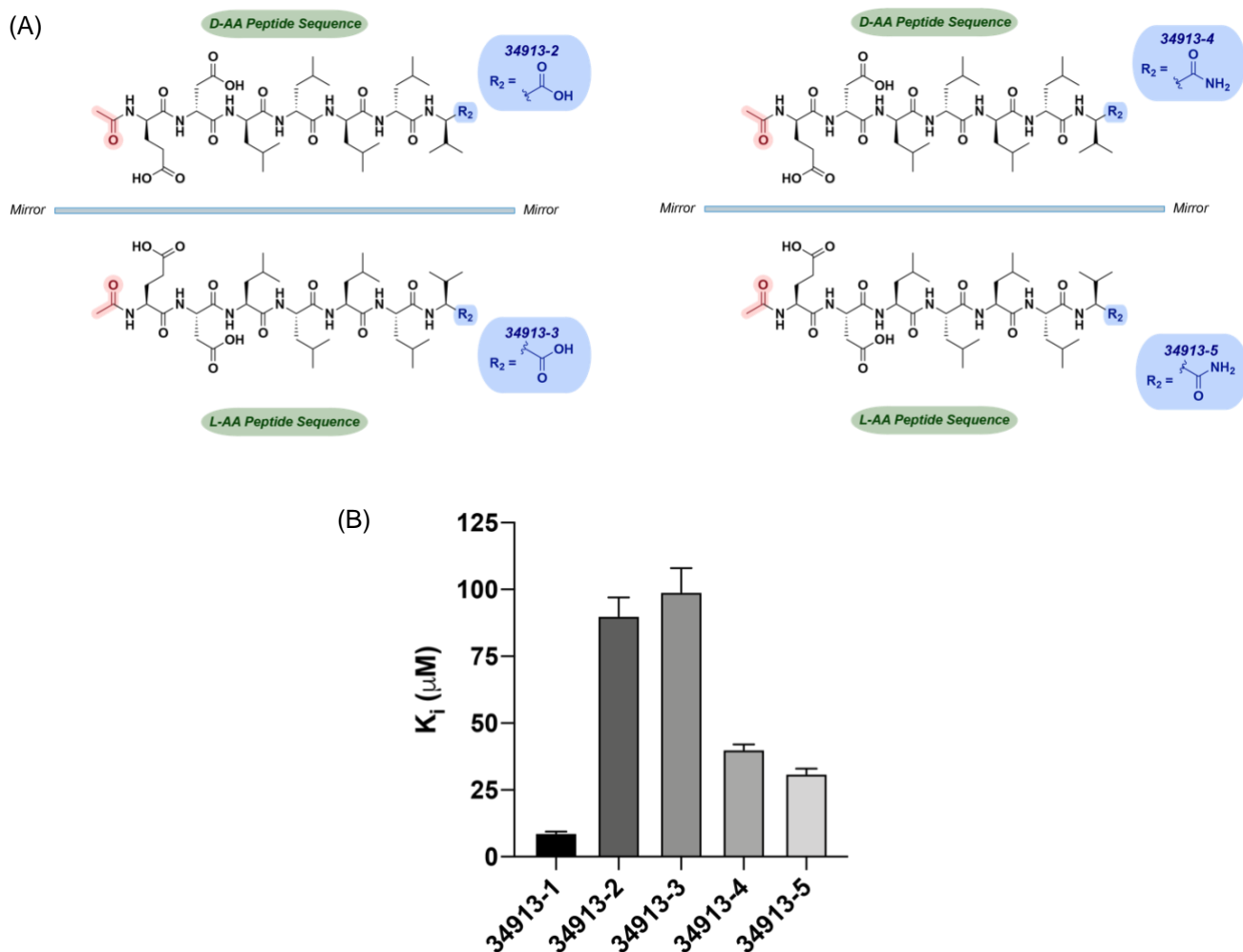


Figure 2.3. Peptides not sufficient for inhibitory activity against Med25•ATF6 α . (A) The following peptide sequence (N'-Glu-Asp-Leu-Leu-Leu-Val) analogs were synthesized with varying C-terminal functional groups and amino acid stereochemistry: D-peptide with C-terminal carboxylic acid (34913-2; top right), L-peptide with C-terminal carboxylic acid (34913-3; top bottom), D-peptide with C-terminal amide (34913-4; top left), and L-peptide with C-terminal amide (34913-5; bottom left). (B) Apparent EC₅₀ values were determined through titration of the lipopeptide analogs for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM). The EC₅₀ values were converted to K_i values using the apparent K_d value based on the direct binding of Med25 AcID•ATF6 α PPI using a K_i calculator³⁴ with the indicated error (SDOM). Results shown represent the average K_is of two biological duplicates.

Upon determination of the importance of the lipophilic tail, it was hypothesized that the branching at the proximal and distal ends of 34913-1 may not play a significant role in the inhibition of Med25•ATF6 α (Figure 2.3). Attachment of a simplified lipophilic tail was initially investigated by coupling undecanoic acid to the N-terminus of both the D-amino acid peptide sequence (34913-6) and the L-amino acid peptide sequence (34913-7). Improved inhibitory activity for the Med25 AcID•ATF6 α PPI was observed for both 34913-6 and 34913-7 compared to the parent lipopeptide 34913-1 (Figure 2.4).

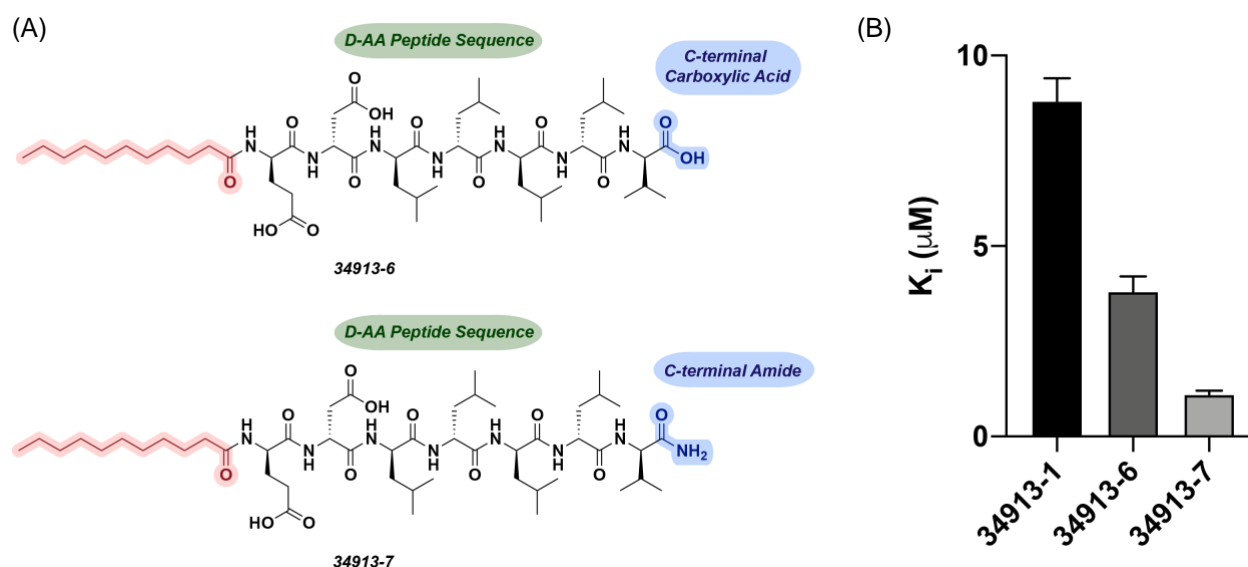


Figure 2.4. Simplified lipid tail increases inhibitory activity for Med25•ATF6 α . (A) Two additional lipopeptides with synthesized composed of D-amino acids and a simplified lipophilic tail, undecanoic acid. The difference between these two analogs was the C-terminus, with one containing a C-terminal carboxylic acid (34913-6; top) and the other containing a C-terminal amide (34913-7; bottom). (B) Apparent EC_{50} values were determined through titration of the lipopeptide analogs for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM) The EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of Med25 AcID•ATF6 α PPI using a K_i calculator.³⁴ Results shown represent the average K_i s of two biological duplicates

Visible precipitation of 34913-6 and 34913-7 from solution suggested that the increase in potency of these analogs may be due to aggregation, tested in a detergent screen using increasing percentages of NP40. This data demonstrated that the increased potency of these two analogs containing undecanoic acid were the result of aggregation,

indicating that further branching of the lipophilic tail may be necessary for improved solubility of further analogs (Figure 2.5).

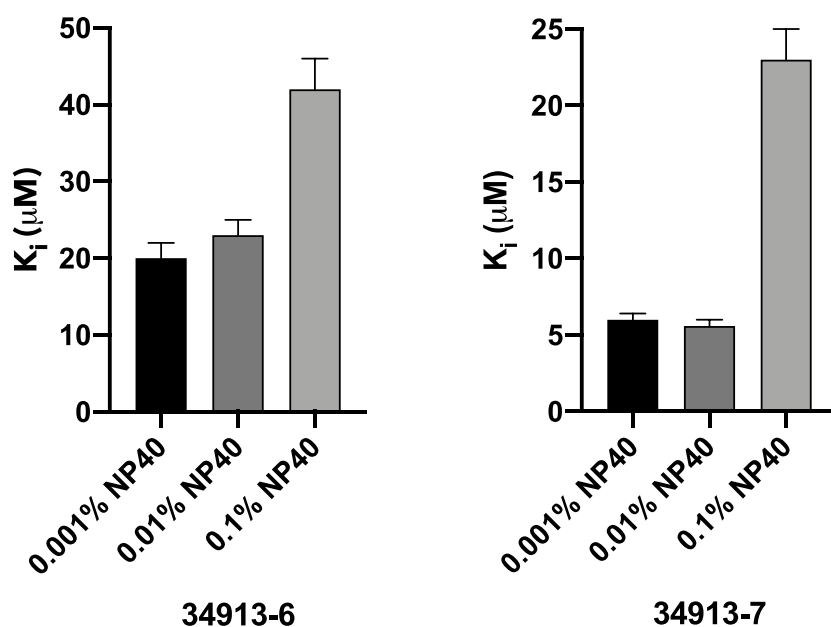


Figure 2.5. NP40 detergent determines aggregation contributes to potency. The addition of detergent results in an increase in inhibitory activity of 34913-6 (left) and 34913-7 (right) for Med25 AcID•ATF6 α , indicating that aggregation may be a contributing factor. Apparent EC₅₀ values were determined through titration of the lipopeptide analogs for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM). The EC₅₀ values were converted to K_i values using the apparent K_d value based on the direct binding of Med25 AcID•ATF6 α PPI using a K_i calculator.³⁴ Results shown represent the average K_is of two biological duplicates.

Substitution in the lipid tail leads to improved activity

Due to visible aggregation of the undecanoic acid-based simplified analogs, a substituted lipophilic tail was designed and synthesized. The stereochemistry of the lipophilic tail of 34913-1 is undefined, as the small quantity of the originally isolated natural product was insufficient for full structural characterization. Noting the presence of a secondary alcohol and a methyl group proximal to the C-terminus, the synthesis of an analog containing these moieties was pursued using standard aldol approaches. As shown in Figure 2.6, one enantiomer was prepared for these studies using a previously

published procedure (Figure 2.6). This 2-step synthesis resulted in a crude product, (2S, 3R)-3-hydroxy-2-methylundecanoic acid, that was coupled to the N-terminus of the D-amino acid peptide sequence (EDLLLLV), resulting in 34913-8 and 34913-9. These two lead analogs vary solely in the C-terminus, with one analog containing a C-terminal carboxylic acid (34913-8) and the other containing a C-terminal amide (34913-9) (Figure 2.7A).

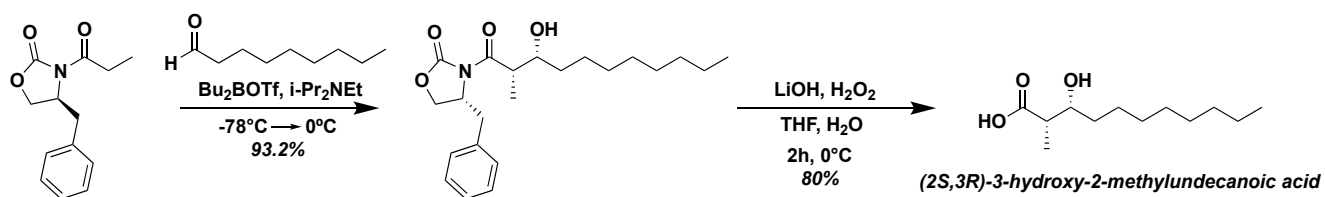


Figure 2.6. Synthesis of (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Shown here is the two step synthesis of the lipophilic tail (2S,3R)-3-hydroxy-2-methylundecanoic acid.

Initial analysis of these two analogs determined that 34913-8 and 34913-9 inhibits the Med25•ATF6 α PPI with apparent K_i values of $10 \pm 1 \mu\text{M}$ and $6.6 \pm 0.3 \mu\text{M}$, respectively, demonstrating comparable activity to 34913-1 (Figure 2.7B). Though there is no observed insolubility of these compounds in solution, the possible effects of aggregation were investigated by using increasing percentages of NP40 detergent solution. These results indicate that aggregation is not contributing to the activity of 34913-8, though there is a slight effect on the activity of 34913-9 (Figure 2.8). Taken together, this data suggests the addition of the substituents in the proximal region of the lipid tail was sufficient for the replication of the inhibitory activity of 34913-1 for Med25•ATF6 α *in vitro*. Future studies will investigate the importance of stereochemistry for inhibition with the synthesis of the enantiomer and two diastereomers of (2S, 3R)-3-hydroxy-2-methylundecanoic acid.

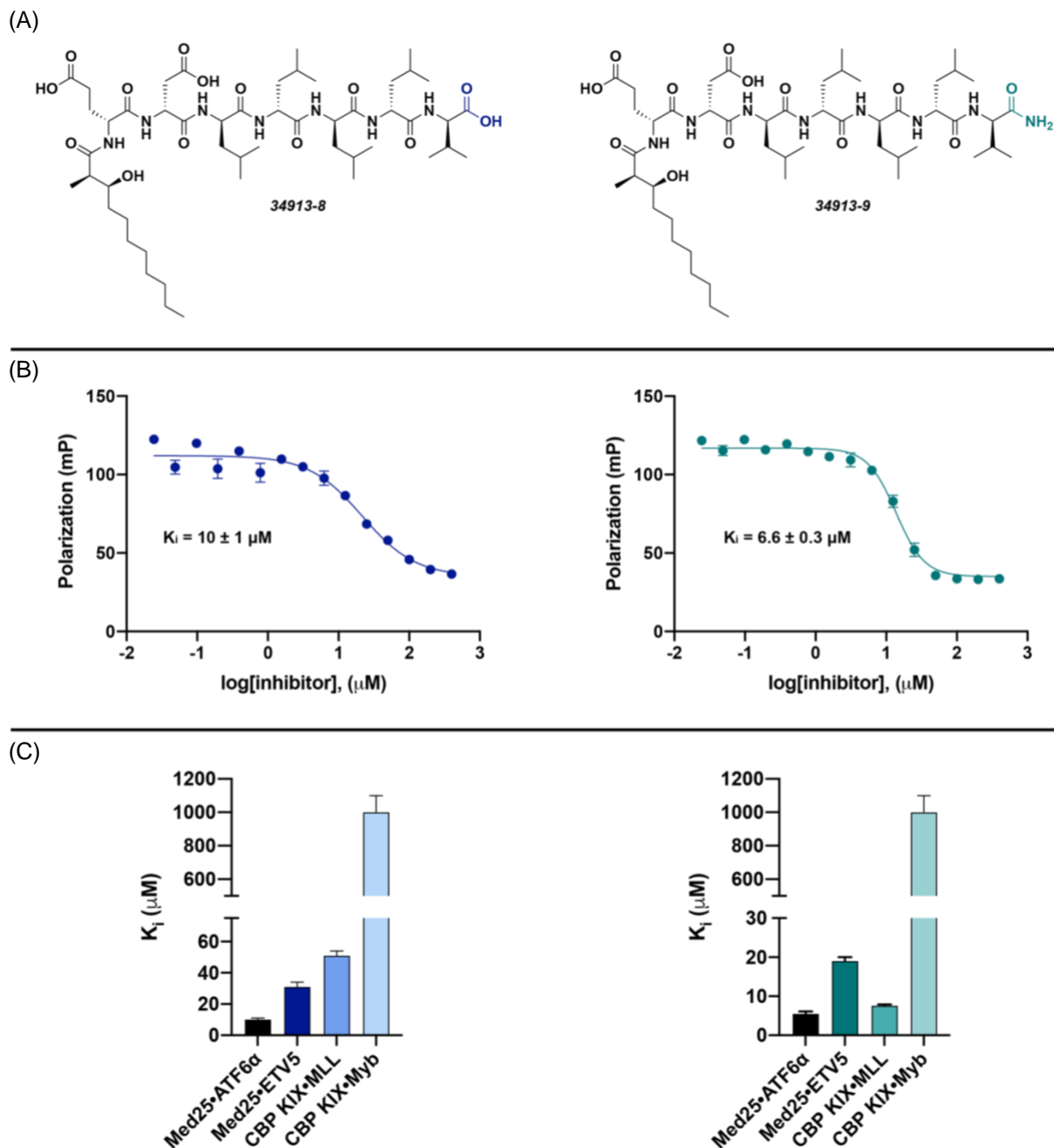


Figure 2.7. Analog 34913-8 is potent and selective inhibitor for Med25 PPIs. (A) Two lead analogs were synthesized containing a substituted lipophilic tail, (2S, 3R)-3-hydroxy-2-methylundecanoic acid, attached to the N-terminus of the peptide sequence. The difference lies solely in the C-terminus, with one analog containing a C-terminal carboxylic acid (34913-8) highlighted in blue and the other containing a C-terminal amide (34913-9) highlighted in red. (B) Apparent EC_{50} values were determined through titration of 34913-8 and 34913-9 for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM). (C) Apparent EC_{50} values were determined through titration of 34913-8 and 34913-9 analog for Med25 AcID•ATF6 α , Med25 AcID•ERM, CBP KIX•MLL, and CBP KIX•Myb performed in experimental triplicate with the indicated error (SDOM). The EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of Med25 AcID•ATF6 α , Med25 AcID•ERM, CBP KIX•MLL, and CBP KIX•Myb using a K_i calculator.³⁴

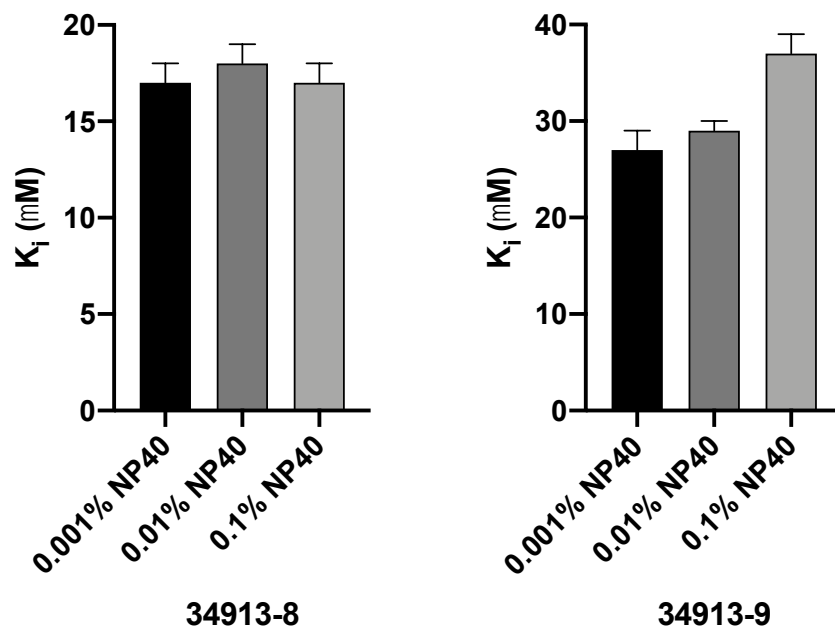


Figure 2.8. NP40 detergent determines aggregation doesn't contribute to potency. The addition of NP40 detergent results in relatively consistent inhibitory activity of 34913-8 (left) or 34913-9 (left) for Med25 AcID•ATF6 α , indicating there is not aggregation. Apparent EC_{50} values were determined through titration of the lipopeptide analogs for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM). The EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of Med25 AcID•ATF6 α PPI using a K_i calculator.³⁴

Biophysical analysis was performed to assess the binding location of analogs 34913-8 and 34913-9 on Med25 AcID. Based on the potency of these analogs for the Med25•ATF6 α PPI and previous data using the lead lipopeptide 34913-1, it was hypothesized that these analogs would bind to the H2 face of Med25 AcID, acting as an orthosteric inhibitor of the Med25 AcID•ATF6 α PPI.^{13,25} The interaction of these analogs with both surface and buried residues on the surface of Med25 AcID was determined by performing 1H , ^{13}C - and 1H , ^{15}N -heteronuclear single quantum coherence (HSQC) NMR experiments in which 34913-8 and 34913-9 were complexed with ^{15}N , ^{13}C -labeled Med25 AcID at 0, 0.2, 0.5, 0.8, 1.1, 2, and 3 equivalents of lipopeptide relative to protein. Peaks in each of the collected spectra were assigned to specific residues using previously published NMR assignments from our lab^{8,9} and the chemical shift perturbations (CSP)

were calculated with a CSP ≥ 0.02 ppm considered to be significant and above the signal-to-noise (S/N) ratio.

Analysis of Med25 AcID ^1H , ^{13}C -HSQC spectra of 34913-8 bound to Med25 AcID revealed a significant shift perturbation of residue L525 on the β -barrel of the H2 face of Med25 AcID, while observing no statistically significant perturbations to any residues on the H1 face at 0.8 equivalents (Figure 2.9A). Further saturation with 1.1 equivalents of 34913-8 revealed additional residues that were perturbed, including residue L514 on the H2 face and L406 on the H1 face (Figure 2.9B). After complete saturation with 3 equivalents of 34913-8, several significant changes in CSP patterns were observed on the H2 face, while only two residues, L406 and L448, was perturbed to varying degrees on the H1 binding face of Med25 AcID (Figure 2.9C). This suggests that 34913-8 preferential binds to the H2 face of Med25 AcID, with secondary binding at the H1 face occurring at higher concentrations of compound.

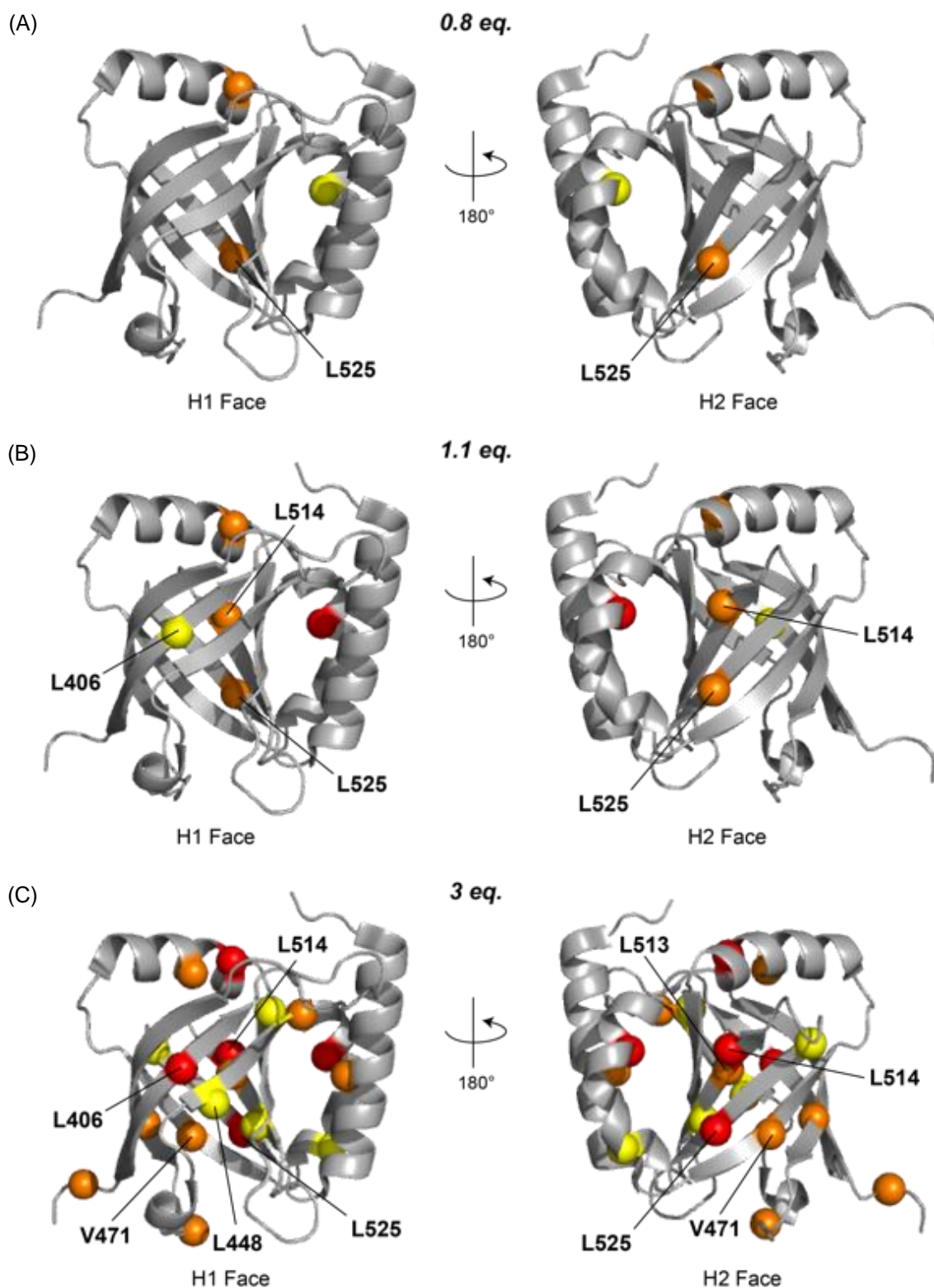


Figure 2.9. ^1H , ^{13}C CSPs map induced by binding of 34913-8. CSPs from dosing with (A) 0.8 eq (B) 1.1 eq (C) 3 eq of 34913-8 were plotted onto the structure Med25 AcID (PDB ID 2XNF). Yellow = 0.02 ppm – 0.0249 ppm, orange = 0.025 ppm – 0.049 ppm, red \geq 0.0491.

Additional ^1H , ^{15}N -HSQC experiments were performed to map the conformational changes observed when 34913-8 interacts with Med25 AcID. Analysis of the Med25 AcID ^1H , ^{15}N HSQC spectra at 0.8 equivalents of 34913-8 in complex with Med25 showed numerous residues on the H2 face of Med25 were perturbed, with three residues (S516, K518, and G524) having significant CSPs of over 0.0851 ppm. Analysis of the H1 face revealed that there were no residues in the β -barrel that were perturbed though there were several residues weakly perturbed (yellow = 0.021 – 0.085 ppm) in the flanking flexible regions surrounding the β -barrel, suggesting that the binding of 34913-8 to Med25 may cause allosteric shifts on the H1 face of the protein (Figure 2.10A). This was consistent when further saturation with 1.1 equivalents of 34913-8, exhibited additional residues on the H2 face that underwent significant chemical shift perturbations, while the H1 face did not (Figure 2.10B). Finally, 3 equivalents of 34913-8 in complex with Med25 revealed low CSP (0.021 – 0.085 ppm) of several residues in the H1 face, including W408, S426, L448, I449, and M500 (Figure 2.10C).

Based on this ^1H , ^{13}C -, ^1H , ^{15}N -HSQC data, 34913-8 binds to the H2 face of Med25 at low equivalents of compound, resulting in significant conformational changes to the residues on the β -barrel and flanking α -helices on the H2 face. Taken together, this supports the hypothesis that 34913-8 is an orthosteric inhibitor of Med25 AcID•ATF6 α PPI. Additionally, the ^1H , ^{15}N -HSQC spectra reveal weak conformation changes in the flexible regions surrounding the H1 face, which suggest allosteric inhibition of the H1 face with 34913-8.^{8,9} According to the ^1H , ^{13}C -HSQC spectra, at saturating concentrations of 34913-8 we begin to see significant perturbations on the β -barrel of the H1 face. As this binding occurs, several residues on the β -barrel of the H1 face in the ^1H , ^{15}N -HSQC

spectra begin to shift suggesting that at saturated concentrations, 34913-8 directly binds to the H1 face, orthosterically inhibiting Med25 AcID•ETV5.

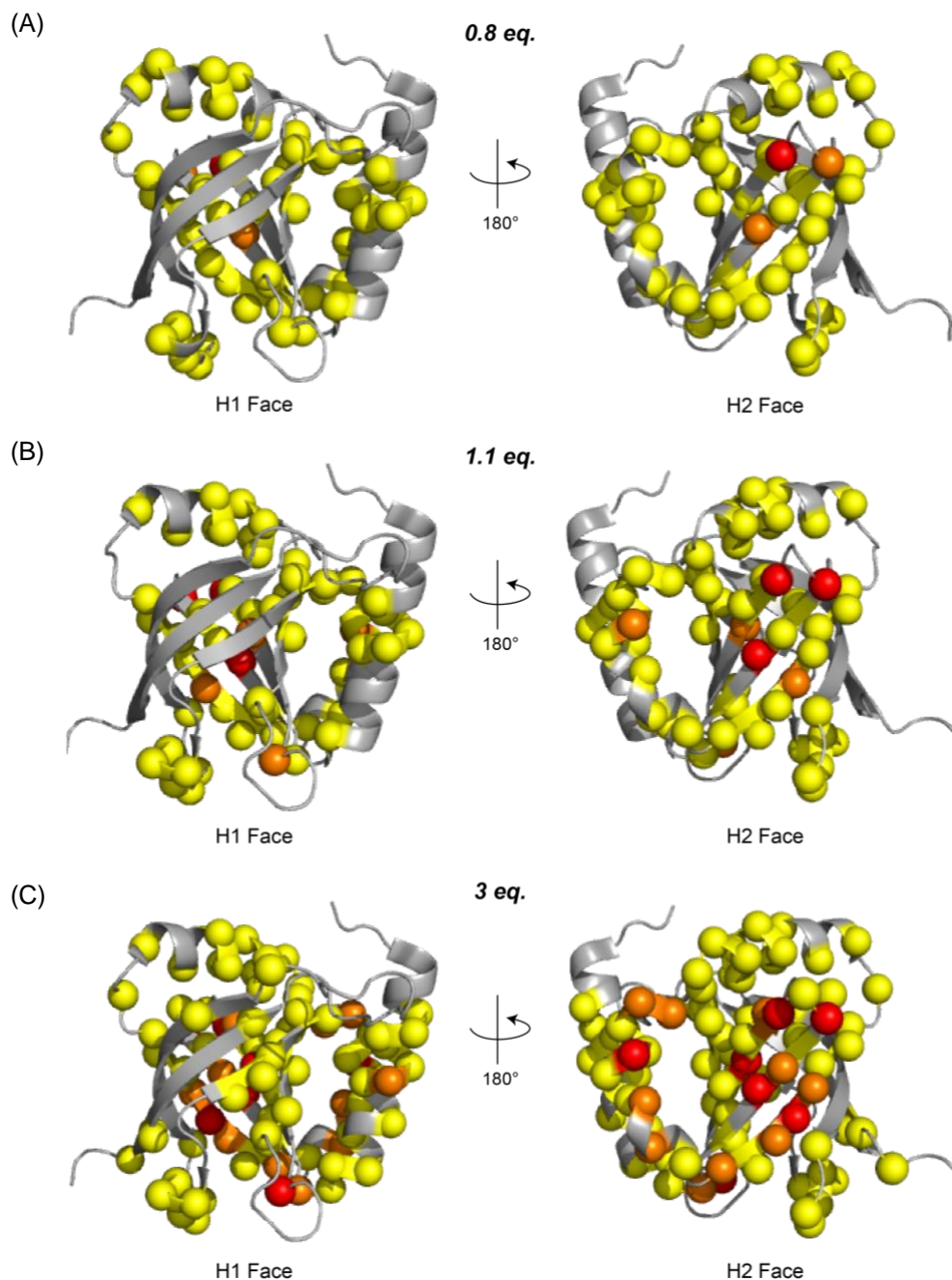


Figure 2.10. ^1H , ^{15}N CSPs map induced by binding of 34913-8. CSPs from dosing with (A) 0.8 eq (B) 1.1 eq (C) 3 eq of 34913-8 were plotted onto the structure Med25 AcID (PDB ID 2XNF). Yellow = 0.02 ppm – 0.085 ppm, orange = 0.0851 ppm – 0.14 ppm, red \geq 0.141.

In comparison to 34913-8, the ^1H , ^{13}C and ^1H , ^{15}N -HSQC spectra of 34913-9 bound to Med25 AcID differed drastically. ^1H , ^{13}C -HSQC analysis of a titration of 34913-9 bound to Med25 revealed that 3 equivalents were needed to see perturbation of Med25 AcID, with only one residue (L458) perturbed above the S/N ratio (Figure 2.11A, yellow = 0.02 – 0.0249 ppm). Further analysis using ^1H , ^{15}N -HSQC demonstrates that there are several residues perturbed at both 0.8 eq and 3 eq of 34913-9, though these perturbations are weak (Figure 2.11B, yellow = 0.02 – 0.085 ppm). A considerable amount of these perturbations occur on the β -barrel of the H2 face and the flanking α -helix, $\alpha 1$, with the residues on the H1 face of Med25 seeing minimal CSPs. This HSQC data initially suggests that 34913-9 is unable to bind to Med25 AcID, though previous competition fluorescence polarization assays have demonstrated that 34913-9 is able to inhibit Med25 AcID with similar potency to 34913-8. Further experiments will be performed to explore this hypothesis further. These results are interesting as they demonstrate the importance of the C-terminal carboxylic acid for specific interactions with Med25 AcID, indicating potential for the synthesis and characterization of further lipopeptides analogs with optimized specificity for Med25 AcID.

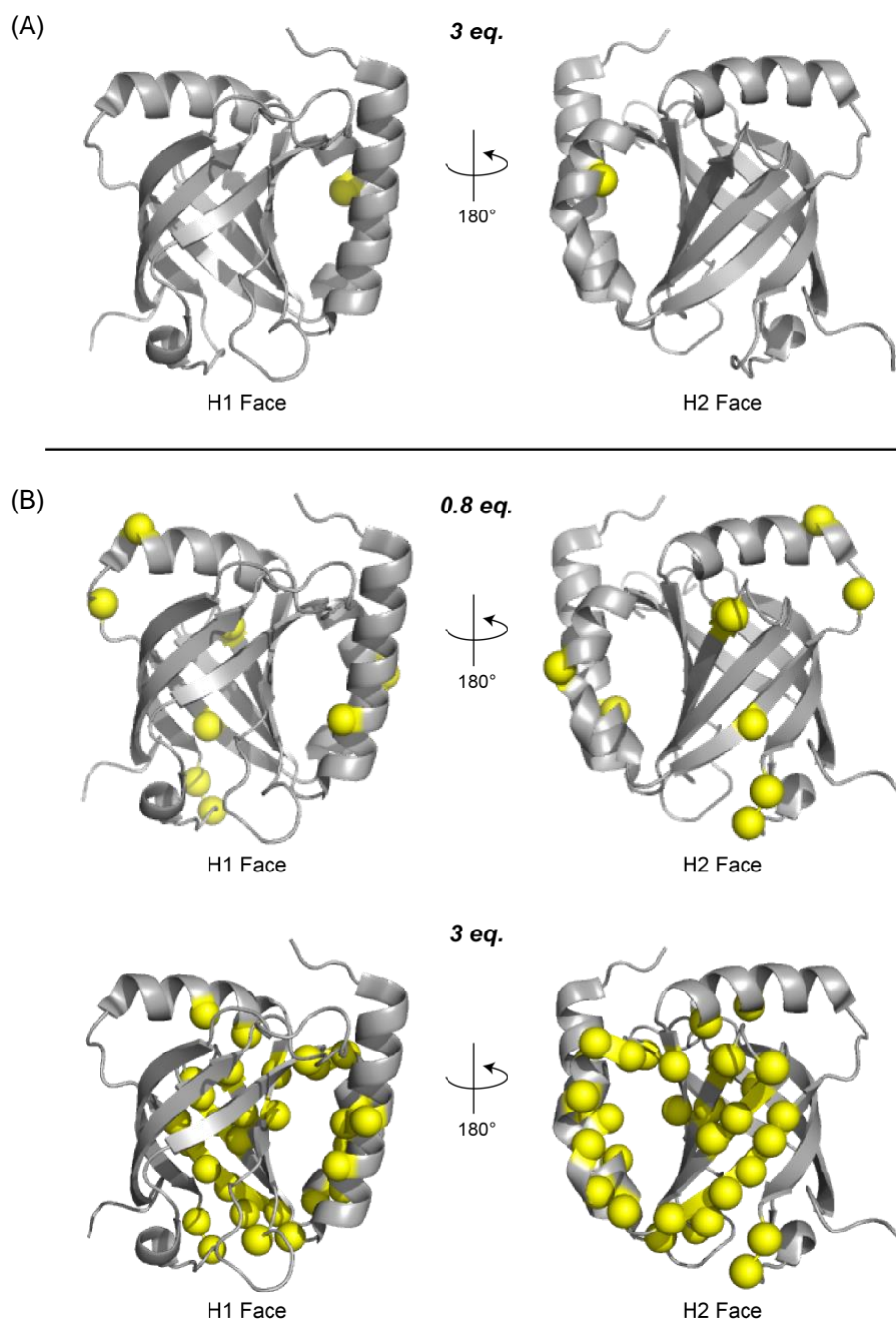


Figure 2.11. ^1H , ^{13}C and ^1H , ^{15}N CSPs map induced by binding of 34913-9. (A) ^1H , ^{13}C spectra CSPs with 3 eq. 34913-8 plotted onto the structure of Med25. Yellow = 0.02 ppm – 0.0249 ppm, orange = 0.025 ppm – 0.049 ppm, red \geq 0.0491. (B) ^1H , ^{15}N spectra CSPs with 0.8 eq. (top) and 3 eq. (bottom) of 34913-9 plotted onto the structure of Med25. Yellow = 0.02 ppm – 0.085 ppm, orange = 0.0851 ppm – 0.14 ppm, red \geq 0.141. Med25 AcID (PDB ID 2XNF).

Based on the HSQC-NMR data, it is hypothesized that 34913-8 binds specifically to the H2 face of Med25 AcID and causes conformational changes on both faces, possibly affecting the binding of the activator binding partner ETV5. To investigate further, competition fluorescence polarization assays were performed to determine the inhibitory activity of 34913-8 and 34913-9 against the Med25•ETV5 PPI. Data demonstrated that both 34913-8 and 34913-9 show 3-fold selectivity for Med25•ATF6 α compared to Med25•ETV5 (Figure 2.7C), further supporting the HSQC-NMR data that 34913-8 orthosterically inhibits Med25•ATF6 α and allosterically inhibits Med25•ETV5. Though the HSQC-NMR determined that 34913-9 does not bind to Med25 AcID, this compound has effectively inhibited both Med25•ATF6 α and Med25•ETV5 using fluorescence polarization. Further experiments will need to be performed to draw more conclusive results.

Lipopeptide analog 34913-8 demonstrates selectivity for Med25 PPIs

Med25 is an especially unique coactivator as it contains a novel structural fold, consisting of a central, seven-stranded β -barrel, that varies from typical coactivators, such as CBPKIX, which are comprised primarily of α -helices. Selectivity of 34913-8 and 34913-9 for CBPKIX and two of its activator binding partners, MLL and Myb, were determined to evaluate the significance of this novel structural fold. Both 34913-8 and 34913-9 showed minimal inhibitory activity against the CBPKIX•Myb PPI, while 34913-8 showed 5-fold selective inhibition for Med25•ATF6 α compared to CBPKIX•MLL (Figure 2.6C). Interestingly, 34913-9 showed virtually no selectivity for Med25•ATF6 α over CBPKIX•MLL. This demonstrates that 34913-8 exhibits selectivity for Med25 PPIs over

other coactivators with similar binding surfaces while 34913-9 does not. It is also important to note that both 34913-8 and 34913-9 inhibit Med25 PPIs at both binding surfaces, including those involved in the binding of transcriptional activators ETV5 (H1 binding surface) and ATF6 α (H2 binding surface).

Lipopeptide 34913-8 engages with Med25 in a cellular context

To investigate the target engagement of these lead analogs with full-length Med25, thermal shifts assays were performed using prepared nuclear extracts from a patient-derived triple-negative breast cancer (TNBC) cell line, VARI068.^{24,25} Results show increased band density when dosed with 34913-8 compared to the control samples, indicating thermal stabilization and target engagement with endogenous Med25 (Figure 2.12A). Analog 34913-9, on the other hand, did not demonstrate this same stabilization, as demonstrated by the decrease in band density comparable to the control sample, indicating a lack of engagement with full-length Med25 (Figure 2.12B). This corroborates previous data suggesting that the C-terminal carboxylic acid of 34913-8 may play a key role in the molecular recognition to Med25 AcID, resulting in a more selective inhibitor.

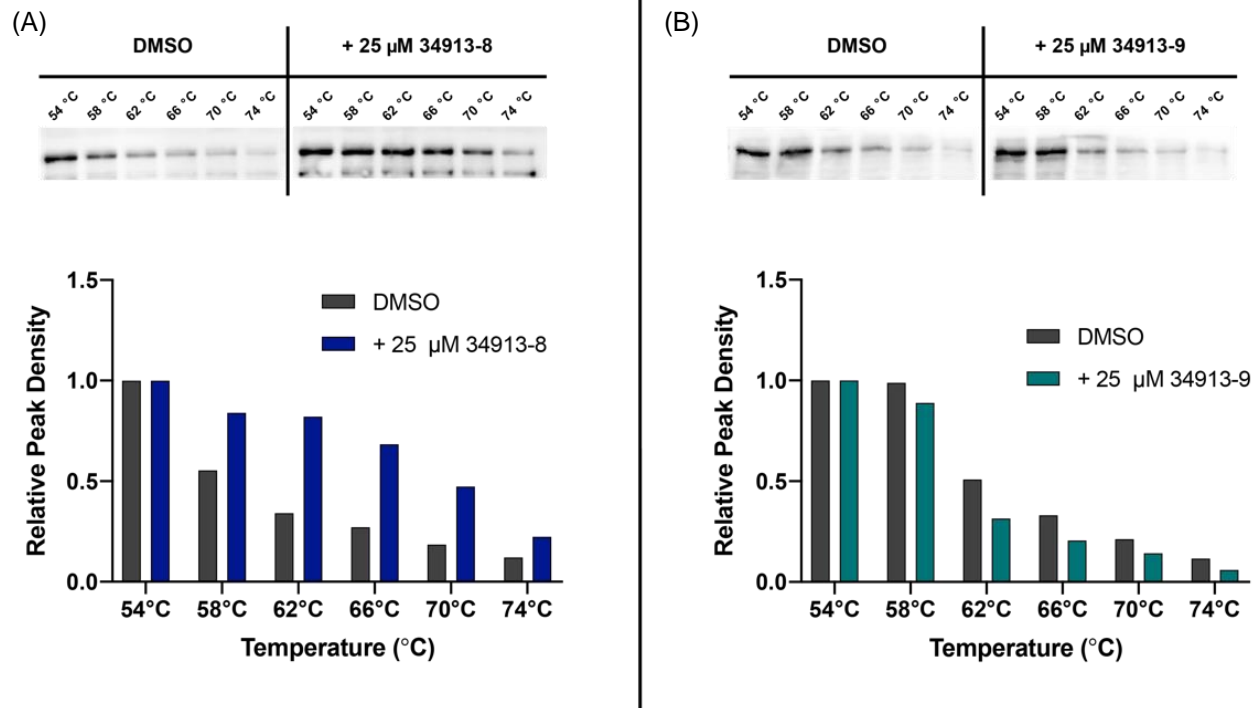


Figure 2.12. Analog 34913-8 stabilizes full length Med25 in VARI068 cell extracts. Cellular thermal shift assays (CETSA) were performed by dosing VARI068^{24,25} cell nuclear extracts with 25 μM 34913-8, 34913-9, or equivalent DMSO and subjecting the samples to a range of temperatures. Western blots were performed using a Med25 antibody. (A) 34913-8 stabilizes full length Med25 using mammalian cell nuclear extracts. It was determined that 34913-8 engages with Med25, stabilizing it as the temperature increases. (B) Data indicated that as the temperature of the samples increases, the Med25 band decreases in a similar manner to the DMSO control suggesting that 34913-9 did not appear to engage with Med25 using mammalian cell nuclear extracts. Data in the bar graph is normalized to the DMSO control (grey) that is equal to 1. These experiments have been repeated in duplicate and this data is representative.

Following the successful determination of 34913-8 engagement of full length Med25, the effects of inhibiting Med25 PPIs regulation in cancer were explored using both 34913-8 and 34913-9. Due to the implication of the Med25•ETV/PEA3 PPIs in cell proliferation, migration, and invasion pathways that is frequently associated with carcinogenesis, the gene transcripts of the Med25•ETV5-regulated MMP2 were analyzed using qPCR. It was determined that the treatment of VARI068 cells with increasing concentrations of 34913-8 resulted in the downregulation of MMP2 gene expression (Figure 2.13A), while treatment with increasing concentrations of 34913-9 did not cause significant change (Figure 2.13B). Collectively, this data is consistent with the

understanding that 34913-8 engages Med25 in cells and alters its PPI network resulting in downstream effects on gene expression. 34913-9 is unable to engage Med25 in cells due to the C-terminal amide resulting in minimal disruption of its PPIS and their effect on downstream gene expression.

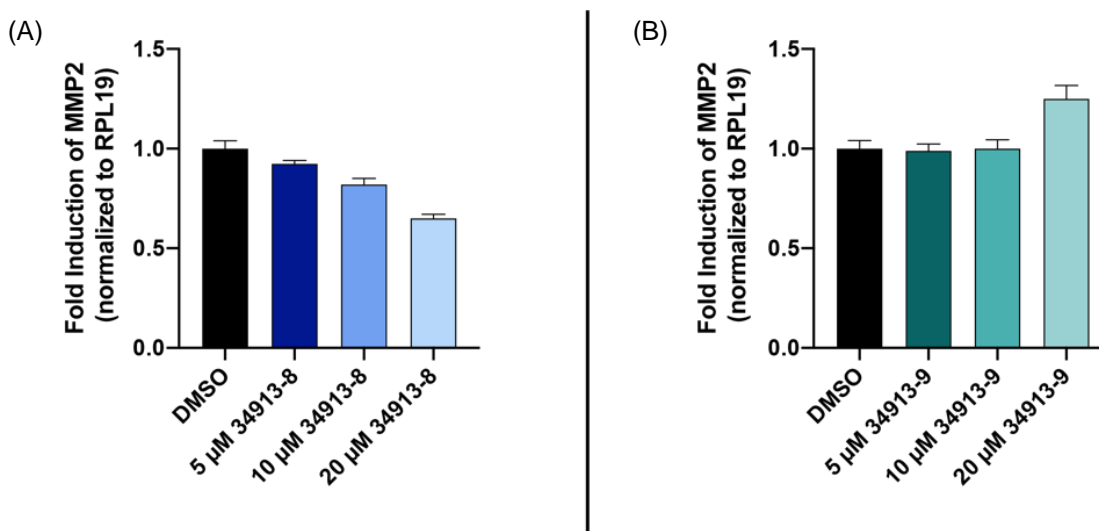


Figure 2.13. Analog 34913-8 shows inhibition of Med25 in a cellular context. (A) Analysis and quantification of the MMP2 transcript levels by qPCR in VARI068 cells^{24,25} indicates that treatment with increasing concentrations of 34913-8 results in the decrease of MMP2 levels. (B) Analysis and quantification of the MMP2 transcript levels by qPCR indicates that treatment with increasing concentrations of 34913-9 does not see much difference in the MMP2 levels. MMP2 transcript levels are normalized to the reference gene RPL19. Results shown are the average of technical triplicate experiments.

2.4 Conclusions and Future Directions

The overarching hypothesis of this work was to develop and characterize a potent and/or selective inhibitor of Med25, resulting in promising probes for the dissection of Med25•activator PPIs. Both the ETV/PEA3 and ATF6 α activators have significant impacts on several cellular processes involved in human disease making them excellent therapeutic targets for molecular intervention. Lipopeptides act as a unique class of natural products that frequently contain amphipathic peptide sequences connected to lipophilic tail that can vary in length and composition of substituents. Med25 was an

interesting potential target for lipopeptides due to several lines of evidence revealing that lipids and amphipathic TADs bind to the protein. The data presented in this chapter demonstrates the potential for developing an inhibitor to better understand Med25's key role in the transcription of various disease related genes and phenotypes.

In this chapter, a promising lipopeptide natural product was investigated as an effective inhibitor of Med25 and its PPIs. Due to this compound's modular structure, each component was studied to investigate the contribution it had on its inhibitory activity for Med25•activator PPIs. It was determined that the presence of both peptide sequence and a substituted lipid tail were important for inhibition. Dissection of the structure and activity of this compound lead to the discovery of two lead analogs which differed solely in the C-terminus. Several experiments determined that the C-terminal carboxylic acid (34913-8) offered selectivity for the AcID domain using competition fluorescence polarization assays, HSQC NMR, and engagement with full-length Med25 using mammalian cells. Treatment of mammalian cells with 34913-8 also demonstrated downregulation of the gene MMP2, which is regulated by the Med25•ETV5 PPI.

The lipopeptide analog, 34913-8, represents the first potent, selective, and noncovalent small molecule that targets Med25 *in vitro* and *in cellulo*. This is also the first lipopeptide to target a coactivator•activator PPI, signifying potential for the use of lipopeptides as promising probes for the dissection of Med25•activator PPIs as well as other coactivator•activator PPIs. For example, there are limited gene targets that have been identified to study these Med25•activator PPIs and their impact on disease. Moving forward, these lead analogs, 34913-8 and 34913-9, can be used as probes to disrupt Med25 PPIs, resulting in the identification of previously unknown Med25 PPI dependent

genes using RNA-seq experiments. Mammalian cells treated with/without a specific (34913-8) and non-specific (34913-9) probe should result in differential expression profiles because of Med25 PPI inhibition. The gene targets identified will be used to highlight cellular processes affected by Med25, representing a broader sense of its role in human disease.

2.5 Materials and Methods

Reagents and Instrumentation

Unless otherwise noted, chemical and biological reagents were obtained from commercial sources and were used without additional modification. Protein, peptide, and DNA concentrations were determined using a NanoDrop ND-1000 UV-Vis Spectrophotometer and/or a Beckman Spectrophotometer.

Plasmids for Protein Expression

Prof. Patrick Cramer generously provided the Med25 expression plasmid pET21b-Med25(394-543)-His6. Plasmid sequence identity was confirmed via standard Sanger sequencing methods on an Applied Biosystems 3730xl DNA Analyzer at the University of Michigan DNA Sequencing Core and analyzed using SeqMan Pro from the Lasergene DNASTAR software suite.

Expression of Med25 AcID/13C, 15N labeled Med25 AcID

Med25 AcID and was expressed as follows. Plasmids were transformed into chemically competent *E. coli* BI21 (DE3) cells, plated onto LB/ampicillin agar, and

incubated at 37°C overnight. The next day, plates were stored at 4°C until further use. In the evening, a single colony from the plate was selected and placed in 50 mL of LB with 0.1 mg/mL ampicillin (Gold Bio Technology) and incubated at 37°C at 250 rpm. The following morning, 5 mL was removed from the starter culture and added to 1 L TB Broth (24 g yeast extract, 12 g tryptone, 4 mL glycerol, 100 mL 0.17 M KH₂PO₄/0.72 M K₂HPO₄, 900 mL water) with 0.1 mg/mL ampicillin and was grown to an OD₆₀₀ of 0.6-0.8 at 37°C, 250 rpm. Once OD₆₀₀ was reached, temperature was reduced to 20°C for a minimum of thirty minutes before induction at which point isopropyl β-D1-thiogalactopyranoside (IPTG, Research Products International) was added to a final concentration of 250 μM to induce expression of IPTG. Bacteria were shaken overnight at 20°C, 250 rpm. The next morning, cultures were centrifuged at 7000 x g for 20 minutes at 10°C. Cell pellets were stored at -80°C until purification.

Uniformly ¹³C,¹⁵N-labeled Med25 for NMR experiments was expressed identically except for the use of M9 minimal media supplemented with 1 g/L ¹⁵NH₄Cl, 2 g/L ¹³C-D-glucose, and 0.5% ¹³C,¹⁵N-labeled Bioexpress media for the 1 L growth (all labeled components were purchased from Cambridge Isotopes).

Purification of Med25 AcID/13C, 15N labeled Med25 AcID

Both proteins were purified using the same fast protein liquid chromatography (FPLC) purification methods and were subjected to affinity and ion exchange chromatography. To purify Med25, cell pellets were resuspended in 25 mL lysis buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.2, 1.4 μL/mL β-mercaptoethanol (BME), 1 Roche complete mini protease inhibitor tablet) and lysed by sonication. Insoluble material was

then pelleted by centrifugation (9500 rpm, 20 min), the supernatant was removed and then re-sonicated and centrifuged again (9500 rpm, 10 min), and then filtered using a 0.45 µM syringe filter (CellTreat) and loaded onto an AKTA Pure FPLC equipped with a 5 mL Ni HisTrap HP column (GE Healthcare) pre-equilibrated with wash buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.2). Med25 was then purified using a gradient of 10–300 mM imidazole (other buffer components were constant), and fractions containing Med25 were pooled and subjected to secondary purification using a HiTrap SP HP cation exchange column (GE Healthcare) using a gradient of 0–1 M NaCl (50 mM sodium phosphate, 1 mM dithiothreitol (DTT), pH 6.8). Pooled fractions were dialyzed into storage buffer (10 mM sodium phosphate, 100 mM NaCl, 0.001% NP-40, 1% glycerol, 1mM DTT, pH 6.8) or NMR buffer (20 mM sodium phosphate, 150 mM NaCl, pH 6.5). Concentration was determined via ultraviolet/visible (UV/Vis) spectroscopy on a NanoDrop instrument at 280 nm using an extinction coefficient of 22,460 M⁻¹cm⁻¹. Aliquots were flash frozen in liquid N₂ and stored at –80°C until use. Protein identity was confirmed by mass spectrometry (Agilent Q-TOF) and purity was assessed via sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) on a 4-12% bis-tris gel stained using Quick Coomassie (Anatrace).

Expression and purification of CBP GACKIX

CBP GACKIX was expressed and purified using previously described protocols.

Solid-Phase synthesis and subsequent HPLC purification of activator peptides

Table 2.1. Sequence of fluorescein labeled activator peptides

Peptide	Sequence (N'-C')
<i>ATF6α</i> (38-75)	FITC- β Ala-FTDDELQLEAANETYENNFDNLDLDFDLMPWESDIWD
<i>ETV5</i> (38-68)	FITC- β Ala-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
<i>MLL</i> (2840-2858)	FITC- β Ala-DCGNILPSDIMDFVLKNTP
<i>Myb</i> (219-316)	FITC- β Ala-KEKRIKELELLLMSTENELKGQQLVP

Table 2.2. Masses of fluorescein labeled activator peptides

Peptide	Expected Mass (g/mol)	Observed Mass (g/mol)
<i>ATF6α</i> (38-75)	5,042.0324	5042.0394
<i>ETV5</i> (38-68)	4027.76	4027.7046
<i>MLL</i> (2840-2858)	2550.09	2550.0906
<i>Myb</i> (219-316)	3524.8047	3524.8056

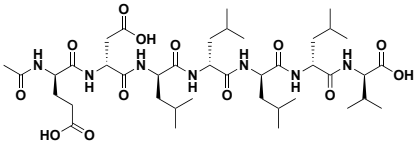
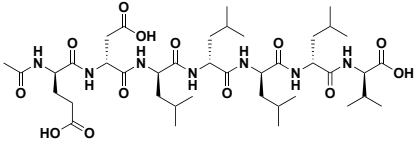
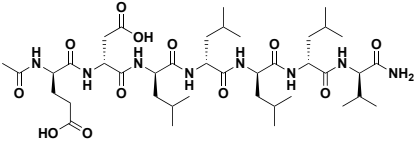
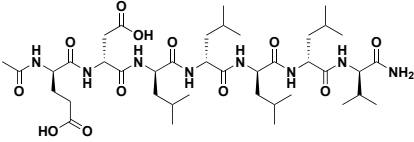
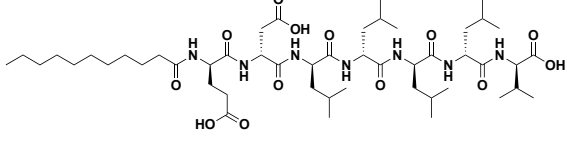
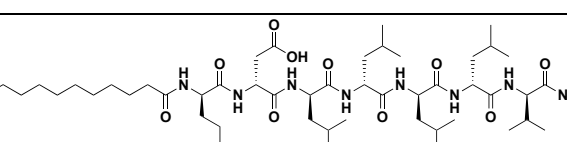
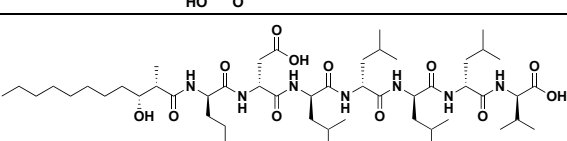
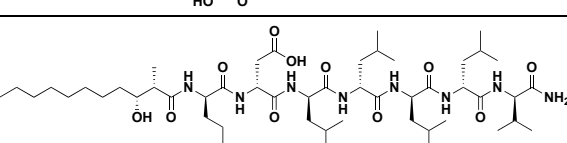
Peptides listed in Table 2.1 were made using standard fluorenylmethyloxycarbonyl (Fmoc) solid-phase synthesis methods on a Liberty Blue Microwave Synthesizer (CEM). Deprotection of the Fmoc groups was done in 20% piperidine (ChemImpex) in dimethylformamide (DMF) that was supplemented with 0.2 M Oxyma Pure (CEM) and irradiating under variable power to maintain a temperature of 90°C for one minute. For coupling reactions, 5 eq of amino acids (ChemImpex and NovaBiochem) relative to Rink Amide Protide resin (CEM) were added to diisopropylcarbodiimide (DIC, 7 eq, ChemImpex) and Oxyma Pure (5 eq) in DMF and irradiated at variable power to maintain a temperature of 90°C for 4 min. Between all couplings and deprotections, the resin was

rinsed four times with excess DMF. Fmoc-protected β -Alanine was added to the N-terminus of all four peptides.

Following synthesis, all four peptides were deprotected for 30 minutes in 20% piperidine, rinsed 3x with DMF, and then treated with fluorescein isothiocyanate (FITC, ThermoFisher) in the presence of N,N-diisopropylethylamine (DIPEA, Sigma Aldrich). All peptides were cleaved in 95% trifluoroacetic acid (TFA, Sigma Aldrich), 2.5% H₂O and 2.5% triisopropylsilane (TIPS, Sigma Aldrich) for four hours, then filtered. The remaining solution was concentrated under nitrogen and precipitated in cold diethyl ether. The precipitated peptide was dissolved in 50/50 0.1% TFA in water and acetonitrile with minimal ammonium hydroxide to help with solubility, syringe filtered using a 0.45 μ M filter, and purified via high-performance liquid chromatography (HPLC) purification. The peptide was purified using reversed phase HPLC on an Agilent 1260 Series instrument with a 250 x 10 mm Luna Omega 5 μ m PS C18 column (Phenomenex) using a 10-50 gradient elution of acetonitrile in 0.1% TFA in water over 50 min. Fractions were collected based off the peaks visible using the 495 nm fluorescence signal. Purified peptide were tested for purity using analytical HPLC, and identity was determined via mass spectrometry under negative ion mode (Agilent qTOF). Once determined to be pure, peptides were dissolved in minimal dimethyl sulfoxide (DMSO) and quantified via UV/Vis spectroscopy using FITC absorbance at 495 nm ($\epsilon_{495}=72,000 \text{ M}^{-1}\text{cm}^{-1}$). Full characterization of activators, including HPLC analytical and mass spectrometry chromatograms, can be found in Appendix I.

Solid-phase synthesis and HPLC purification of lipopeptide analogs

Table 2.3. Structure of lipopeptide analogs and their masses

Peptide	Structure	Expected Mass (g/mol)	Observed Mass
34913-2		855.4953	855.4943
34913-3		855.4953	855.494
34913-4		854.5113	854.5108
34913-5		854.5113	854.5106
34913-6		981.6362	981.6345
34913-7		980.6522	980.6507
34913-8		1011.6468	1011.6476
34913-9		1010.6627	1010.6637

The peptide sequence of the lipopeptide analogs was synthesized on Rink Amide Protide resin (C-terminal amide, CEM) or CI-MPA Protide resin) C-terminal carboxylic

acid. Resins were swelled for 30 minutes in dichloromethane (DCM, Fisher Chemical). Cl-MPA Protide resin was coupled with the first amino acid (5eq, ChemImpex and NovaBiochem) three times for 1.5 hours each without a deprotection using 0.125 M KI in 1.0 M DIPEA in DMF solution. Rink Amide Protide resin was deprotected for 30 min using 20% piperidine and the first amino acid (5eq, ChemImpex and NovaBiochem) was coupled using 0.5 M hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) in DMF, 0.49 M hydroxybenzotriazole (HOBT) in DMF, and 1.0 M DIPEA in DMF. Couplings proceeded for 1.5 hours. After this initial coupling, both resins were rinsed 3x with DMF, deprotected for 30 min with 20% piperidine, and then the next amino acid was coupled for 1.5 hours. This was repeated until the peptide sequence was complete. A final deprotection was performed and analogs 34913-2–34913-5 were acetylated at the N-terminus using a cocktail of acetic anhydride and triethylamine (TEA, Fisher Scientific) in DCM. Analog 34913-6 and 34913-7 were coupled with undecanoic acid (5eq) using 0.5 M HBTU in DMF, 0.49 M HOBT in DMF, and 1.0 M DIPEA in DMF. Analog 34913-8 and 34913-9 were coupled to crude (2S, 3R)-3-hydroxy-2-methylundecanoic acid (5eq) using 0.5 M HBTU in DMF, 0.49 M HOBT in DMF, and 1.0 M DIPEA in DMF.

All peptides were cleaved in 95% trifluoroacetic acid (Sigma Aldrich), 2.5% H₂O and 2.5% TIPS (Sigma Aldrich) for two hours, then filtered. The remaining solution was concentrated under nitrogen and precipitated in cold diethyl ether. The precipitated lipopeptide was dissolved in 100% methanol and syringe filtered using a 0.45 µm filter, and purified via HPLC purification. The peptide was purified using reversed phase HPLC on an Agilent 1260 Series instrument with a 250 x 10 mm Luna Omega 5 µm PS C18 column (Phenomenex) using a 10-100 gradient elution of acetonitrile in 0.1% TFA in water

over 45 min. Fractions were collected based off the peaks visible using the 214 nm fluorescence signal. Purified peptide were tested for purity using analytical HPLC, and identity was determined via mass spectrometry under negative ion mode (Agilent qTOF). Once determined to be pure, the mass of the purified peptide was taken and a 10-20 mM stock were made DMSO (low enough concentration to assure lipopeptide is fully dissolved). These stocks were aliquoted and stored at -2°C. Full characterization of lipopeptide analogs, including HPLC analytical and mass spectrometry chromatograms, can be found in Appendix I.

Direct Binding Assay

Direct binding experiments measured by fluorescence polarization were performed as previously described. Binding experiments were performed in black, round-bottom 384-well plates (Corning) and read on a PHERAStar multi-mode plate reader with polarized excitation at 485 nm and emission intensity measured through parallel and perpendicular polarized 535 nm filters. Wells were filled with 10 μ L of storage buffer (10 mM sodium phosphate, 100 mM NaCl, 0.001% NP40, 10% glycerol, pH 6.8). Protein was added to top well and diluted vertically with the last well having storage buffer only. FITC-labeled peptides were diluted in storage buffer (10 mM sodium phosphate, 100 mM NaCl, 0.01% NP40, 10% glycerol, pH 6.8) and 10 μ L were added to every well, resulting is a final peptide concentration of 20 nM. Data was analyzed using Graphpad Prism 8.0. For direct binding experiments, a binding curve that accounts for ligand depletion was fit to the observed polarizations values as a function of protein to obtain the apparent dissociation constant, K_d :

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

“x” and “a” are the total concentrations of protein and fluorescently labeled peptide respectively, “y” is the observed polarization at a given protein concentration, “b” is maximum observed polarization value, and “c” is minimum observed polarization value.

Competition Binding Assay

Inhibition assays were performed in triplicate with a final sample volume of 20 μ L in a low volume, non-binding 384-well black plate (Corning). Protein was precomplexed with FITC-labeled peptide using two-fold the concentrations of protein and peptide to reach 50% of the tracer bound in binding buffer (5 mM sodium phosphate, 100 mM NaCl, 10% glycerol, 0.001% NP40, pH 6.8). Small molecule, dissolved in DMSO and diluted in assay buffer was serially diluted down two-fold with assay buffer going down the plate to row N. 10 μ L of the pre-complexed peptide-protein complex was then added to each well for a final volume of 20 μ L (Row O should contain only the peptide-protein complex-no inhibitor). The final well, row P, served as a negative control, tracer (FITC-labeled peptide) only. The plate was incubated for thirty minutes at room temperature before fluorescence polarization was measured on a PHERAstar plate reader. The experiment was conducted using polarized excitation at 485 nm and emission intensity was measured through a parallel and a perpendicular polarized 535 nm filter. Polarization values were plotted against log[inhibitor] and inhibition curves were then fit to a non-linear regression using Prism’s built-in equation ‘log(agonist) vs response – variable slope (four parameters)’ with the bottom constraint set to ‘constant or equal to 35 (as set on the Pherastar) from which

the EC₅₀ values were calculated. The EC₅₀ values were converted to K_i values using the apparent K_d value based on the direct binding of the specific coactivator•activator PPI using a K_i calculator from the Nikolovska-Coleska lab.³⁴ Each data point is an average of three independent replicates with the indicated error representing the standard deviation of the three replicates.

NMR Spectroscopy

Constant time ¹H,¹³C-HSQC experiments were performed with 75 μM uniformly ¹³C,¹⁵N labeled Med25 in NMR buffer (20 mM sodium phosphate pH 6.5, 150 mM NaCl, 3 mM DTT, 10% D₂O, and 2% DMSO) on a Bruker 600 MHz instrument equipped with a cryoprobe. HSQC experiments were processed in NMRPipe³⁵ and visualized with NMRFAM-Sparky.³⁶ All chemical shift perturbation analyses were performed on samples with 1.1 equivalents of unlabeled binding partner. Peak assignments of lipopeptide analog-bound complexes were achieved by titration experiments with the lipopeptide analog (34913-8/34913-9) with titration points of 0.2, 0.5, 0.8, 1.1, 2, and 3 equivalents of lipopeptide analog. Chemical shift perturbations (Δδ) were calculated from the proton (Δδ_H) and carbon (Δδ_C) chemical shifts by:

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

Mammalian Cell Culture

HeLa cells were purchased from ATCC. VARI068 cell line arose from culturing a triple negative breast cancer patient-derived xenograft as described previously.^{24,25} HeLa cells were grown using Dulbecco's Modified Eagle Medium (DMEM, Gibco, cat.#:

11965092) supplemented with 10% fetal bovine serum (FBS). VARI068 cells were grown using DMem (Gibco, cat. #: 11965092) supplemented with 10% FBS, 1x Antibiotic-Antimycotic, and 10mg/mL Gentamycin. Both cells lines were grown at 37°C and 5% CO₂.

Cellular Thermal Shift Assays

VARI068 cells^{24,25} were harvested using standard protocols and pelleted (~4 million cells per pellet) in 1.5 mL microcentrifuge tubes at 1500 G for 5 minutes at 4°C. Nuclear extracts were generated using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific, cat #: 78833) and following manufacturer's protocol. Nuclear extracts were then buffer exchanged into phosphate buffered saline (PBS) using Zeba Desalting Column, 7K MWCO (Thermo Fisher Scientific, cat #: 89882). Prepared nuclear extracts were split into 3 epitubes. 34913 Amide and 34913 Acid (dissolved in DMSO) were added to nuclear extracts to reach proper concentration, with equivalent volume of DMSO being added to the final epitube. Final concentration of DMSO was 0.1% v/v. Dosed nuclear extracts were incubated at room temperature for 30 mins. Following incubation, extracts were aliquoted into thin-walled PCR tubes (20 µL per tube, ~300,000 cells per tube).

A Labnet Multigene OPTIMAX PCR was used to heat each sample for 3 minutes. Six temperatures were tested: 54°C, 58°C, 62°C, 66°C, 70°C, 74°C. Once heated, contents of the PCR tubes were transferred to epitubes and centrifuged at 17000 x g for 2 minutes at 4°C to remove precipitated protein. About 18 µL of supernatant was transferred to clean epitubes, making sure not to disturb the precipitated protein pellet. LDS loading dye was added and samples were boiled at 95°C for 10 minutes. 15 µL of

sample was loaded into homemade gel (4% stacking, 12% resolving gel) and run at 170V for 1 hour or until dye front runs off bottom of the gel. Protein was transferred from gel to PDVF membrane using a Bio-Rad Trans-Blot Turbo Transfer System following the standard protocols. Membrane was then blocked for 1 hour at room temperature with gentle shaking using SuperBlock™ Blocking Buffer in PBS (Thermo Scientific, 37515). Super Block was removed and Med25 antibody (Novus Biologicals, NBP2-55868) was added to membrane (1:500 dilution with 1:500 Tween 20 in SuperBlock™) and incubated overnight at 4°C with gentle shaking. Primary antibody was removed and membrane was washed three times with PBST for 5 minutes each. Secondary antibody (Santa Cruz, sc-2357, 1:1000 with 1:500 Tween 20 in SuperBlock™) was added to the membrane and incubated for 1 hour at room temperature with gentle shaking. Secondary antibody was removed and membrane was washed three times with PBST for 5 minutes each. HRP substrate (Thermo Scientific, 34095) was added to the membrane and incubated for 1 minute at room temperature. Western blot was visualized using Chemiluminescence on an Azure Biosystems c600 imager. Analysis was conducted on ImageJ.

Quantitative Polymerase Chain Reaction

For endogenous gene expression analysis, VARI068 cells^{24,25} were seeded in a 24-well plate (1x10⁵ cells/well) and allowed to adhere overnight at 37°C. Media was removed and replaced with fresh media containing vehicle or compound delivered in DMSO (0.2% v/v) at the indicated concentrations. After 3 hours, the media was removed, and the total RNA was isolated using RNeasy Plus Mini Kits (Qiagen, 74134) following the manufacturer's protocol. Each RNA sample was used to synthesize cDNA using

iScript Reverse Transcription Supermix (Bio-Rad, 1708841). RT-qPCR samples were run in triplicate using the cDNA from the previous step, GoTaq qPCR Master Mix (Promega, A6001), and the following primers:

human *RPL19* Forward. 5':ATGTATCACAGCCTGTACCTG:3'

human *RPL19* Reverse. 5':TTCTTGGTCTCTCTTCCTCCTTG:3'

MMP2 Forward. 5':CATTCCAGGCATCTGCGATGAG:3'

MMP2 Reverse. 5':AGCGAGTGGATGCCGCCTTTAA:3'

HSPA5 Forward. 5':CTGGGTACATTTGATCTGACTGG:3'

HSPA5 Reverse. 5':CTTACCGACCTTTCGGTGGTCCTACG:3'

2.6 References

1. Raaijmakers, J.M., De Bruijn, I., Nybroe, O., Ongena, M. Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol Rev.* **2010**, 34(6), 1037-1062.
2. Cameotra, S.S., Makkar, R.S. Recent applications of biosurfactants as biological and immunological molecules. *Curr. Opin. Microbiol.* **2004**, 7(3), 262-266.
3. Dey, G., Bharti, R., Banerjee, I., et al. Pre-clinical risk assessment and therapeutic potential of antitumor lipopeptide 'Iturin A' in an *in vivo* and *in vitro* model. *RSC Adv.* **2016**, 6, 71612-71623.
4. Dey, G., Bharti, R., Sen, R., Mandel, M. Microbial amphiphiles: a class of promising new-generation anticancer agents. *Drug Discov. Today.* **2015**, 20(1), 136-146.
5. Meena, K.R., Kanwar, S.S. Lipopeptides as the antifungal and antibacterial agents: applications in food safety and therapeutics. *Biomed. Res. Int.* **2015**, 2015, 473050.
6. Park, S.Y., Kim, J., Lee, Y.J., et al. Surfactin suppresses TPA-induced breast cancer cell invasion through the inhibition of MMP-9 expression. *Int. Jour. Oncol.* **2012**, 42(1), 287-296.
7. Ahmad, I.Z., Parvez, S., Tabassum, H. Cyanobacterial peptides with respect to anticancer activity: Structural and functional perspective. *Stud. Nat. Prod. Chem.* **2020**, 67, 345-388.
8. Henderson, A.R., Henley, M.J., Foster, N.J., et al. Conservation of coactivator engagement mechanism enables small molecule allosteric modulators. *Proc. Natl. Acad. Sci.* **2018**, 115(36), 8960-8965.
9. Henley, M.J., Linhares, B.M., Morgan, B.S., et al. Unexpected specificity within dynamic transcriptional protein-protein complexes. *Proc. Natl. Acad. Sci.* **2020**, 117(44), 27346-27353.
10. Sela, D., Conkright, J.J., Chen, L., et al. Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6 α . *J. Biol. Chem.* **2013**, 288(36), 26179-26187.
11. Landrieu, I., Verger, A., Baert, J.L., et al. Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* **2015**, 43(14), 7110-7121.
12. Vojnic, E., Mourão, A., Seizl, M., et al. Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **2011**, 18(4), 404-409.

13. Bontems, F., Verger, A., Dewitte F, et al. NMR structure of the human Mediator MED25 ACID domain. *J. Struct. Biol.* **2011**, *174*(1), 245-251.
14. Milbradt, A.G., Kulkarni, M., Yi, T., et al. Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol.* **2011**, *18*(4), 410-415.
15. Benz, C.C., O'Hagan, R.C., Richter, B., et al. HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene.* **1997**, *15*(13), 1513–1525.
16. de Launoit, Y., Chotteau-Lelievre, A., Beaudoin, C., et al. The PEA3 group of ETS-related transcription factors. *Adv. Exp. Med. Biol.* **2000**, 480, 107–116.
17. Pellecchia, A., Pescucci, E., De Lorenzo, E., et al. Overexpression of ETV4 is oncogenic in prostate cells through promotion of both cell proliferation and epithelial to mesenchymal transition. *Oncogenesis.* **2012**. *1*(7), e20.
18. Myers, E., Hill, A.D.K., Kelly, G., et al. A positive role for PEA3 in HER2-mediated breast tumour progression. *Br. J. Cancer.* **2006**, *95*(10), 1404–1409.
19. Fauquette, V., Perrais, M., Cerulis, S., et al. The antagonistic regulation of human MUC4 and ErbB-2 genes by the Ets protein PEA3 in pancreatic cancer cells: implications for the proliferation/differentiation balance in the cells. *Biochem. J.* **2005**, *386*(1), 35–45.
20. Haze, K., Yoshida, H., Yanagi, H., Yura, T., Mori, K. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol. Biol. Cell.* **1999**, *10*(11), 3787–3799.
21. Yoshida, H., Okada, T., Haze, K, et al. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol. Cell. Biol.* **2000**, *20*(18), 6755–6767.
22. Wang, Y., Shen, J., Arenzana, W., et al. Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. *J. Biol. Chem.* **2000**, *275*(35), 27013–27020.
23. Fernandez, P. M., Tabbara, S.O., Jacobs, L.K., et al. Overexpression of the glucose-regulated stress gene GRP78 in malignant but not benign human breast lesions. *Breast Cancer Res. Treat.* **2000**, *59*(1), 15–26.
24. Aw Yong, K.M., Ulintz, P.J., Caceres, S., et al. Heterogeneity at the Invasion Front of Triple Negative Breast Cancer Cells. *Sci. Rep.* **2020**, *10*(1), 5781.
25. Liu, M., Liu, Y., Deng, L., et al. Transcriptional Profiles of Different States of Cancer Stem Cells in Triple-Negative Breast Cancer. *Mol. Cancer.* **2018**, *17*(1), 65.

26. Cress, W.D., Triezenberg, S.J. Critical structural elements of the VP16 transcriptional activation domain. *Science*. **1991**, 251(4989), 87-90.
27. Regier, J.L., Shen, F., Triezenberg, S.J. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci.* **1993**, 90(3), 883-887.
28. Hope, I.A., Mahadevan, S., Struhl, K. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature*. **1998**, 333(6174), 635-640.
29. Hope, I.A., Struhl, K. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell*. **1986**, 46(6), 885-894.
30. Lum, K.M., Sato, Y., Beyer, B.A., et al. Mapping Protein Targets of Bioactive Small Molecules Using Lipid-Based Chemical Proteomics. *ACS Chem. Bio.* **2017**, 12(10), 2671-2681.
31. Niphakis, M.J., Lum, K.M., Cognetta, A.B. 3rd, et al. A Global Map of Lipid-Binding Proteins and Their Ligandability in Cells. *Cell*. **2015**, 161(7), 1668-1680.
32. Beyersdorf, M.S., Mechanistic Investigation of the Transcriptional Coactivator Med25 and its Protein-Protein Interaction Network. *Deep Blue Doc.* **2017**.
33. Feng, Z., Xu, B. Inspiration from the mirror: D-amino acid containing peptides in biomedical approaches. *Biomol. Concepts*. **2016**, 7(3) 179-187.
34. Nikolovska-Coleska, Z., Wang, R., Fang, X., et al. Development and Optimization of a Binding Assay for the XIAP BIR3 Domain Using Fluorescence Polarization *Anal. Biochem.* **2004**, 332(2), 261-273.
35. Delaglio, F., Grzesiek, S., Vuister, G.W., et al. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. **1995**, 6(3), 277-293.
36. Lee, W., Tonelli, Markley, J.L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*. **2015**, 31(8), 1325-1327.
37. Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature*. **1988**, 333(6170), 210-212.
38. Brzovic, P.S., Heikaus, C.C., Kisselev, L., et al. The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell*. **2011**, 44(6), 942-953.

39. Warfield, L., Tuttle, L.M., Pacheco, D., et al. A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci.* **2014**, 111(34), 3506-3513.
40. Tuttle, L.M., Pacheco, D., Warfield, L., et al. Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep.* **2018**, 22(12), 3251-3264.
41. Tanaka, M. Modulation of promoter occupancy by cooperative DNA binding and activation-domain function is a major determinant of transcriptional regulation by activators in vivo. *Proc. Natl. Acad. Sci.* **1996**, 93(9), 4311-4315.
42. Minter, A.R., Brennan, B.B., Mapp, A.K. A small molecule transcriptional activation domain. *J. Am. Chem. Soc.* **2004**, 126(34), 10504-10505.
43. Rowe, S.P., Mapp, A.K. Assessing the permissiveness of transcriptional activator binding sites. *Biopolymers.* **2008**, 89(7), 578-581.
44. Thuerauf, D.J., Morrison, L.E., Hoover, H., Glembotski, C.C. Coordination of ATF6-mediated transcription and ATF6 degradation by a domain that is shared with the viral transcription factor, VP16. *J. Biol. Chem.* **2002**, 277(23), 20734-20739.
45. Kitamata, M., Inaba, T., Suetsugu, S. The roles of the diversity of amphipathic lipids in shaping membranes by membrane-shaping proteins. *Biochem. Soc. Trans.* **2020**, 48(3), 837-851.

Chapter 3

Lipopeptide 34913-8 is a Transcriptional Domain Mimic

3.1 Abstract

The protein-protein interactions formed between transcriptional coactivators and transcriptional activators are critical for the activation of transcription. The most common models of activator•coactivator PPIs invoke non-specific electrostatic and hydrophobic forces that drive their function, with no requirement of specific contacts between the two binding partners. Recent work from our lab has countered this model, determining that small variations in a TAD peptide sequence can lead to the binding of structurally distinct coactivator conformations. The coactivator Med25 and the ETV/PEA3 family of transcriptional activators can act as a model system for this hypothesis as these activators have sequence homology.

The main objective of this Chapter was to use lipopeptide analog, 34913-8, identified in Chapter 2, as a molecular probe to gain insight into how individual residues of the 34913-8 peptide sequence affect its binding to transcriptional coactivators, particularly Med25 AcID. This was accomplished by synthesizing a series of alanine and neutrally charged substitution analogs and evaluating their inhibitory activity against Med25•ATF6 α using competition fluorescence polarization assays. The selectivity of these analogs was then examined to further identify individual residues that are important for selectivity of 34913-8 for the Med25 PPIs over the CBPKIX PPIs. Additionally, 34913-8 was tested

against the Med25•ETV PPIs, resulting in the selective inhibition of specific Med25•ETV PPIs over others. The data presented in this Chapter suggests that 34913-8 can be used to inhibit specific conformations of Med25. It also demonstrates the potential in synthesizing further lipopeptide analogs to potently and selectively target Med25 PPIs, as well as other coactivator•activator PPIs.

3.2 Introduction

Activator•coactivator molecular recognition

The PPIs between transcriptional coactivators and activators are crucial for the proper activation of gene expression. This is accomplished when a coactivator interacts with the TAD of an activator, thus facilitating recruitment of additional transcriptional machinery to the promoter region of the gene of interest. The historical model of coactivator•activator PPIs is that they are non-specific, instead proposing that the structurally well-defined DNA-binding domain of the transcriptional activator is the driving force for gene-specific function (Figure 3.1).¹⁻⁵

This non-specific activator-coactivator PPI recognition model originated due to the intrinsically disordered structure of transcriptional activators and early mutagenesis studies hypothesizing that these PPIs are not sequence-specific.⁶ One foundational example of the latter utilized the fusion of random peptides sequences to the Gal4 DNA binding domain to demonstrate that peptides containing several acidic residues activated transcription. Additionally, an increase in net negative charge of these peptides resulted in stronger activation of transcription.⁷ Mutational studies of several transcriptional activators, including VP16 and Gcn4, determined that hydrophobic regions of these TADs

may also be critical for transcriptional activity.⁷⁻¹² Although amphipathic TADs are almost exclusively unstructured in the absence of binding partners, several NMR studies and computational work determined that TADs adopt a helical conformation when bound to coactivators.^{7,9,13}

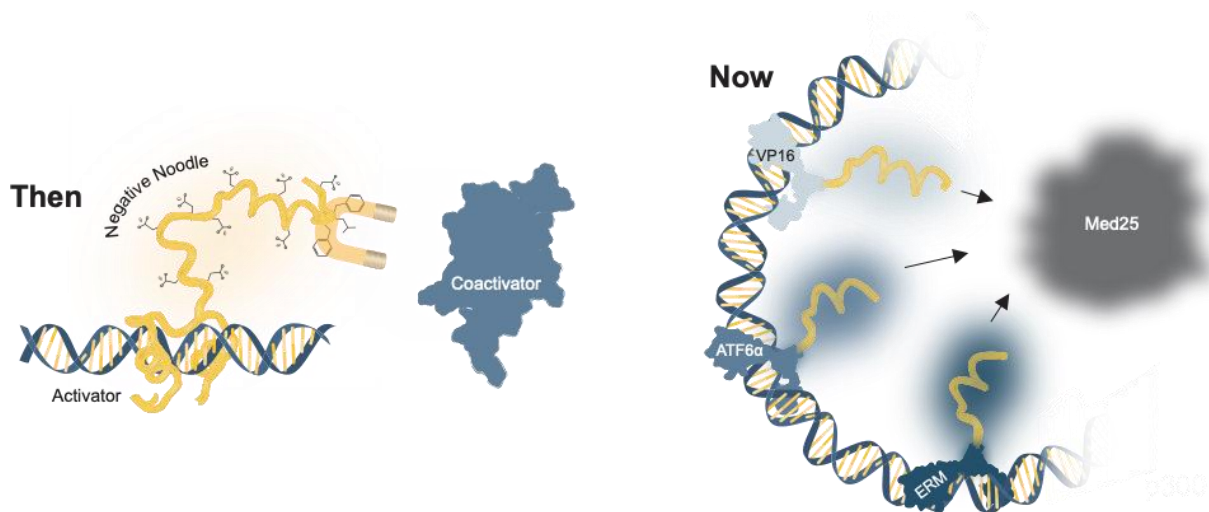


Figure 3.1. Models of molecular recognition. The model for the mechanism by which activators•coactivator PPIs occur is often debated. The earliest hypothesis suggested these interactions were driven by largely nonspecific interactions, with intrinsically disordered, amphipathic TADs binding to coactivators. Recent biophysical and structural data suggests this model is oversimplified. In addition to TADs existing in conformational ensembles, the ABD also undergoes conformational changes due to internal malleability. The result is “fuzzy” complexes, such as Med25•activator PPIs that can undergo conformational changes.

Moving away from a non-specific binding model

Over the past few decades, there have been several studies that have countered the non-specific binding model. Mutagenesis studies have been used to identify specific residues within TADs that are critical for the function of transcriptional activators.¹⁰⁻¹² In one case, a hydrophobic residue, F442, in the TAD of VP16, was identified as critical for transcriptional activity. Substitution of this F442 with nonaromatic residues, serine and alanine, saw complete inactivation of VP16. The aromatic residue, tyrosine, was able to functionally replace F422, though the transcriptional activity of VP16 was weakened.¹¹⁻¹²

Additional studies demonstrated through the insertion of helix breaking residues, that the predicted helical structure of these TADs is not critical for transcriptional activation.¹¹ This suggests that while most TADs are amphipathic peptide sequences with the propensity to form α -helices, there are key residues that are critical for specific transcriptional activity and thus interaction with coactivators.

Recently, the molecular recognition mechanisms of coactivator•activator PPIs were dissected using the TADs of the ETV/PEA3 family of transcriptional activators (ETV1, ETV4, and ETV5) and the AcID domain of the coactivator Med25.¹⁶ Considerable NMR data demonstrated that the three ETV activators interact with the H1 binding surface of Med25.^{16,17} Additionally, the core region of these activators differs by only two residues and have the same dissociation constants (Figure 3.2).¹⁶ This largely conserved core region has a nearly identical affinity for Med25 AcID as the full length TADS indicating that the amino terminal sequences are only minor contributors to overall affinity.¹⁶



Figure 3.2. ETV activator TAD sequence alignment. The ETV/PEA3 transcriptional activator peptide sequences have high sequence homology. The sequence alignments of the three TADs are shown above. The differences between the sequences are bolded. The helix denotes the activator core region that undergoes coupled folding and binding with Med25, as determined by NMR chemical shift analysis. This figure is adapted from Henley, et al. (2020).¹⁶

Data from stopped-flow kinetics experiments reveal that despite sharing nearly identical binding affinities for Med25, ETV4 preferentially interacts with a different conformer of Med25 as compared to ETV1 and ETV5. This suggests that these variations in the TAD peptide sequence of ETV4 and ETV1/5 are contributing significantly to these

distinct conformations of Med25. Further analysis of the variable regions of these activators determined that mutations in the amino terminus alter the conformational ensembles of Med25, proposing that these activators engage differently with Med25 (Figure 3.3). NMR studies further support this model, with both N-terminal mutants altering an overlapping subset of residues in the core binding site, which is typically occupied by the TAD of the helical region.¹⁶

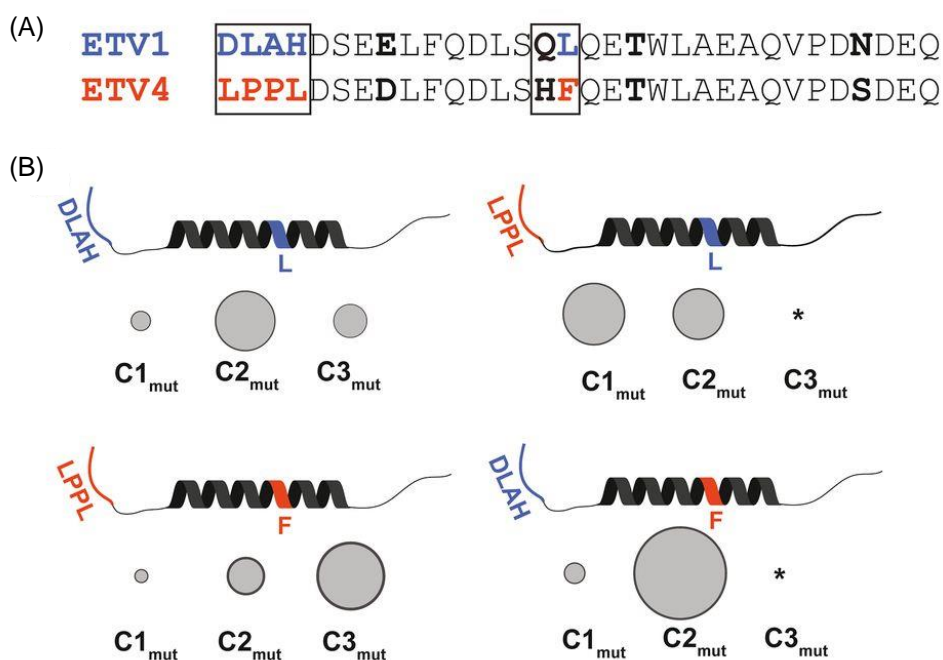


Figure 3.3. Variable residues mediate differences in Med25•ETV conformation behavior. (A) Alignment of ETV1 and ETV4 activators with regions that were selected for mutational analysis boxed. Bolded residues are conserved between ETV1 and ETV5, but not ETV4. Regions/residues that affected the conformational ensemble are color coded to ETV1 (blue) or ETV4 (orange). (B) Results from kinetics experiments of mutant TADs, for native (Left) and nonnative (Right) combinations of variable N-terminus and helical binding regions. Variants were made based on the ETV4 sequence. The data shown are the average across all of the variants tested from each group, with the error (dark gray outer circle) representing the SD. This figure is adapted from Henley, et al. (2020).¹⁶

In the previous Chapter, 34913-8 was identified as an amphipathic lipopeptide that selectively targets Med25 *in vitro* and *in cellulo*. More specifically, the peptide sequence of 34913-8 contains hydrophobic and acidic residues that resemble the amphipathic

peptide sequence of a TAD. Based on these similarities, I hypothesize that 34913-8 can interact with Med25 AcID similarly to the TADs of transcriptional activators. I further propose that specific residues within the peptide sequence of 34913-8 play a role in its inhibition of the Med25•ATF6 α PPI. In this Chapter, the peptide sequence of 34913-8 will undergo several substitutions to identify key residues that contribute to its activity. These analogs will be tested against other coactivator•activator PPIs using competition fluorescence polarization assays to determine the importance of each of these residues in coactivator•activator specificity. Lastly, 34913-8 will be tested against the Med25•ETV activator PPIs using competition fluorescence polarization. This will determine if these lipopeptides can be used as molecular probes to selectively target one Med25 AcID conformation over another, despite the overlapping binding surfaces of several of its activator binding partners. Altogether, this information acts as an improved targeting strategy for the design and synthesis of modulators that can stabilize specific conformations of coactivator ABDs, leading to inhibition of coactivator•activator PPIs and their subsequent gene expression.

3.3 Results and Discussion

Analysis of 34913-8 analogs with alanine substitutions

Several lines of evidence suggest that sequence identity plays an important role in the binding of an activator to its molecular target.^{8,10-12} The lipopeptide analog, 34913-8, contains a peptide sequence (N'-Glu-Asp-Leu-Leu-Leu-Leu-Val-C') composed primarily of acidic and hydrophobic residues, resembling activator peptides known to adopt helices upon binding. At the conclusion of the studies described in Chapter 2, it was understood that 34913-8 successfully inhibits Med25 PPIs, most prominently Med25•ATF6 α , though

little is known about the contribution the individual residues of the peptide sequence have on this inhibitory activity.

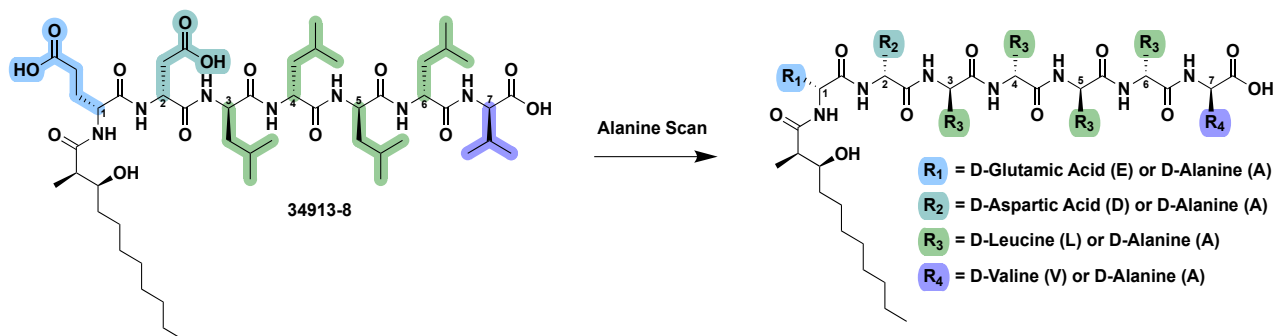


Figure 3.4. Alanine residue substitutions of 34913-8. An alanine scan was performed in which each D-amino acid residue in the 34913-8 peptide sequence is replaced with a D-alanine. The substitutions are numbered starting from the N-terminal end of the peptide sequence as shown, with analogs being named 34913-Ala1, 34913-Ala2, etc.

To investigate the importance of these individual residues, an alanine scan was performed using the peptide sequence of lipopeptide 34913-8 as the template (Figure 3.4). Seven analogs were synthesized with an individual D-alanine substitution replacing every residue in the sequence. The N-terminal residue (D-Glu) substitution was denoted as 34913-Ala1 and subsequent residues were named in sequential order, with the C-terminal residue (D-Val) substitution denoted as 34913-Ala7. The inhibitory activity of these analogs for the Med25•ATF6 α PPI was determined using a competition fluorescence polarization assay. Unfortunately, these 7 analogs were all contaminated with (R)-4-benzyl-3-((2S,3R)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, the starting material from the final reaction in the lipophilic tail synthesis (refer to Appendix I for approximate percent purities of the analogs and mass spectrometry identification of the impurity). Additionally, 34913-Ala7 was unable to be used due to contamination issues. For comparison purposes among the series, all K_i values were normalized against

that of 34913-Ala1. Future experiments will re-evaluate the activity of analogs that are at least 95% pure.

Based on previous literature, it was expected that analogs with an increased net negative charge and bulkier hydrophobic groups, such as leucine, would have more potent inhibitory activity for Med25•ATF6 α compared to analogs that do not.¹⁰⁻¹² Focusing first on the alanine substitutions of the hydrophobic residues of the peptide, substitution of D-alanine for each D-leucine residue in positions 3-6 resulted in varied inhibition of Med25•ATF6 α (Figure 3.5A, green). This suggests that certain leucine residues in the peptide sequence contribute more to the inhibition of Med25 than others. For example, 34913-Ala3 and 34913-Ala4 both have weaker activity when compared to 34913-Ala5 and 34913-Ala6 thus highlighting the importance of bulkier hydrophobic residues at positions 3 and 4 for inhibition of Med25•ATF6 α . This revealed that, like the TADs of transcriptional activators, certain hydrophobic residues within the 34913-8 peptide sequence are more critical for the inhibitory activity against Med25•ATF6 α than others.

Analogs 34913-Ala1 and 34913-Ala2 offered interesting insight into the importance of net negative charge for the inhibition of the Med25 AcID•ATF6 α PPI (Figure 3.5A, blue). Both of these substitution analogs result in the removal of a negative charge, though they differ greatly in observed activity, with 34913-Ala1 displaying a weaker inhibitory activity against Med25•ATF6 α when compared to 34913-Ala2. This observable difference in inhibition of these two analogs despite having the same net negative charge opposes the previously proposed binding models suggesting that increase in net negative charge results in increased inhibitory activity. Overall, these alanine substitution analogs highlight the individual contributions each residue within the 34913-8 peptide sequence has

towards the inhibition of Med25•ATF6 α , likely due to its specific binding to Med25 AcID. It is possible that certain substitutions may even vary in inhibitory activity due to their binding to distinct conformations of Med25, though further experiments would need to be performed to confirm this hypothesis.

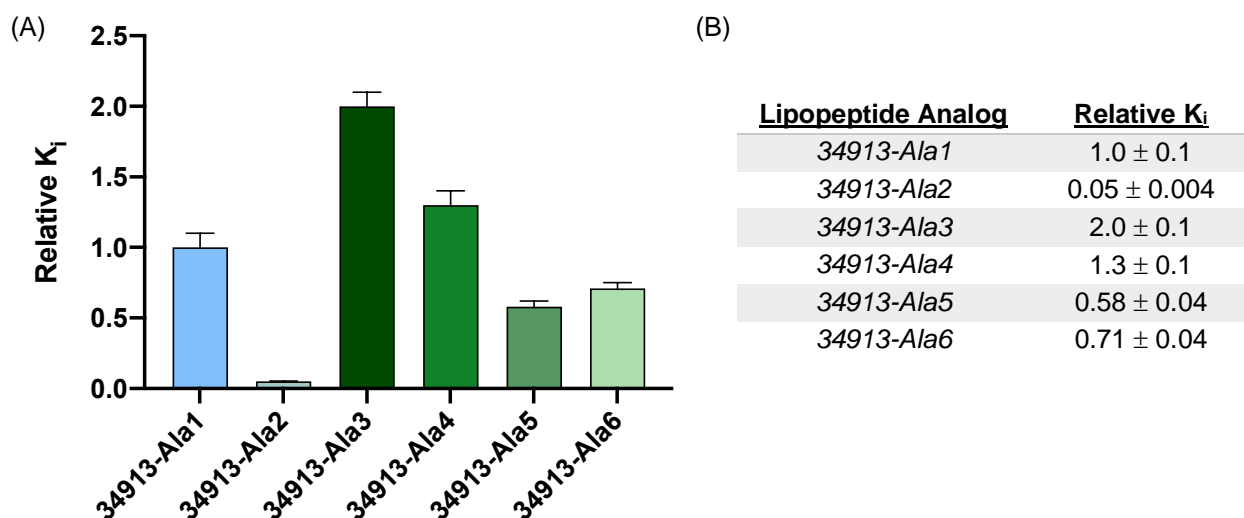


Figure 3.5. Alanine substitutions determine important residues for inhibitory activity. (A) Apparent EC_{50} values were determined through titration of alanine analogs for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM). Apparent EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of the Med25 AcID•ATF6 α PPI using a K_i calculator.¹⁸ K_i values were then normalized against 34913-Ala1. (B) A compiled list of the substitution analogs with their relative K_i values and respective errors.

Importance of charged amino acids for 34913-8 binding

The alanine scan demonstrated that the two negatively charged residues of 34913-8 showed varying inhibitory activity when substituted with an alanine, suggesting that negative charge alone is not sufficient for the inhibition of Med25•ATF6 α . It is possible that the bulkiness of the structure of these residues, instead of the negative charge, is what is contributing to this activity. To investigate this role further, each negatively charged residue in the 34913-8 peptide sequence (D-Glu and D-Asp) was substituted with their neutral counterparts (D-Gln and D-Asn, respectively) (Figure 3.6). Additionally,

a double substitution was made where both acidic residues were replaced with their neutral counterparts. This will result in the removal of the negative charge, with retention of the polarity and overall structure of these individual residues, allowing for a more complete evaluation of the impact these negatively charged residues have on the inhibition of Med25 AcID•ATF6 α (Figure 3.7). These specific neutrally charged analogs also contained the same impurity that was found in the alanine analogs. For comparison purposes among the series, all K_i values were normalized against that of 34913-Ala1. Future experiments will re-evaluate the activity of analogs that are at least 95% pure.

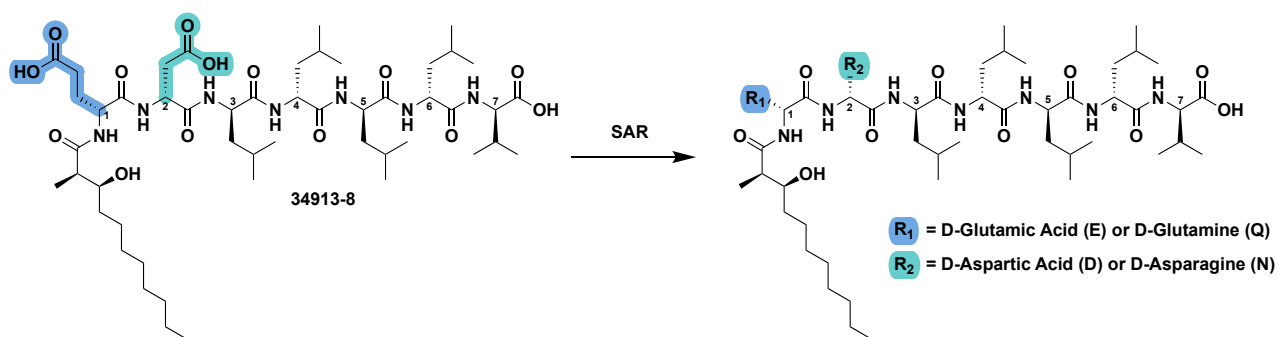


Figure 3.6. Charged residue substitutions of 34913-8. Three charged residue substitutions were performed in which the two negatively charged D-amino acid residues in the 34913-8 peptide sequence were each replaced with their neutral amino acid counterparts. An additionally analog was also synthesized in which both negatively charged residues were replaced with their neutral counterparts. The substitutions are numbered starting from the N-terminal end of the peptide sequence as shown, with analogs being named 34913-Gln1, 34913-Asn2, and 34913-Gln1Asn2.

It was observed that the two individually charged analogs, 34913-Gln1 and 34913-Asn2, had fairly comparable inhibition to that of 34913-Ala1. This again is starkly different from the 34913-Ala2 analog that seems much more potent against Med25•ATF6 α . The same chemical space occupied by both the negatively charged amino acids and their neutral amino acid counterparts coupled with the increase in inhibitory activity of 34913-Gln1Asn2 suggests that the size of the negatively charged amino acids is not the main

contributor for inhibition of the Med25•ATF6 α PPI. The complete removal of negatively charged residues on 34913-Gln1Asn2 demonstrates an increased potency for the Med25•ATF6 α PPI comparable to the activity of 34913-Ala2. This result further disproves the coactivator•activator binding model that suggests that net negative charge of TADs is sufficient for its inhibitory activity.^{3,4} It is possible that this double substitution analog may bind in an alternate binding mode compared to 34913-8 and the other substitution analogs, but additional experiments would need to be performed for further confirmation.

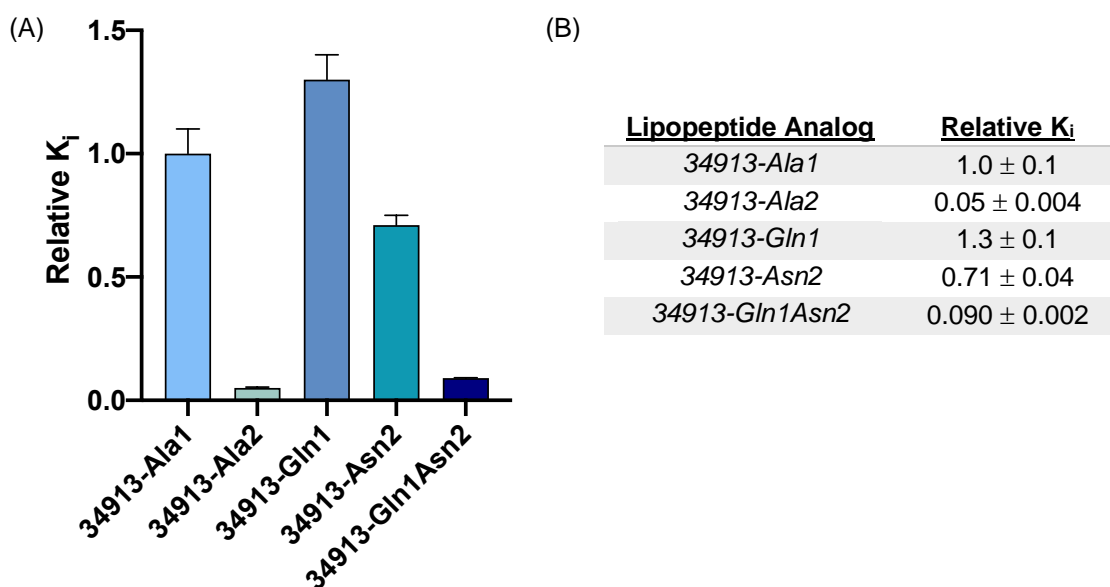


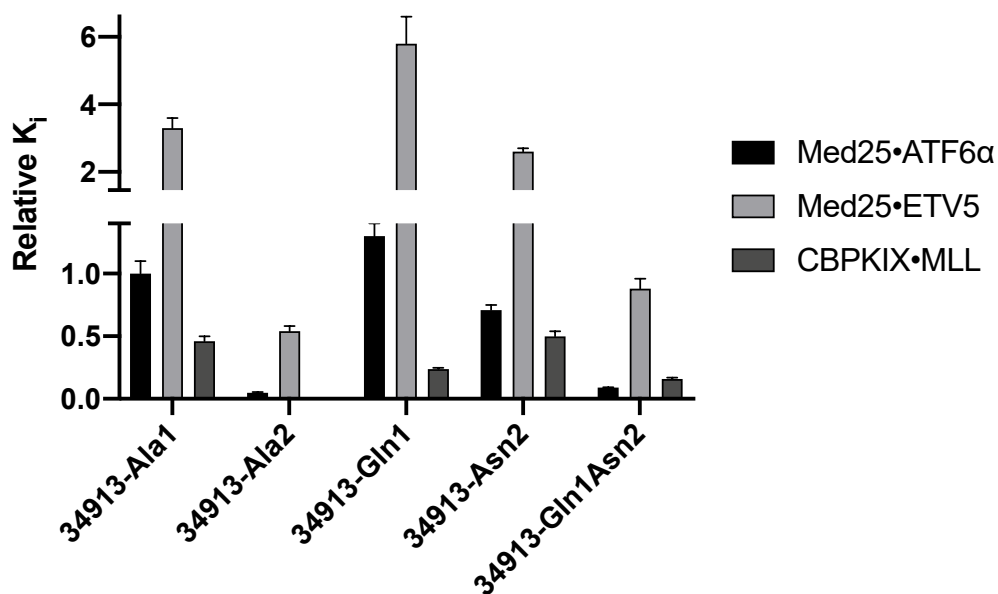
Figure 3.7. Relative inhibitory activity of substitution analogs. (A) Apparent EC_{50} values were determined through titration of the substitution analogs of the negatively charged amino acids for Med25 AcID•ATF6 α PPI performed in experimental triplicate with the indicated error (SDOM). The EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of the Med25 AcID•ATF6 α PPI using a K_i calculator.¹⁸ K_i values were then normalized against 34913-Ala1. (B) A compiled list of the substitution analogs with their relative K_i values and respective errors.

Determination of selectivity of the substitution analogs

Amphipathic transcriptional activation domains are rarely selective for a given coactivator thus presenting difficulty in the development of selective synthetic amphipathic inhibitors.^{6,7} In Chapter 2, 34913-8 demonstrated selectivity for the

Med25•ATF6 α PPI over other coactivator•activator PPIs, highlighting the potential of using 34913-8 as a template to develop selective inhibitors for specific coactivators. Several experiments determined that the C-terminal carboxylic acid of 34913-8 contributed to this selectivity, proposing that specific residues within the peptide sequence of 34913-8 could also contribute to the selectivity of Med25•activator PPIs over other coactivator•activator PPIs.

Based on the results involving a negatively charged functional group in Chapter 2, the substitution analogs of the negatively charged amino acids from Figure 3.7 were further investigated to determine their selectivity for Med25 AcID PPIs. The five analogs with various substitutions at the charged residue positions were tested using fluorescence polarization binding assays against Med25 AcID•ETV5, CBPKIX•MLL, and CBPKIX•Myb (Figure 3.8A). These analogs showed no evidence of binding to CBPKIX•Myb so the relative K_i values were not represented in Figure 3.8. Results indicate that, similar to 34913-8, all these substitution analogs are selective for Med25•ATF6 α over Med25•ETV5, proposing that the negative charge of the C-terminal carboxylic acid of the lipopeptide is more important for this particular selectivity than the negatively charged residues. Additionally, 34913-Gln1Asn2 is the only analog where you see slight selectivity for the Med25•ATF6 α PPI over CBPKIXMLL. This further emphasizes that the peptide sequence of 34913-8 can be specifically modified to optimize for the selectivity of one coactivator over another, highlighting the hypothesis that sequence specificity of 34913-8 is crucial for the selective inhibition of coactivator•activator PPIs.



<u>Lipopeptide Analog</u>	<u>K_i (μM)</u>		
	<u>Med25•ATF6α</u>	<u>Med25•ERM</u>	<u>CBPKIX•MLL</u>
34913-Ala1	1.0 \pm 0.1	3.3 \pm 0.3	0.46 \pm 0.04
34913-Ala2	0.050 \pm 0.004	0.54 \pm 0.04	0.0045 \pm 0.0004
34913-Gln1	1.3 \pm 0.1	5.8 \pm 0.8	.24 \pm 0.01
34913-Asn2	0.71 \pm 0.04	2.6 \pm 0.1	0.50 \pm 0.04
34913-Gln1Asn2	0.090 \pm 0.002	0.88 \pm 0.08	0.16 \pm 0.01

Figure 3.8. Determination of the selectivity of charge substitution analogs. (A) Apparent EC_{50} values were determined through titration of each lipopeptide analog for Med25 AcID•ATF6 α , Med25 AcID•ERM, CBPKIX•MLL, and CBPKIX•Myb performed in experimental triplicate with the indicated error (SDOM). The EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of Med25 AcID•ATF6 α , Med25 AcID•ERM, CBPKIX•MLL, and CBPKIX•Myb using a K_i calculator.¹⁸ K_i values were then normalized against 34913-Ala1. (B) A compiled list of the substitution analogs with their relative K_i values and respective errors.

34913-8 shows selectivity for ETV1/5

Previous work by Henley, et al. revealed the importance of peptide sequence identity of TADs for their recognition and binding to Med25.¹⁶ The activators ETV4 and ETV1/5 vary only slightly in peptide sequence and have nearly identical dissociation constants, though ETV4 is able to engage with Med25 differently than ETV1/5. Lipopeptide 34913-8 was determined to inhibit Med25•ETV5 using competition fluorescence polarization in Chapter 2. Based on what we know about the ETVs binding to Med25, it is hypothesized

that 34913-8 could selectively inhibit one Med25•ETV PPI over the other (Figure 3.9A). Competition fluorescence polarization assays determined the inhibitory activity of lipopeptide 34913-8 for Med25•ETV PPIs (Figure 3.9B). These experiments demonstrated that 34913-8 selectively inhibits the Med25•ETV1/5 PPIs over the Med25•ETV4 PPI by ~2-fold, suggesting that the subset of residues that ETV1/5 interacts with on Med25 are more affected by lipopeptide binding than the ETV4 peptide (Figure 3.9C). This strongly supports the hypothesis that a probe can be utilized to selectively target one Med25•activator PPI over another, even if they are similar in sequence.

Based on the significant sequence similarities of ETV1 and ETV5, it was hypothesized that 34913-8 would inhibit the Med25•ETV1 and Med25•ETV5 PPIs with the same inhibitory activity. While the inhibitory activity for these two PPIs ($19 \pm 2 \mu\text{M}$ and $28 \pm 3 \mu\text{M}$) are similar compared to the Med25•ETV4 PPI, they were not within error of one another suggesting slight selectivity for the inhibition of the Med25•ETV1 PPI. This suggests the possibility that the two amino acid sequence variation between the TADs of ETV1 and ETV5 could be of importance for their molecular recognition of Med25. While 34913-8 does not demonstrate a particularly high degree of selectivity for one of these PPIs over the other, this difference in inhibition highlights the potential for lipopeptide 34913-8 to be further modified to optimize this selectivity for specific coactivator•activator PPIs.

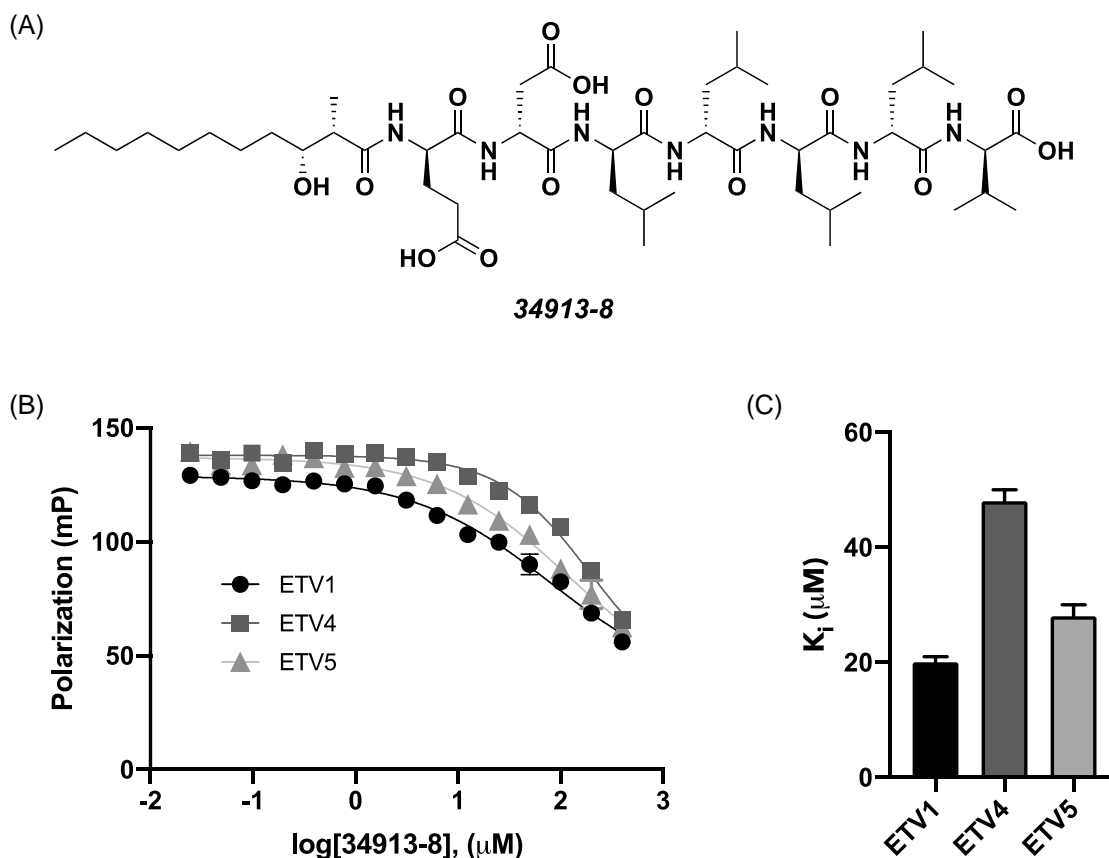


Figure 3.9. 34913-8 selectively inhibits Med25•ETV PPIs. (A) Lipopeptide 34913-8 is an effective inhibitor of Med25 PPIs and contains a D-amino acid heptapeptide N-terminally attached to a lipophilic tail. (B) Apparent EC_{50} values were determined through titration of 34913-8 for Med25 AcID•ETV PPIs performed in experimental triplicate with the indicated error (SDOM). (C) The EC_{50} values were converted to K_i values using the apparent K_d value based on the direct binding of the Med25 AcID•ETV PPIs using a K_i calculator.¹⁸ Results shown represent the average K_i s of two biological duplicates.

3.4 Conclusions and Future Directions

There are several proposed models that attempt to explain the interactions that occur between the TADs of transcriptional activators and the activator binding domains (ABD) of coactivators. The more traditional model is that these activators are driven predominantly by net negative charge with a lack of sequence specificity. Several lines of evidence point to this being inaccurate with the presence of key hydrophobic residues being required for transcriptional activity.^{11,12} The most recent studies from our lab have further expanded on these models of coactivator•activator PPIs by analyzing the

coactivator Med25 and its activator binding partners. It was revealed that Med25 was conformationally sensitive to slight sequence variations in the ETV family of transcriptional activators, suggesting that the TADs of the activators engage in specific interactions with Med25.¹⁶

The lipopeptide analog 34913-8, identified in Chapter Two, revealed a seven amino acid peptide sequence with similarities to the TAD of transcriptional activators that bind to Med25. This resulted in further analysis of the peptide sequence using alanine and neutrally charged amino acids to determine important residues for its inhibitory activity against the Med25•ATF6 α PPI. Several substitutions were made that identified individual residues within the peptide sequence of 34913-8 that were important for its inhibitory activity and further suggested that negative charge is not sufficient for inhibition. Additional analysis needs to be performed to determine if substitutions in these residue positions result in altered binding to Med25 AcID. Selectivity experiments determined that every substitution analog tested was selective for Med25•ATF6 α over Med25•ETV5, and one analog (34913-Gln1Asn2) conveyed selectivity for Med25•ATF6 α over CBPKIX•MLL. This suggests that the sequence specificity of the peptide sequence of 34913-8 is important for its role in selectivity, as well as activity. Altogether, this could lead to the synthesis of several lipopeptides designed to target specific coactivator•activator PPIs.

Further, in this Chapter, the lipopeptide 34913-8 was used to determine if a synthetic modulator could selectively inhibit one certain Med25•ETV PPI over another. 34913-8 demonstrated selectivity for the Med25•ETV1/5 PPIs over the Med25•ETV4 PPI, further establishing individual residue specificity as an important characteristic for Med25•activator PPIs. This also strongly supports the hypothesis that a probe can be

utilized to selectively target Med25, inhibiting a particular Med25•activator PPI and the expression of its related genes.^{18,19}

Several questions have arisen from the analysis of these lipopeptide analogs for Med25•ATF6 α PPI. While fluorescence polarization data may suggest that certain substitutions result in improved inhibitory activity, thus binding directly with Med25, it is known from Chapter 2 that inhibition does not always correlate to specific interactions and direct engagement with the target (i.e. analog 34913-9). To fully explore the effects these residues have on inhibitory activity, CETSA can be performed using both Med25 and CBP to determine if there is full length target engagement of the lipopeptide to these coactivator proteins. Additionally, ¹H, ¹³C-HSQC NMR can be performed using both Med25 AcID and CBP KIX to determine if there is direct binding to the target coactivator protein. These experiments could also determine if the analogs are actually binding to these ABDs, and if they are, it could reveal significant shift perturbations of unique Med25 residues, indicating certain lipopeptides can bind to distinct Med25 conformations.

3.5 Materials and Methods

Reagents and Instrumentation

Unless otherwise noted, chemical and biological reagents were obtained from commercial sources and were used without additional modification. Protein, peptide, and DNA concentrations were determined using a NanoDrop ND-1000 UV-Vis Spectrophotometer and/or a Beckman Spectrophotometer.

Protein expression and purification

Protocols were followed exactly as written in 2.5 Materials and Methods of this dissertation.

Solid-Phase synthesis and subsequent HPLC purification of activator peptides

Protocols were followed exactly as written in 2.5: Materials and Methods of this dissertation. Additional sequences and masses for activators previously used in Chapter 2 can be found in 2.5 Materials and Methods of this dissertation. Full characterization of activators, including HPLC analytical and mass spectrometry chromatograms, can be found in Appendix I.

Table 3.1. Sequence of fluorescein labelled ETV peptides

Peptide	Sequence (N'-C')
<i>ETV1 (38-69)</i>	FITC- β Ala-DLAHDSEELFQDLSQLQETWLAEAQVPDNDEQ
<i>ETV4 (45-76)</i>	FITC- β Ala-LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ

Table 3.2. Masses of fluorescein labelled ETV peptides

Peptide	Expected Mass (g/mol)	Observed Mass (g/mol)
<i>ETV1 (38-69)</i>	4171.7542	4171.766
<i>ETV4 (45-76)</i>	4157.7789	4157.7888

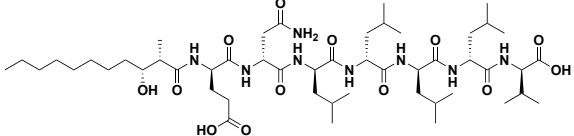
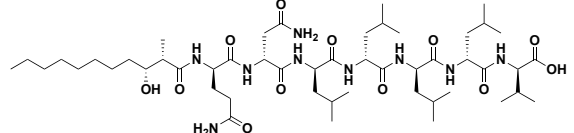
Solid phase synthesis and HPLC purification of lipopeptides analogs

Protocols were followed exactly as written in 2.5: Materials and Methods of this dissertation. Additional structures and masses for lipopeptides previously used in Chapter

2 can be found in 2.5 Materials and Methods of this dissertation. Full characterization of lipopeptide analogs, including HPLC analytical and mass spectrometry chromatograms, can be found in Appendix I. The HPLC analytical chromatograms determined an impurity in these dissolved samples eluting from the column ~17 min. The identity of this impurity was confirmed using mass spectrometry (further details in Appendix I).

Table 3.2. Structure of lipopeptide substitution analogs and their masses

Peptide	Structure	Expected Mass (g/mol)	Observed Mass
34913-Ala1		953.6413	953.6414
34913-Ala2		967.6569	967.6557
34913-Ala3		969.5998	969.5999
34913-Ala4		969.5998	969.6006
34913-Ala5		969.5998	969.5986
34913-Ala6		969.5998	969.5998
34913-Ala7		983.6155	983.6152
34913-Gln1		1010.6627	1010.6608

34913- Asn2		1010.6627	1010.6585
34913- Gln1-Asn2		1009.6787	1009.6769

Fluorescence Polarization/Competition Fluorescence Polarization

Protocols were followed exactly as written in 2.5 Materials and Methods of this dissertation. All values shown represent the average K_i s of two biological duplicates.

3.6 References

- 4 Ma, J., Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell*. **1987**, *48*(5), 847-853.
- 5 Hope, I.A., Struhl, K. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell*. **1986**, *46*(6), 885-894
- 6 Hope, I.A., Mahadevan, S., Struhl, K. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature*. **1998**, *333*(6174), 635–640.
- 7 Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature*. **1988**, *333*(6170), 210-212.
- 8 Sadowski, I., Ma, J., Triezenberg, S., Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature*. **1988**, *335*(6190), 563-564.
- 9 Brzovic, P.S., Heikaus, C.C., Kisselev, L., et al. The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell*. **2011**, *44*(6), 942-953.
- 10 Tuttle, L.M., Pacheco, D., Warfield, L., et al. Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep*. **2018**, *22*(12), 3251-3264.
- 11 Drysdale, C.M., Dueñas, E., Jackson, B.M., et al. The transcriptional activator GCN4 contains multiple activation domains that are critically dependent on hydrophobic amino acids. *Mol Cell Biol*. **1995**, *15*(3), 1220-1233.
- 12 Warfield, L., Tuttle, L.M., Pacheco, D., et al. A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci*. **2014**, *111*(34), 3506-3513.
- 13 Sainz, M.B., Goff, S.A., Chandler, V.L. Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo. *Mol Cell Biol*. **1997**, *17*(1), 115-122.
- 14 Cress, W.D., Triezenberg, S.J. Critical structural elements of the VP16 transcriptional activation domain. *Science*. **1991**, *251*(4989), 87-90.
- 15 Regier, J.L., Shen, F., Triezenberg, S.J. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci*. **1993**, *90*(3), 883-887.

- 16 Giniger, E., Ptashne, M.. Transcription in yeast activated by a putative amphipathic alpha helix linked to a DNA binding unit. *Nature*. **1987**, 330(6149), 670-672
- 17 Ma J, Ptashne M. A new class of yeast transcriptional activators. *Cell*. **1987**, 51(1), 113-119.
- 18 Sugase K, Dyson HJ, Wright PE. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature*. 2007;447(7147):1021-1025.
- 19 Henley, M.J., Linhares, B.M., Morgan, B.S., et al. Unexpected specificity within dynamic transcriptional protein-protein complexes. *Proc. Natl. Acad. Sci*. **2020**, 117(44), 27346-27353.
- 20 Henderson, A.R., Henley, M.J., Foster, N.J., et al. Conservation of coactivator engagement mechanism enables small molecule allosteric modulators. *Proc. Natl. Acad. Sci*. **2018**, 115(36), 8960-8965.
- 21 Zaneta Nikolovska-Coleska, Renxiao Wang, Xueliang Fang, Hongguang Pan, York Tomita, Peng Li, Peter P. Roller, Krzysztof Krajewski, Naoyuki Saito, Jeanne Stuckey and Shaomeng Wang, "Development and Optimization of a Binding Assay for the XIAP BIR3 Domain Using Fluorescence Polarization", *Analytical Biochemistry*, **2004**.
- 22 Sela, D., Conkright, J.J., Chen, L., et al. Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6 α . *J. Biol. Chem*. **2013**, 288(36), 26179-26187.
- 23 Landrieu, I., Verger, A., Baert, J.L., et al. Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res*. **2015**, 43(14), 7110-7121.
- 24 Vojnic, E., Mourão, A., Seizl, M., et al. Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol*. **2011**, 18(4), 404-409.
- 25 Bontems, F., Verger, A., Dewitte F, et al. NMR structure of the human Mediator MED25 ACID domain. *J. Struct. Biol*. **2011**, 174(1), 245-251.
- 26 Milbradt, A.G., Kulkarni, M., Yi, T., et al. Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol*. **2011**, 18(4), 410-415.
- 27 Lee, M.S., Lim, K., Lee, M.K., Chi, S.W. Structural Basis for the Interaction between p53 Transactivation Domain and the Mediator Subunit MED25. *Molecules*. **2018**, 23(10), 2726.
- 28 Haze, K., Yoshida, H., Yanagi, H., Yura, T., Mori, K. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol. Biol. Cell*. **1999**, 10(11), 3787–3799.

- 29 Yoshida, H., Okada, T., Haze, K., et al. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol. Cell. Biol.* **2000**, 20(18), 6755–6767.
- 30 Wang, Y., Shen, J., Arenzana, W., et al. Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. *J. Biol. Chem.* **2000**, 275(35), 27013–27020.
- 31 Fernandez, P. M., Tabbara, S.O., Jacobs, L.K., et al. Overexpression of the glucose-regulated stress gene GRP78 in malignant but not benign human breast lesions. *Breast Cancer Res. Treat.* **2000**, 59(1), 15–26.
- 32 Benz, C.C., O'Hagan, R.C., Richter, B., et al. HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene.* **1997**, 15(13), 1513–1525.
- 33 de Launoit, Y., Chotteau-Lelievre, A., Beaudoin, C., et al. The PEA3 group of ETS-related transcription factors. *Adv. Exp. Med. Biol.* **2000**, 480, 107–116.
- 34 Pellecchia, A., Pescucci, E., De Lorenzo, E., et al. Overexpression of ETV4 is oncogenic in prostate cells through promotion of both cell proliferation and epithelial to mesenchymal transition. *Oncogenesis.* **2012**. 1(7), e20.

Chapter 4

Conclusions and Future Directions

4.1 Summary

Historically, targeting protein-protein interactions (PPIs) with small molecule modulators has been a significant challenge. This has been particularly true of coactivator•activator PPIs that are often dynamic and have weak binding affinities and shallow binding surfaces.^{1,2} There have been several successful natural product inhibitors that have been discovered to effectively target these coactivator•activator PPIs, with a number coming from our lab, demonstrating the potential of these molecules.^{3,4}

This dissertation sought to leverage a natural lipopeptide, 34913-1, and its analogs to develop an effective and selective inhibitor that targets the activator interaction domain (AcID) of the coactivator, Med25, and its activator binding partners, ATF6 α and the ETVs. These Med25 PPIs are particularly interesting due to their implications in several diseases, including cancer, developmental disorders, and diseases of protein homeostasis.⁵⁻⁷ To first address this goal, a lipopeptide natural product was divided into three different components that could be independently modified to determine their functional necessity. Through this process, two lead analogs, 34913-8 and 34913-9, were identified that demonstrated comparable potency to the parent lipopeptide for the Med25•ATF6 α PPI, though 34913-8 exhibited selectivity for Med25 AcID *in vitro* and full length Med25 in mammalian cell extracts. Following these analyses, several substitution analogs were synthesized to determine the importance of specific residues in 34913-8 for

binding and inhibition of Med25. similarities in inhibition between our lipopeptide analogs and the TADs of transcriptional activators. This revealed the sequence specificity in lipopeptide analogs binding to the Med25•ATF6 α PPI and further supported the notion that these lipopeptides can be modified to selectively target certain coactivator•activator PPIs over others.

4.2 Conclusions

The bulk of this thesis, as described in Chapters Two and Three, have utilized lipopeptide analogs to better understand the structural characteristics necessary for a small molecule to inhibit the coactivator Med25, with the larger goal of targeting ‘undruggable’ coactivator•activator PPIs. Chapter Two focuses on the analysis of a lipopeptide natural product, 34913-1, to determine the importance of certain functional elements for the inhibition of the Med25•ATF6 α PPI. This compound was separated into three main components: the lipophilic tail, the peptide sequence, and the C-terminus. It was determined that the lipophilic tail and peptide sequence are critical for inhibition of the Med25•ATF6 α PPI, while branching of the lipophilic tail may be important for the solubility of these compounds.

This functional analysis guided the synthesis of two lead lipopeptides, 34913-8 and 34913-9, with comparable inhibitory activity to the parent lipopeptide. These two compounds vary solely in their C-terminus with one containing a C-terminal carboxylic acid and the other containing a C-terminal amide, respectively. HSQC NMR experiments showed that 34913-8 binds to the H2 binding face of Med25 AcID, like ATF6 α ⁸, while 34913-9 does not appear to bind at all suggesting that the C-terminal carboxylic acid

offers specificity of the lipopeptide for Med25 AcID. This hypothesis was further supported when 34913-8 demonstrated selectivity for the Med25 PPIs over other CBP KIX•activator PPIs using competition fluorescence polarization assays, while 34913-9 did not. Additionally, 34913-8 demonstrated engagement with full-length Med25 using mammalian cell extracts and effective inhibition of the Med25 PPIs in a cellular context, while 34913-9 did not. This data reveals 34913-8 as an effective and selective inhibitor of the Med25•ATF6 α PPI, with the C-terminal carboxylic acid conveying specificity for Med25.

Chapter Three focuses on the several hydrophobic and acidic residues of the peptide sequence of the lipopeptide, resembling the TAD of transcriptional activators.¹⁰⁻¹⁷ Due to their shared amphipathic nature, we hypothesized that the lipopeptide may interact with Med25 AcID similarly to these activators due to their shared amphipathic nature, therefore alanine scanning was used to identify key residues important for inhibition of the Med25•ATF6 α PPI. It was observed that an alanine residue substitution at several residue positions within the peptide sequence resulted in a loss of inhibitory activity, while other residue positions saw increased inhibitory activity. Taken, together, this confirms that individual residues of the peptide sequence of 34913-8 are important for its inhibition of Med25•ATF6 α while others are not, opposing the traditional binding model suggesting that these coactivator•activator PPIs are driven primarily by non-specific interactions. This also suggests that certain residue positions are more sensitive to structural changes than others, which can be used to guide the design of further lipopeptide analogs.

By understanding the importance of individual elements of both the lipopeptide and Med25 for inhibition, we hope to develop more potent and selective inhibitors to target

Med25, as well as other coactivator•activator PPIs. To further investigate the importance of negatively charged residues in the peptide sequence, neutrally charged substitution analogs were synthesized. These analogs resulted in varying degrees of inhibitory activity, suggesting that further experiments would need to be performed to gather a complete evaluation of the role these negative charges have on inhibition of the Med25•ATF6 α PPI. Additionally, these neutrally charged substitution analogs were further investigated for their selectivity for Med25•PPIs over other coactivator•activator PPIs. It was observed that none of these neutrally charged analogs demonstrated selectivity for Med25 AcID PPIs over CBP KIX PPIs except 34913-Gln1Asn2. Further 34913-8 was able to selectively inhibit the Med25•ETV1/5 PPIs over the Med25•ETV4 PPI, despite these three activators varying only slightly in sequence, suggesting that 34913-8 can be used as a probe to selectively target one Med25•activator PPI over another. Overall, the data collected in Chapter Three of this dissertation suggests that 34913-8 acts as a transcriptional domain mimic, further demonstrating its potential for the successful inhibition of the Med25 PPIs. Additionally, these analogs can inform the design of additional analogs that can be applied to selectively inhibit Med25•PPIs as well as target several other coactivator•activator PPIs synthetically.

4.3 Future Directions

Further dissection of the lipopeptide peptide sequence

The alanine scan and neutrally charged substitution analogs introduced in Chapter Three demonstrated the importance of individual residues in the inhibitory activity of these lipopeptides for the Med25•ATF6 α PPI. Additionally, several lines of evidence suggest the importance of aromatic residues for the inhibitory activity of TADs of transcriptional

activators suggesting incorporation of these types of residues could improve binding affinity and/or selectivity for the coactivator target.^{16,17} Future work could focus on the synthesis of a small lipopeptide library featuring amino acid variations at identified residues, with an emphasis on the incorporation of aromatic and small hydrophobic residue side chains, such as isoleucine, leucine, valine, phenylalanine, tryptophan, and tyrosine. The affinity and selectivity of these lipopeptides can then be tested against the Med25•activator PPIs as well as the CBP KIX•activator PPIs. Further, analogs showing improved inhibition and/or selectivity would be subjected to HSQC NMR experiments to determine the binding site of these analogs and cellular thermal shift assays to determine engagement with full length Med25. This library could further explore the role that structural changes in the peptide sequence would have on the inhibitory activity of these lipopeptides. Results could highlight specific structural features that may be important for inhibition or selectivity, which could result in the development of more successful lipopeptide analogs to target Med25 PPIs.

Diversification of lipophilic tails

It was determined within this dissertation that a branched lipophilic tail similar to the one present in the parent lipopeptide, 34913-1, was important for successful inhibition of the Med25•ATF6 α PPI. The lipophilic tail, (2S, 3R)-3-hydroxy-2-methylundecanoic acid, has defined stereochemistry, though the importance of this stereochemistry for inhibition was not fully understood. Synthesis of additional lipophilic tails, including the enantiomer of (2S, 3R)-3-hydroxy-2-methylundecanoic acid and its two diastereomers

(Figure 4.1), could provide further detail. Additionally, these new lipophilic tails can also explore the importance of stereochemistry in compound selectivity.

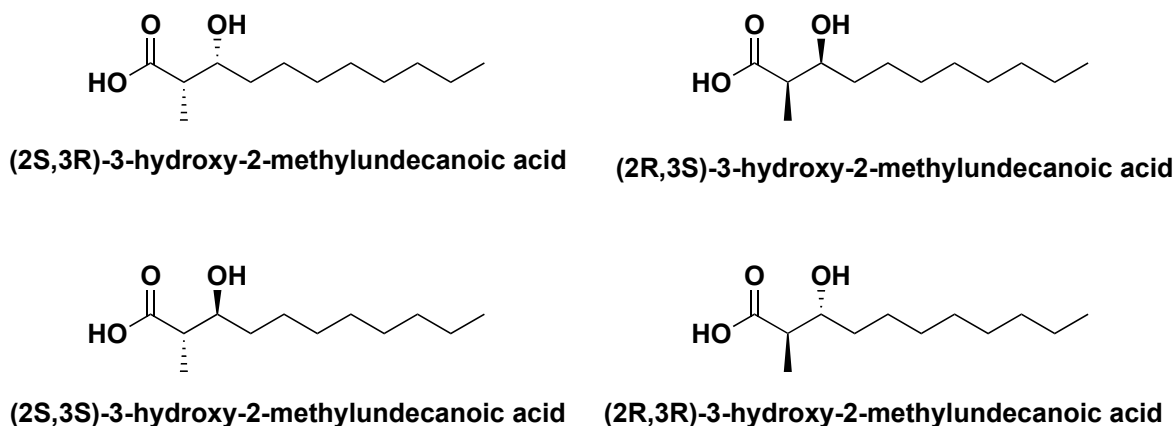


Figure 4.1. Synthesis of lipid tail stereoisomers. The structure of the current lead lipid tail used in this Dissertation is shown in the top left. The synthesis of the enantiomer of (2S,3R)-3-hydroxy-2-methylundecanoic acid (shown in the top right) and its two diastereomers (bottom) can be synthesized to investigate the importance of stereochemistry on inhibition of the Med25•ATF6 α PPI.

Additional fatty acyl (FA) tails can also be synthesized to further investigate the effect that acyl chain length, conjugation, additional hydroxylation, or cyclic rings has on the potency and selectivity of these lipopeptides for Med25. The FA tail of 34913-1 resembles several lipids of lipopeptides found in nature, suggesting several potential lipophilic tails that can be synthesized and tested for activity when coupled to the D-peptide sequence, including those present in gageotetrin C, jahanyne, and carmabin A.¹⁸⁻
²¹ These lipid tail derivatives can be used to determine the reactivity of various changes made in this structural component of the lipopeptide, providing insight on how these changes may affect its binding to Med25 AcID. This may lead to the identification of promising lipophilic tails that can further inform the design and synthesis of more potent and selective inhibitors of these coactivator•activator PPIs.

A high-resolution structural model of Med25 with 34913-8

Several lines of evidence indicate that the tail module of the Mediator complex contains several exchangeable and conformationally dynamic subunits that interact primarily with activators and repressors.²²⁻²⁴ Most recently, Zhao, et al. reported a 4 Å cryo-EM structure that provides a more detailed model of the Tail module of mMed, providing key insights into the protein interactions of these subunits. This study showed that the subunit Med25 is not essential for the interaction of the Tail module with core mMed or for this module's integrity.²⁵ Additionally, it is known that the well-folded C-terminal AcID domain of Med25 is separated from the N-terminal VWA domain by an intrinsically disordered region (IDR), which has been suggested to play a role in the recruitment of transcription factors to active genes.²⁵⁻²⁸ Taken together, this further supports the hypothesis that the larger coactivator Med25 undergoes universal conformational changes when bound to specific activator proteins which allows for the activation of transcription, though this has not been confirmed due to a lack of high-resolution structural data.^{22, 25}

Studies have indicated Med25 and its activator binding partners play a significant role in the dysregulation of gene expression implicated in several diseases.⁴⁻⁷ Due to this importance, an understanding of the full structural dynamics and global architecture of Med25 would be valuable, though it is difficult to characterize structurally. Recently, cryo-EM has emerged as a technique that is well-suited for the structural analysis of dynamic protein complexes, such as Med25.²⁹ Throughout this dissertation, lipopeptide 34913-8 was selective and effective inhibitor of Med25 making it an attractive synthetic ligand to complex with Med25 to determine the full structural aspects of Med25 interactions. This

will allow us to understand the global architectural changes that contribute to the mechanism and regulation of the complex, ideally resulting in the identification of more promising regions within Med25 to modulate synthetically.

4.4 References

1. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **2015**, *11*(12), 891–894.
2. Cesa, L. C., Mapp, A. K., & Gestwicki, J. E. Direct and Propagated Effects of Small Molecules on Protein–Protein Interaction Networks. *Front. Bioeng. Biotechnol.* **2015**, *3*(119).
3. Majmudar, C. Y. et al. Sekikaic Acid and Lobaric Acid Target a Dynamic Interface of the Coactivator CBP/p300. *Angew. Chem. Int. Ed.* **2012**, *51*(45), 11258–11262.
4. Garlick, J.M., Sturlis, S.M., Bruno, P.A. et al. Norstictic Acid Is a Selective Allosteric Transcriptional Regulator. *J. Am. Chem. Soc.* **2021**, *143*(25), 9297-9302.
5. Bushweller, J. H. Targeting Transcription Factors in Cancer — from Undruggable to Reality. *Nat. Rev. Cancer.* **2019**, *19*(11), 611–624.
6. Brennan, P., Donev, R., Hewamana, S. Targeting Transcription Factors for Therapeutic Benefit. *Mol. Biosyst.* **2008**, *4*(9), 909–919.
7. Lambert, M; Jambon, S.; Depaw S.; David-Cordonnier M.H. Targeting Transcription Factors for Cancer Treatment. *Molecules*, **2018**, *23*(6), 1479.
8. Henderson, A. R. et al. Conservation of coactivator engagement mechanism enables small molecule allosteric modulators. *Proc. Natl. Acad. Sci.* **2018**, *115*(36), 8960-8965.
9. Henley, M.J., Linhares, B.M., Morgan, B.S., Cierpicki, T., Fierke, C.A., Mapp, A.K. Unexpected specificity within dynamic transcriptional protein-protein complexes. *Proc. Natl. Acad. Sci.* **2020**, *117*(44), 27346-27353.
10. Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature.* **1988**, *333*(6170), 210-212.
11. Brzovic, P.S, Heikaus CC, Kisselev L, et al. The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell.* **2011**, *44*(6), 942-953.
12. Hope, I., Mahadevan, S. & Struhl, K. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature.* **1988**, *333*(6174), 635–640.
13. Hope, I.A., Struhl, K. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell.* **1986**, *46*(6), 885-894.

14. Sadowski, I., Ma, J., Triezenberg, S., Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature*. **1988**, 335(6190), 563-564.
15. Ma, J. & Ptashne, M. Deletion Analysis of GAL4 Defines Two Transcriptional Activating Segment. *Cell*. **1987**, 48(5), 847-853..
16. Cress, W.D., Triezenberg, S.J. Critical structural elements of the VP16 transcriptional activation domain. *Science*. **1991**, 251(4989), 87-90.
17. Regier, J.L., Shen, F., Triezenberg, S.J. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc Natl Acad Sci U S A*. **1993**, 90(3), 883-887.
18. Tareq, F.S., Lee, M.A., Lee, H.S., Lee, Y.J., Lee, J.S., Hasan, C.M., Islam, Md.T., Shin, H.J. Gageotettrins A–C, Noncytotoxic Antimicrobial Linear Lipopeptides from a Marine Bacterium *Bacillus Subtilis*. *Org. Lett.* **2014**, 16(3), 928–931.
19. Iwasaki, A., Ohno, O., Sumimoto, S., Ogawa, H., Nguyen, K.A., Suenaga, K. Jahanyne, an Apoptosis-Inducing Lipopeptide from the Marine Cyanobacterium *Lyngbya* Sp. *Org. Lett.* **2015**, 17(3), 652–655.
20. Chooi, Y.H., Tang, Y. Adding the Lipo to Lipopeptides: Do More with Less. *Chem. Biol.* **2010**, 17(8), 791–793.
21. Huber, F.M., Pieper, R.L., Tietz, A.J. The Formation of Daptomycin by Supplying Decanoic Acid to *Streptomyces Roseosporus* Cultures Producing the Antibiotic Complex A21978C. *J. Biotechnol.* **1988**, 7(4), 283–292.
22. El Khattabi, L., Zhao, H., Kalchschmidt, J. et al. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell*. **2019**, 178(5), 1145-1158.
23. Asturias, F.J., Jiang, Y.W., Myers, L.C., Gustafsson, C.M., Kornberg, R.D. Conserved structures of mediator and RNA polymerase II holoenzyme. *Science*. **1999**, 283(5404), 985–987.
24. Taatjes, D.J., Tjian, R. Structure and function of CRSP/Med2; a promoter-selective transcriptional coactivator complex. *Mol. Cell*. **2004**, 14(5), 675–683.
25. Zhao, H., Young, N., Kalchschmidt, J. et al. Structure of mammalian Mediator complex reveals Tail module architecture and interaction with a conserved core. *Nat. Commun.* **2021**, 12(1), 1355.
26. Sabari, B.R., Dall’Agnese, A., Boija, A., et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science*. **2018**, 361(6400), 3958.

27. Cho, W. K., Spille, J-H., Hecht, M., et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*. **2018**, 361(6400), 412–415.
28. Chong, S., Dugast-Darzacq, C., Liu, Z., et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science*. **2018**, 361(6400), 2555.
29. Takizawa, Y., Binshtein, E., Erwin, A.L., et al. While the revolution will not be crystallized, biochemistry reigns supreme. *Protein Sci*. **2017**, 26(1), 69-81.

Appendix A

Peptide and Lipopeptide Characterization

This appendix contains analytical HPLC and LC-MS chromatograms for the activator synthesized peptides and lipopeptide analogs used throughout this thesis.

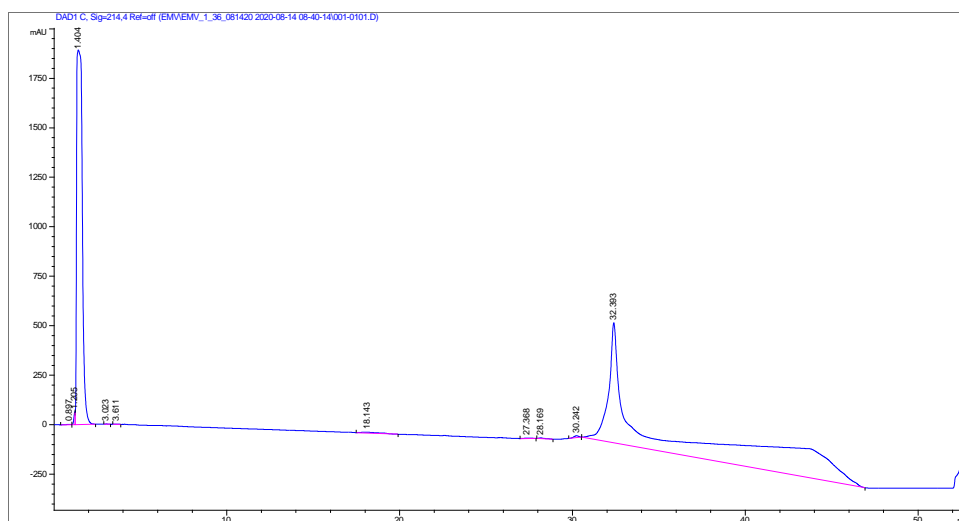
A.1 Activator Peptide Characterization

All HPLC chromatograms of fluorescein labeled peptides were monitored at both 214 and 280 nm.

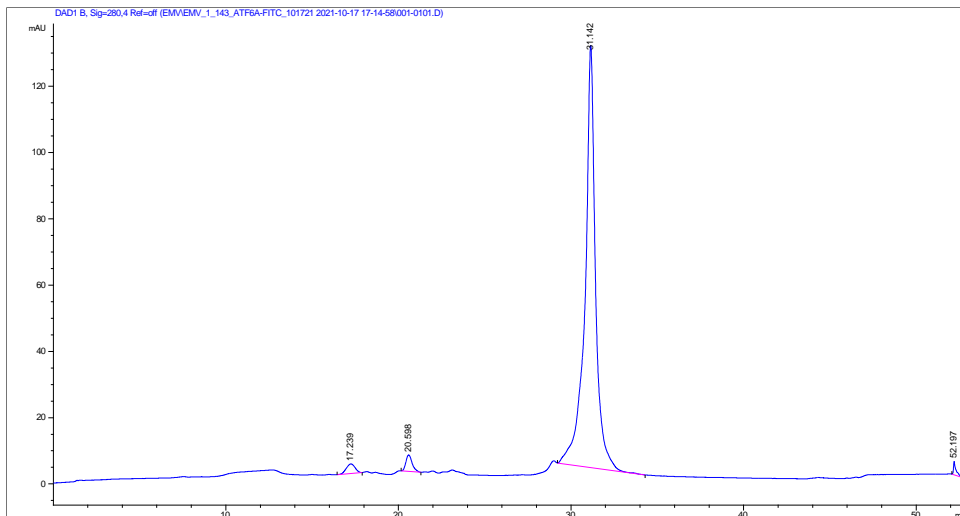
Fluorescein-labeled ATF6 α (38-75)

Sequence: N'-FITC- β Ala-FTDTDELQLEAANETYENNFDNLDLDFLDLMPWESDIWD-C'

HPLC:

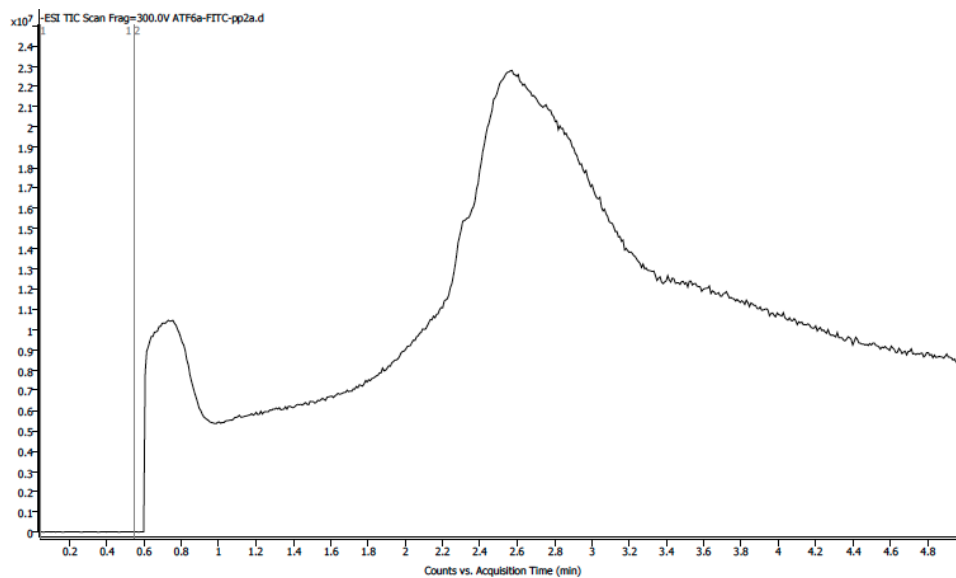


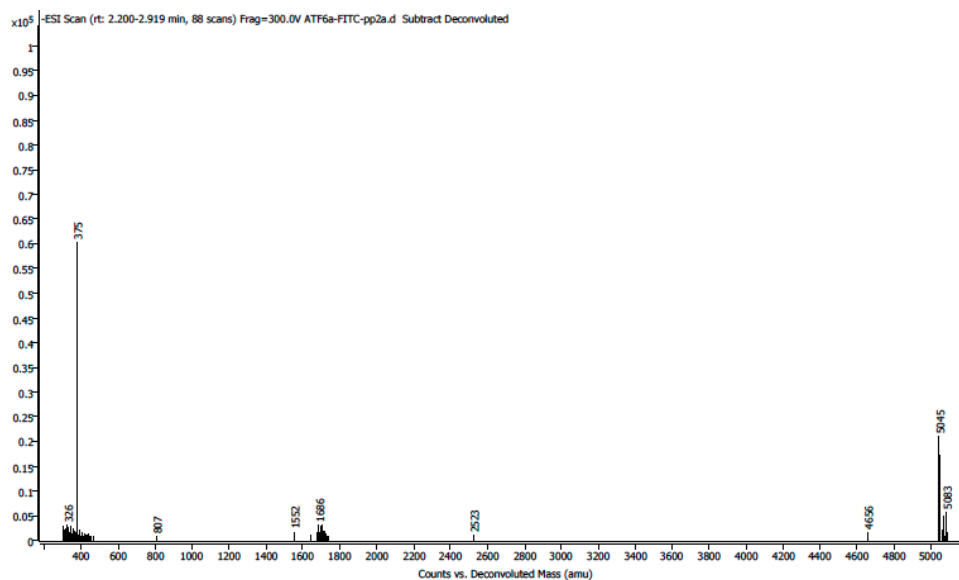
peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the activator peptide sample.



Analytical HPLC chromatogram of **FITC-ATF6 α (38-75)** monitored at 280 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

LC-MS:



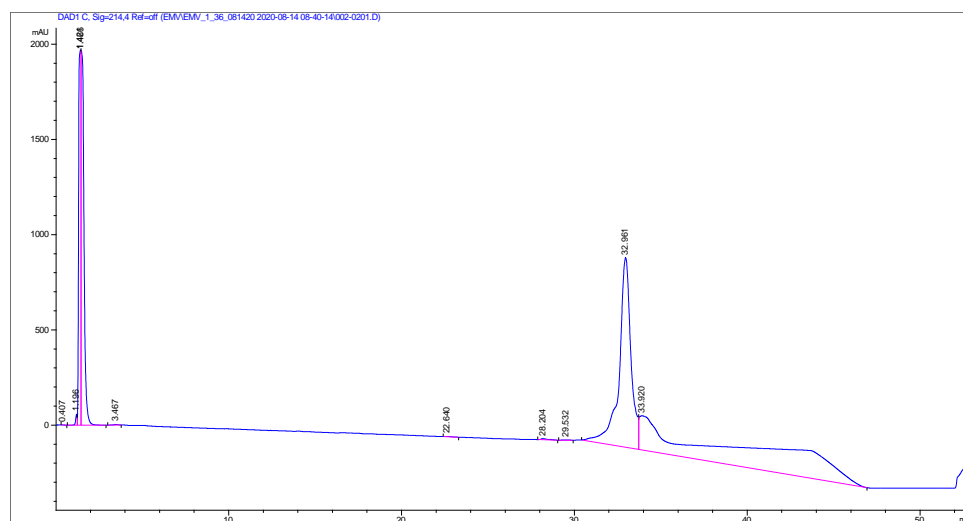


LC-MS of **FITC-ATF6 α (38-75)** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

Fluorescein-labeled ETV5 (38-68)

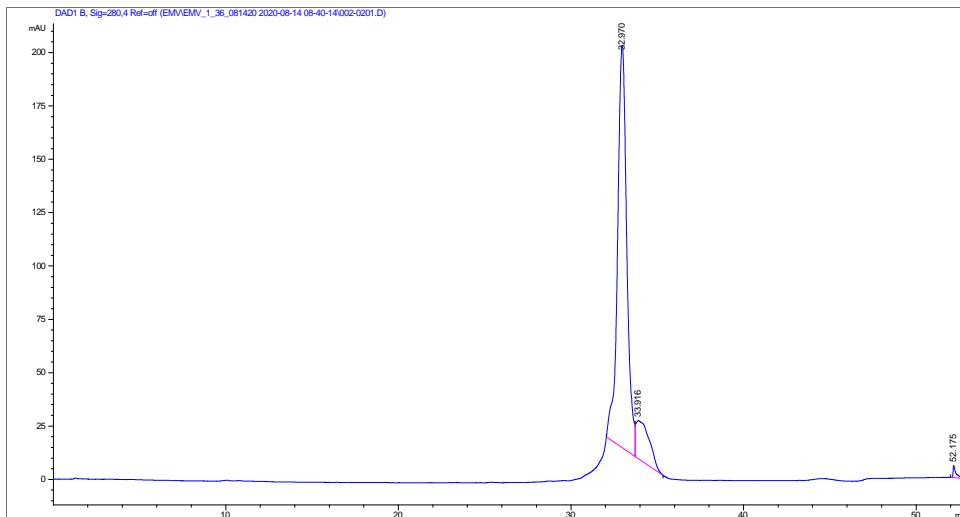
Sequence: N'-FITC- β Ala-ADLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ-C'

HPLC:



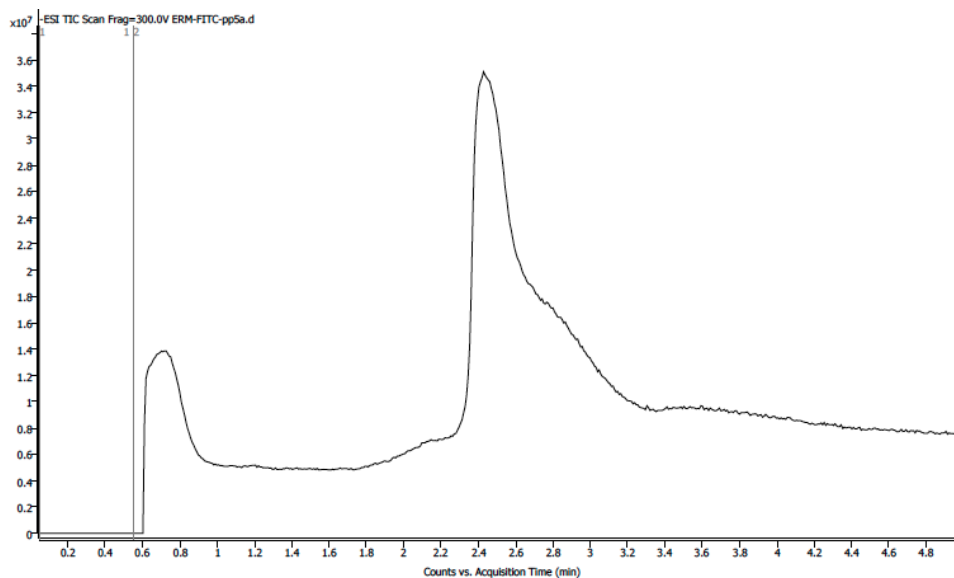
Analytical HPLC chromatogram of **FITC-ETV5 (38-68)** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large

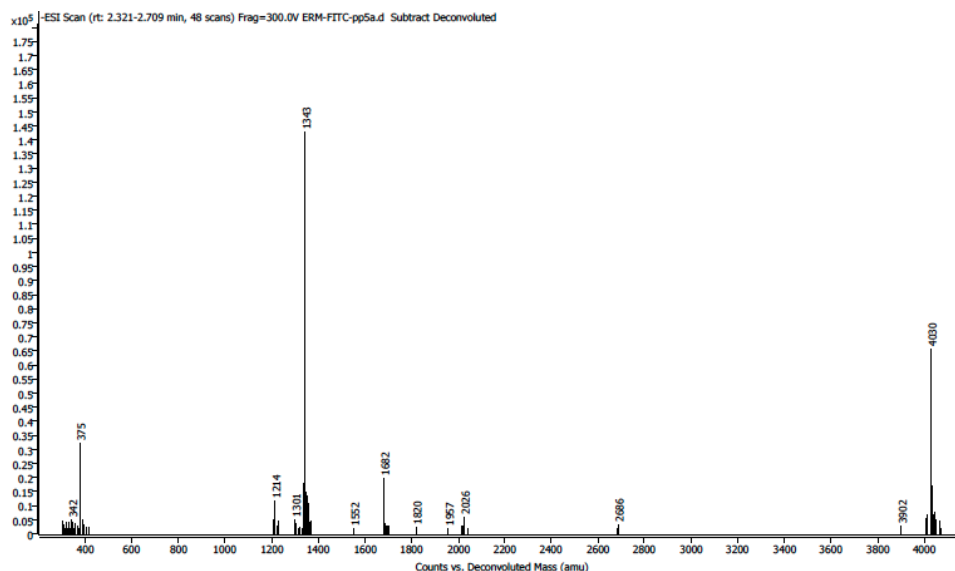
peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the activator peptide sample.



Analytical HPLC chromatogram of **FITC-ETV5 (38-68)** monitored at 280 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

LC-MS:



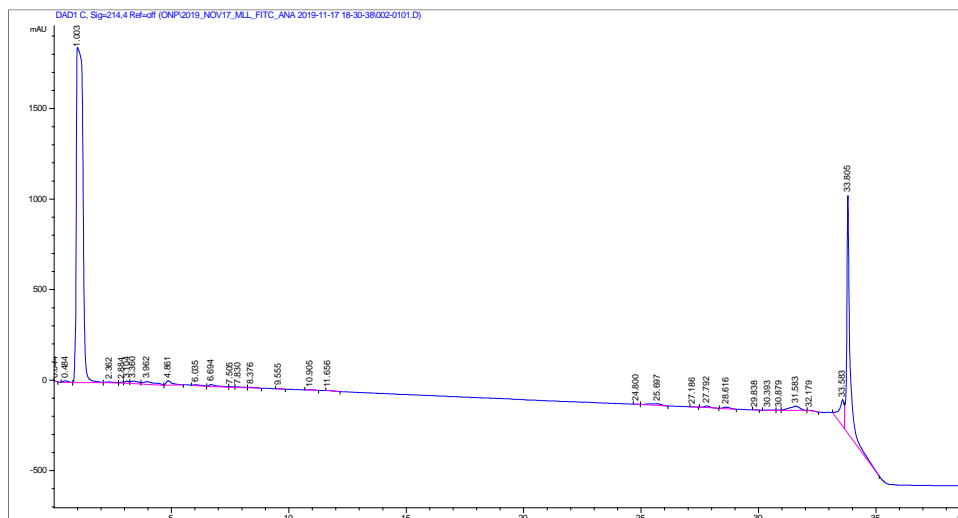


LC-MS of **FITC-ERM (38-68)** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

Fluorescein-labeled MLL (2840-2858)

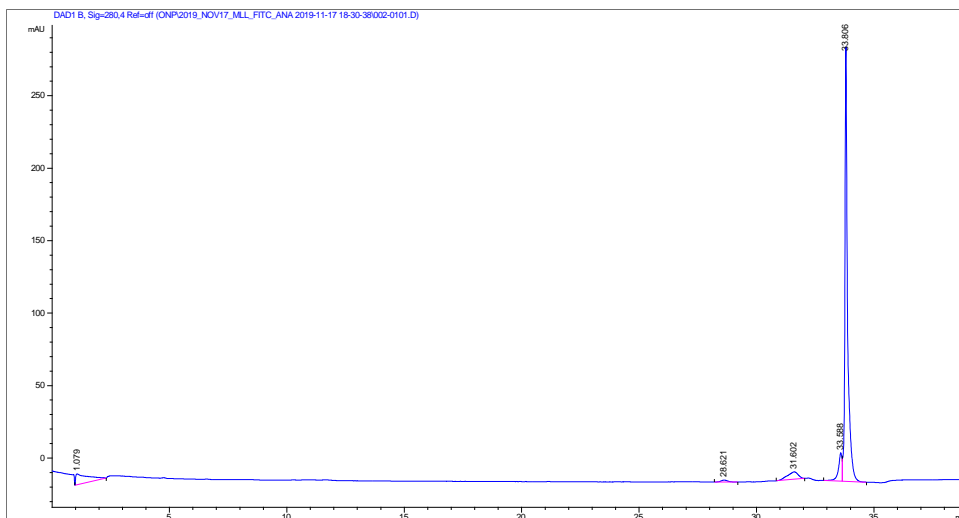
Sequence: N'-FITC-βAla-DCGNILPSDIMDFVLKNTP-C'

HPLC:



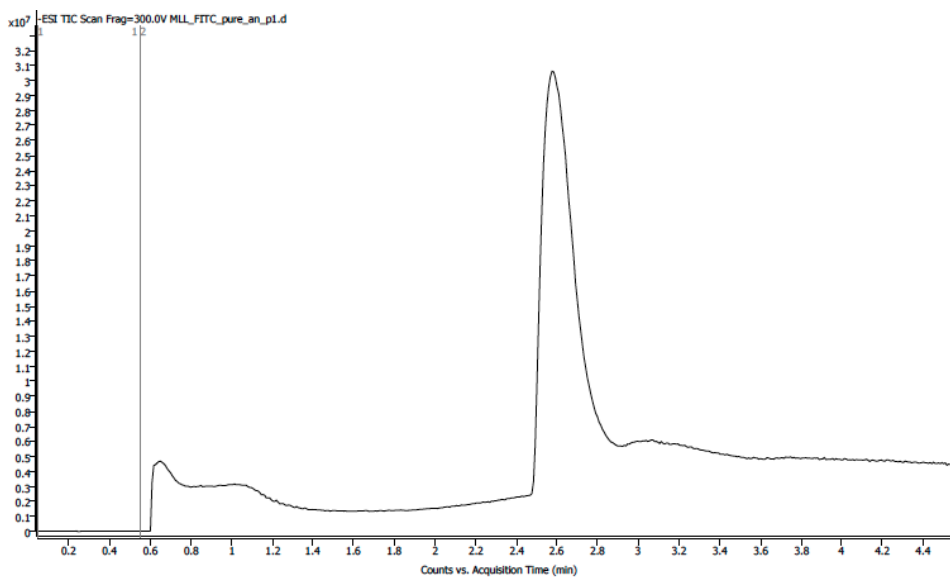
Analytical HPLC chromatogram of **FITC-MLL (2840-2858)** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

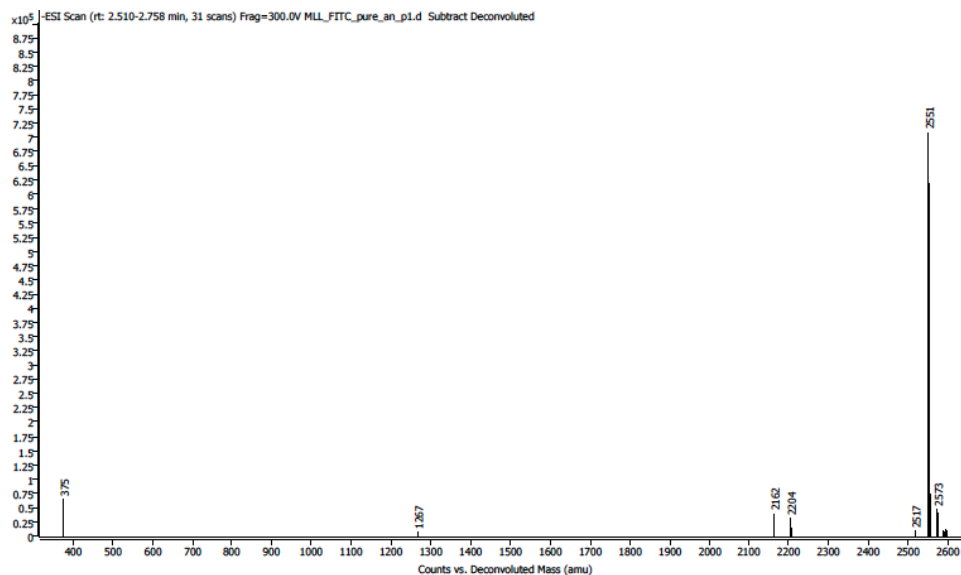
The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the activator peptide sample.



Analytical HPLC chromatogram of **FITC-MLL (2840-2858)** monitored at 280 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

LC-MS:



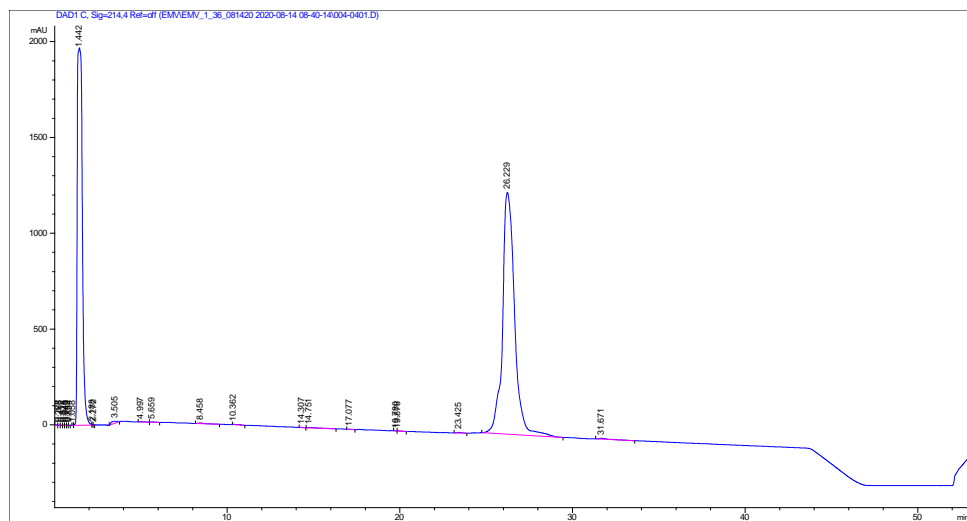


LC-MS of **FITC-MLL (2840-2858)** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

Fluorescein-labeled Myb (219-316)

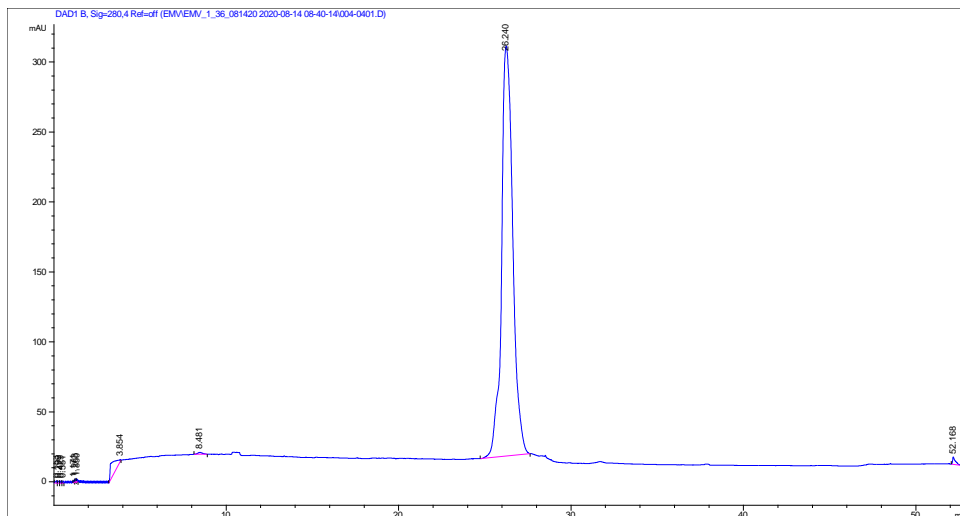
Sequence: N'-FITC-βAla-KEKRIKELELLLMSTENELKGQQVLP-C'

HPLC:



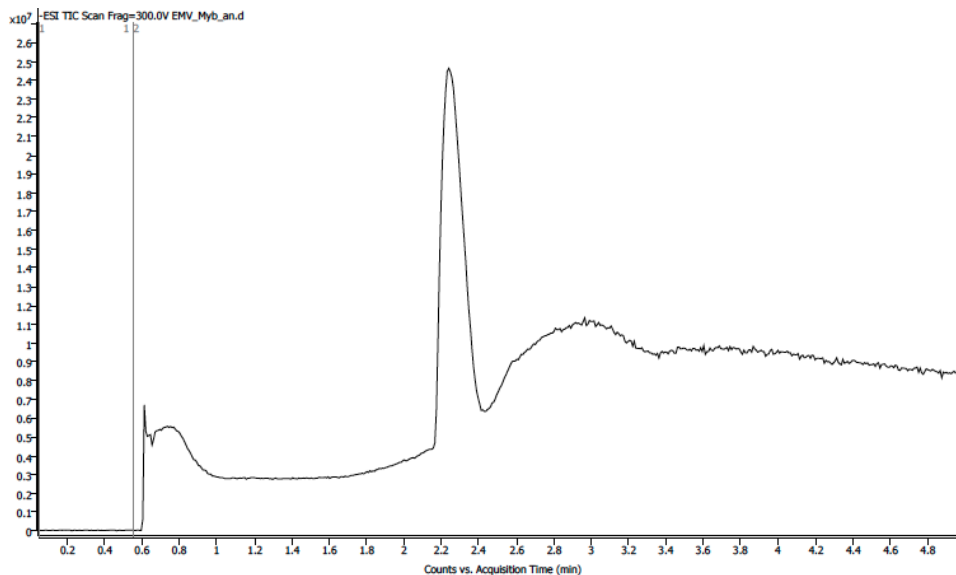
Analytical HPLC chromatogram of **FITC-Myb (219-316)** monitored at 214 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large

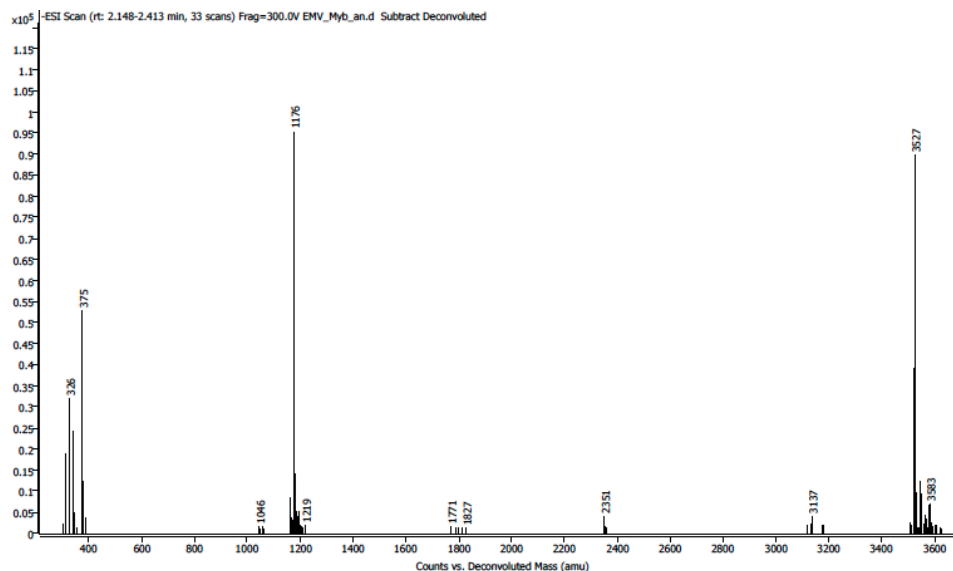
peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the activator peptide sample.



Analytical HPLC chromatogram of **FITC-Myb (219-316)** monitored at 280 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

LC-MS:



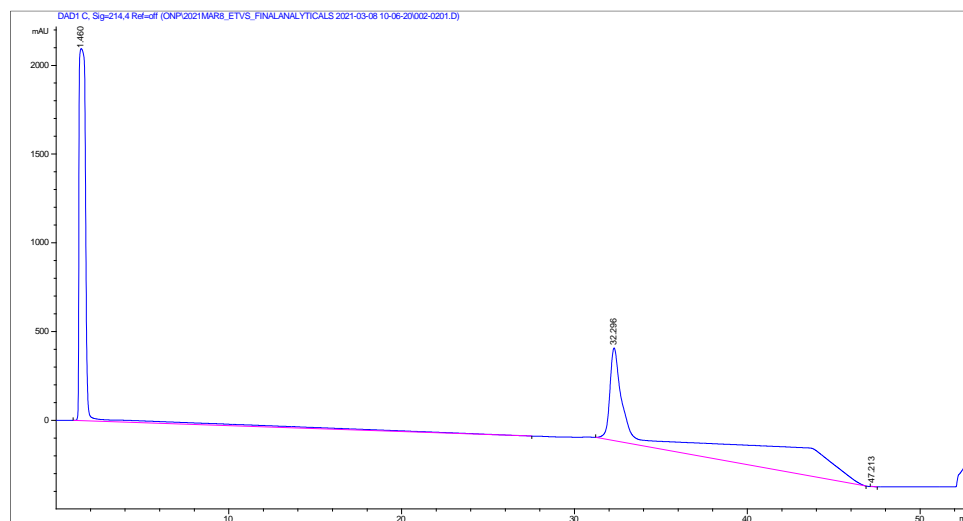


LC-MS of **FITC-Myb (219-316)** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

Fluorescein-labeled ETV1 (38-69)

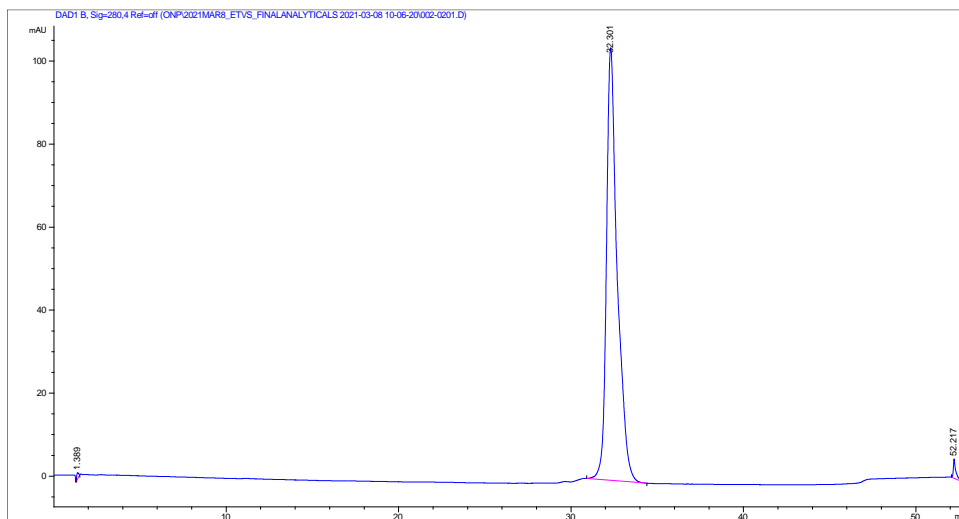
Sequence: N'-FITC-βAla-DLAHDSEELFQDLSQLQETWLAEAQ/VPDNDEQ-C'

HPLC:



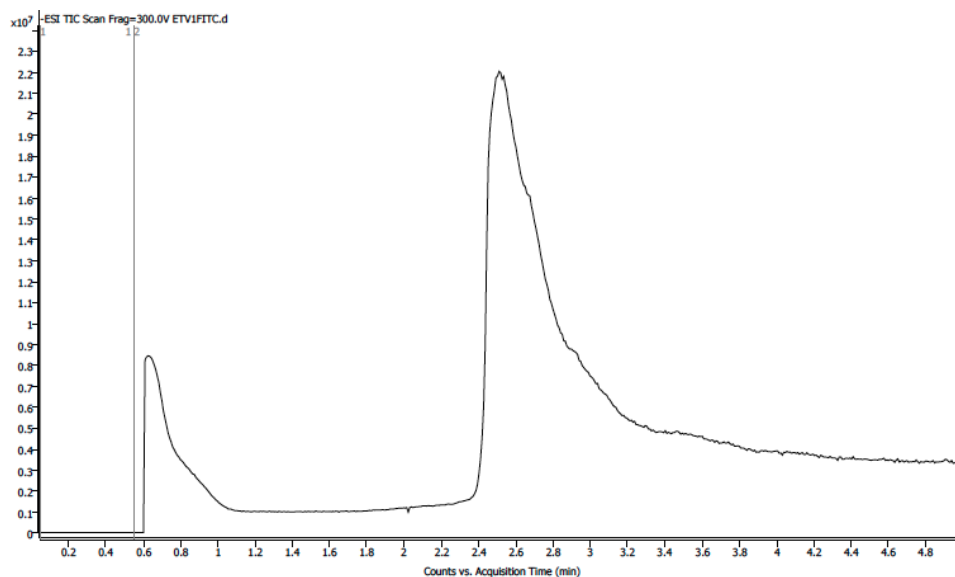
Analytical HPLC chromatogram of **FITC-ETV1 (38-69)** monitored at 214 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large

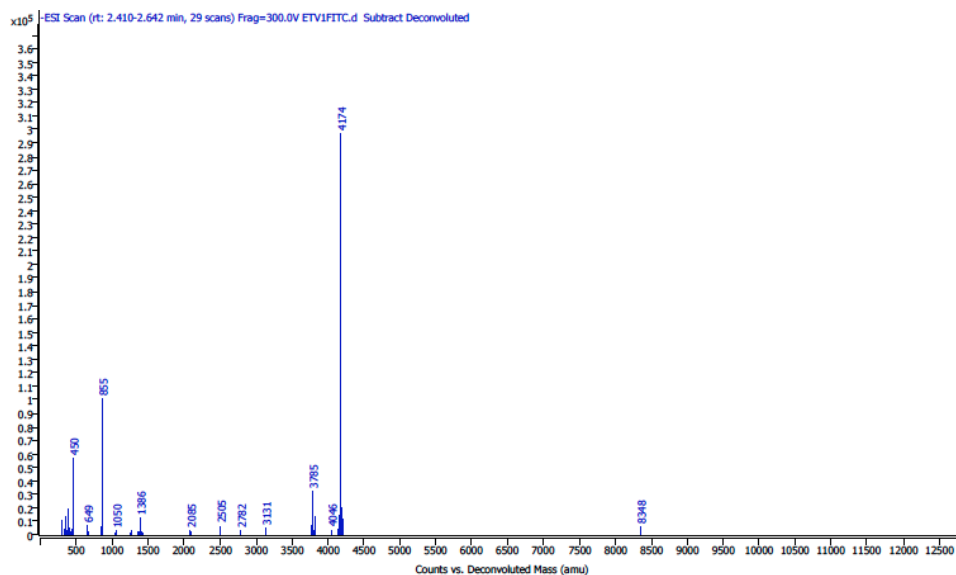
peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the activator peptide sample.



Analytical HPLC chromatogram of **FITC-ETV1 (38-69)** monitored at 280 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

LC-MS:



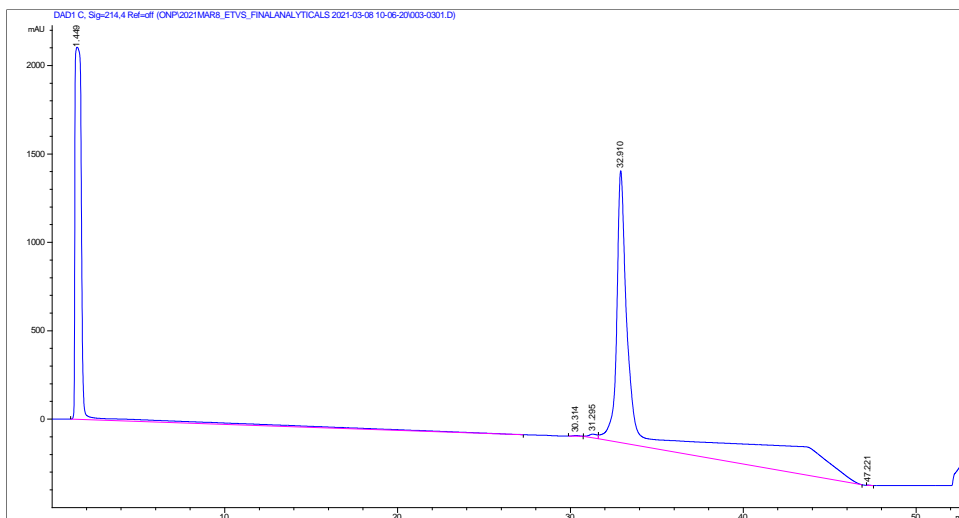


LC-MS of **FITC-ETV1 (38-69)** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

Fluorescein-labeled ETV4 (45-76)

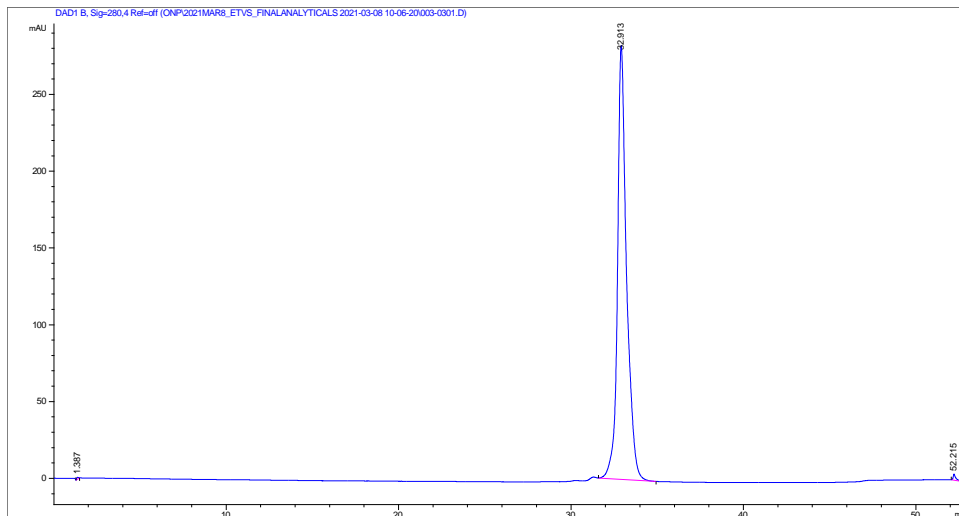
Sequence: N'- FITC-βAla-LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ-C'

HPLC:



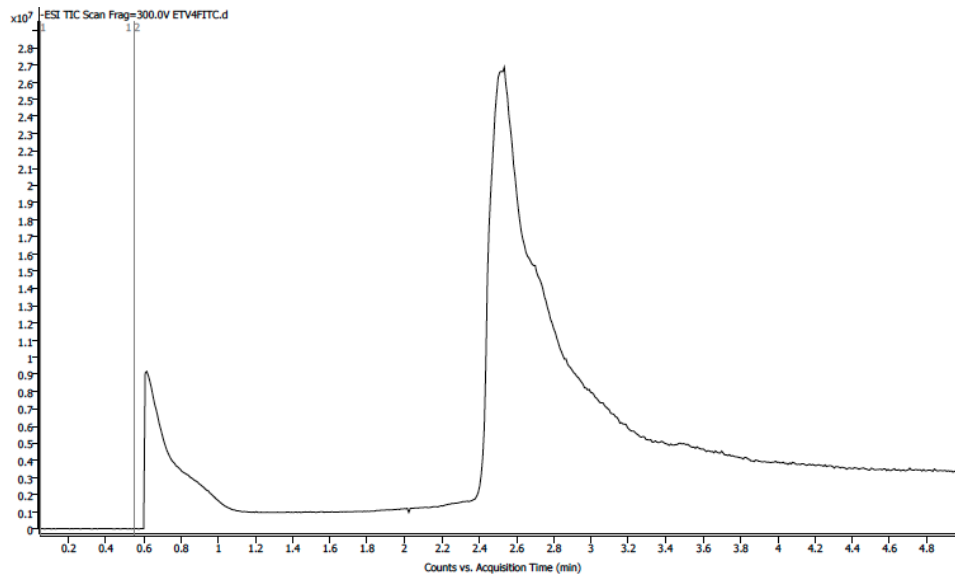
Analytical HPLC chromatogram of **FITC-ETV4 (45-76)** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large

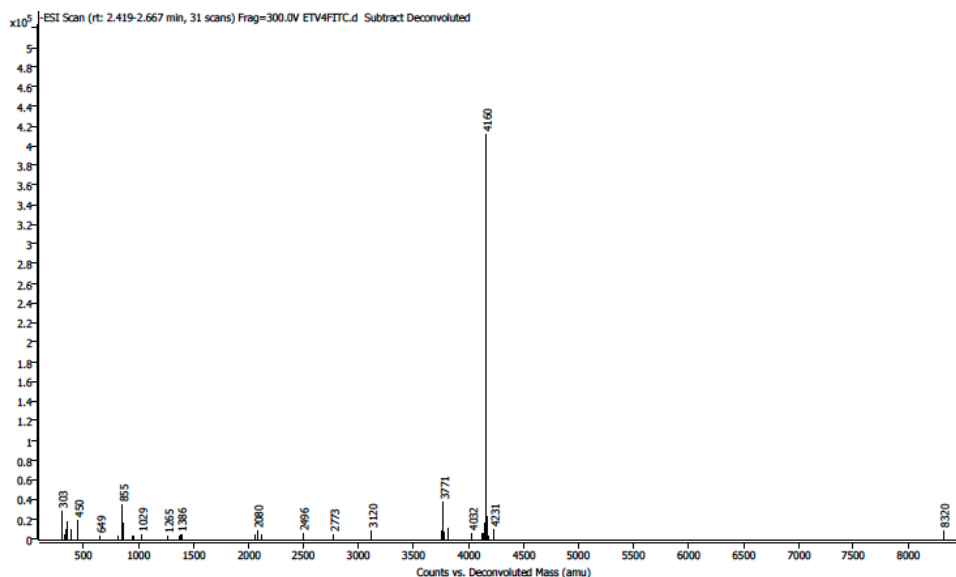
peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the activator peptide sample.



Analytical HPLC chromatogram of **FITC-ETV4 (45-76)** monitored at 280 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.

LC-MS:





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

A.2 Lipopeptide Characterization

All HPLC chromatograms of lipopeptide analogs were monitored at only 214 nm due to a lack of aromatic residues. This resulted in peaks with a lower mAU, especially compared to the solvent peak within the first 5 min of each HPLC chromatogram that is a result of these analogs being dissolved in DMSO.

34913-2

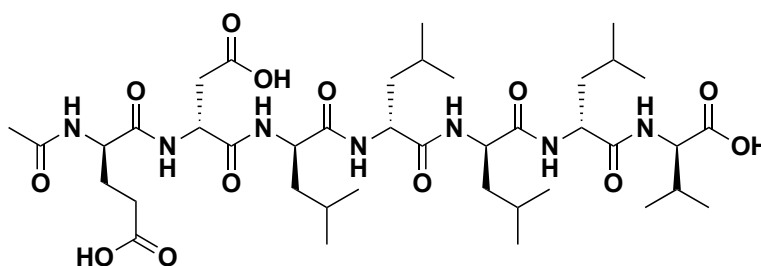
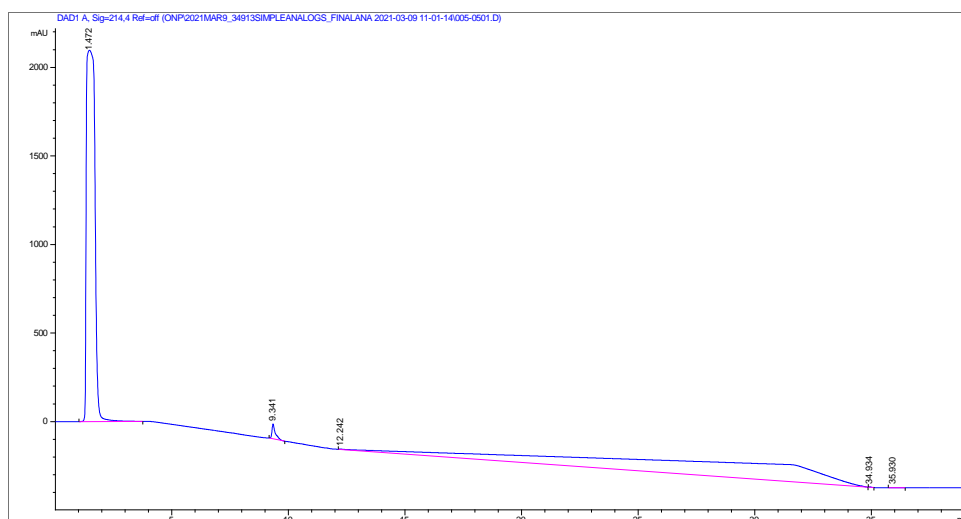


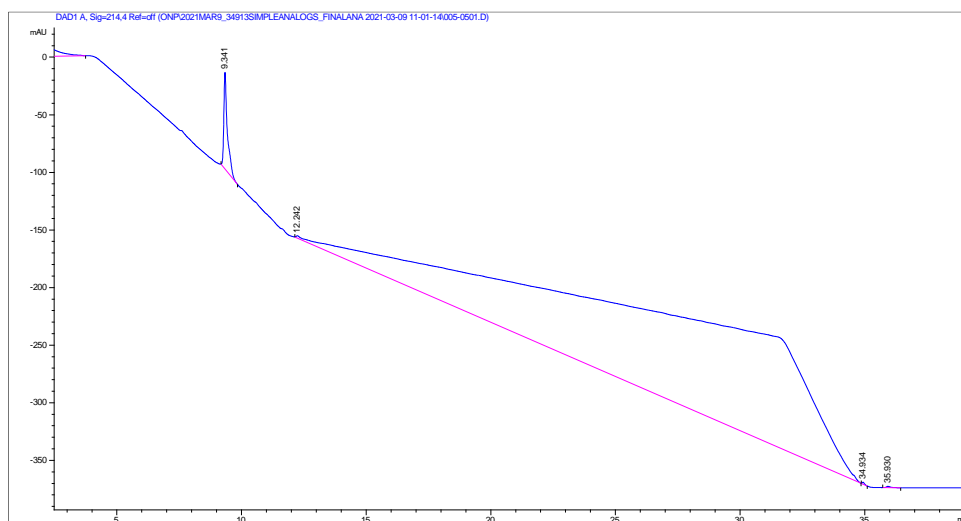
Figure A.1. Structure of analog 34913-2. Structure of simplified lipopeptide analog 34913-2 contains a C-terminal carboxylic acid and a N-terminal acetylation. Peptide is composed of only D-amino acids.

HPLC:

Full Chromatogram:

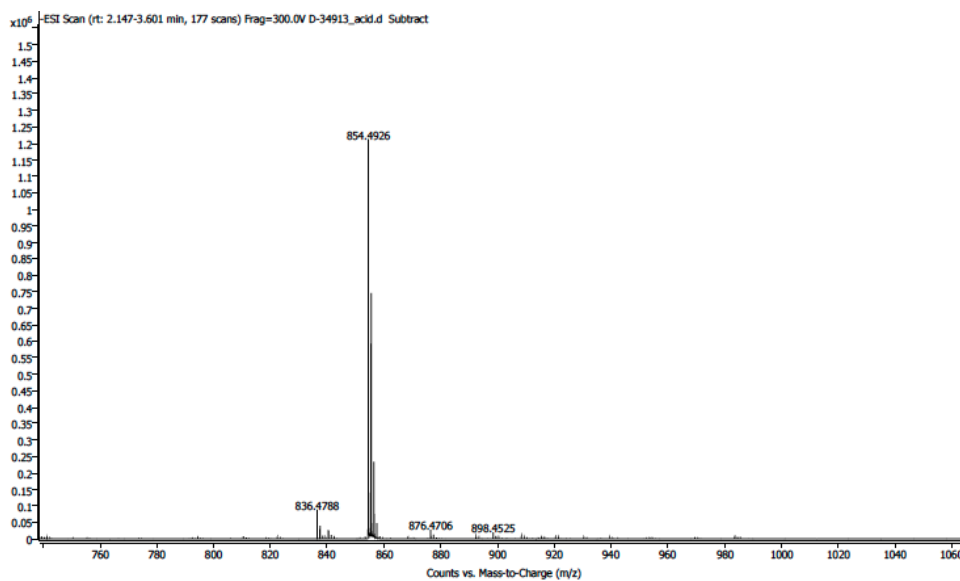
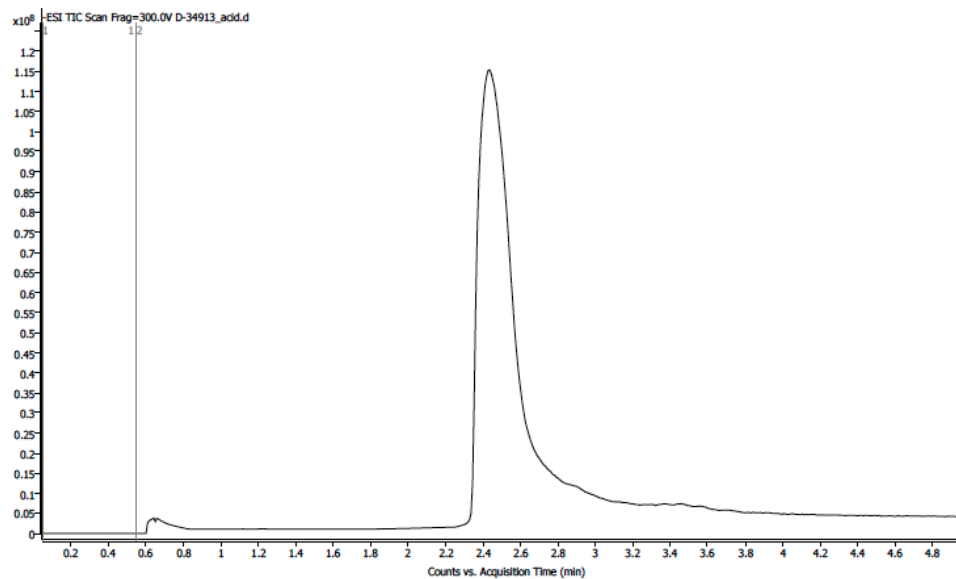


Zoomed in Chromatogram:



Analytical HPLC chromatogram of **34913-2** monitored at 214 nm. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:



LC-MS of **34913-2** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-3

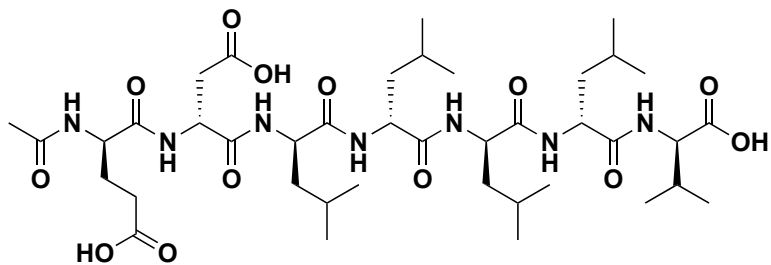
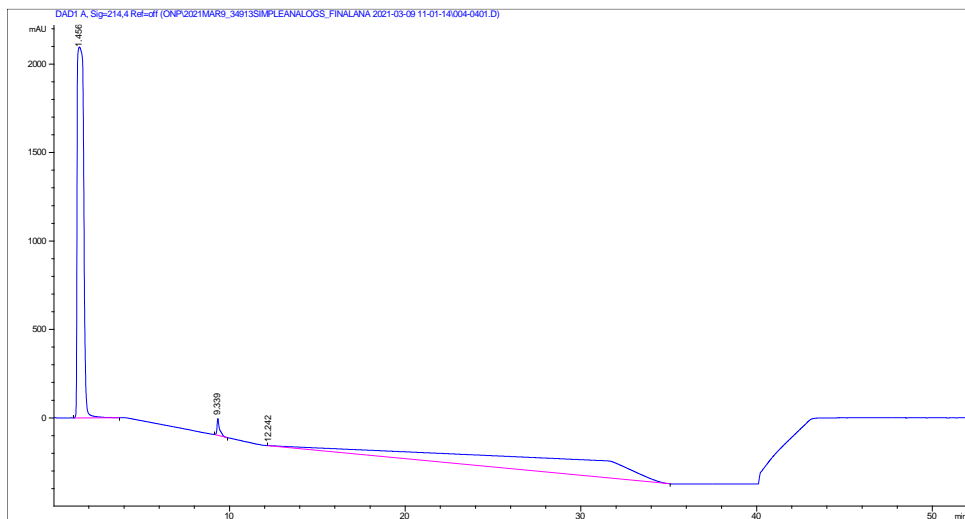


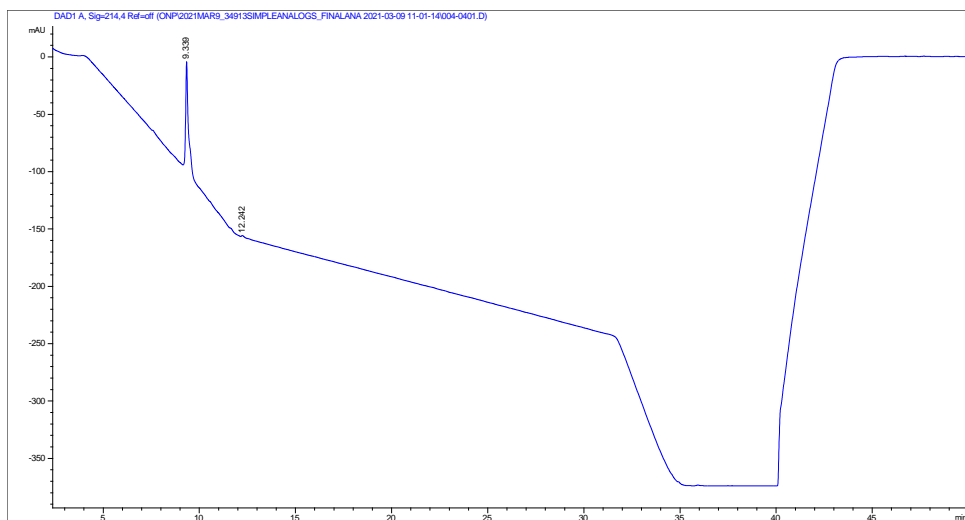
Figure A.2. Structure of analog 34913-3. Structure of simplified lipopeptide analog 34913-3 contains a C-terminal carboxylic acid and a N-terminal acetylation. Peptide is composed of only L-amino acids.

HPLC:

Full Chromatogram:

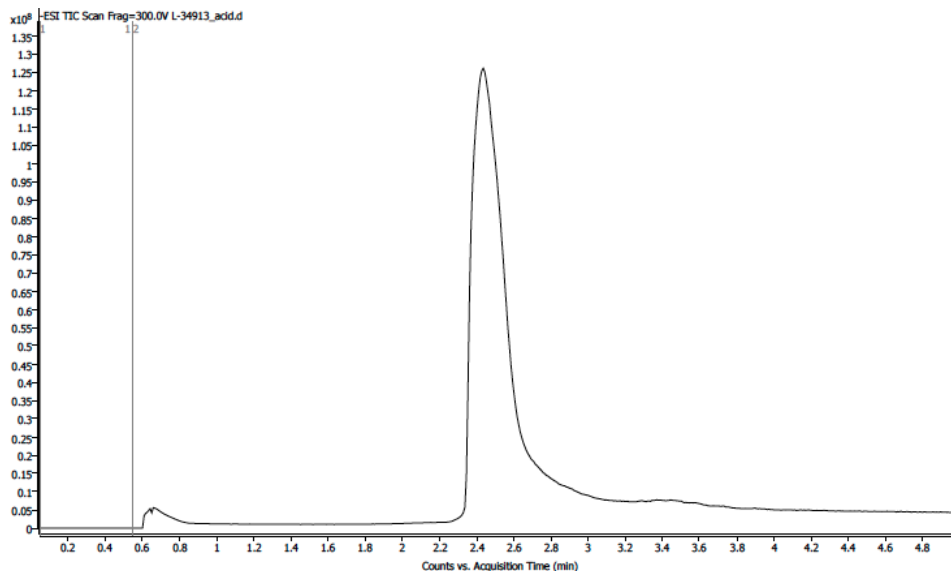


Zoomed in Chromatogram

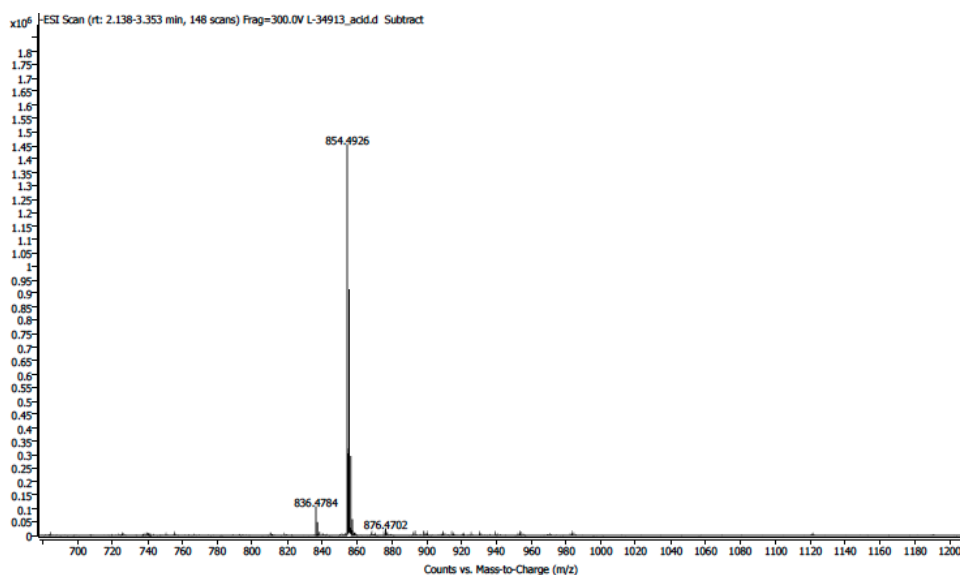


Analytical HPLC chromatogram of **34913-3** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:



LC-MS of **34913-3** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



34913-4

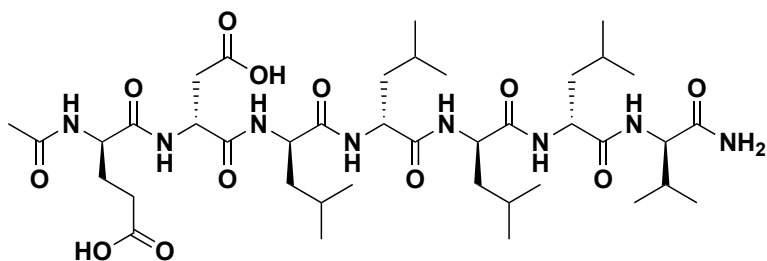
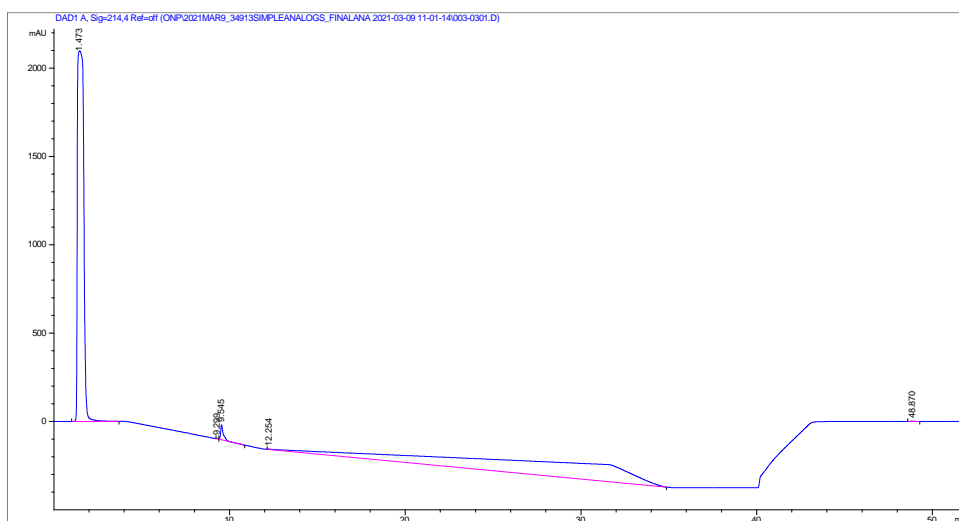


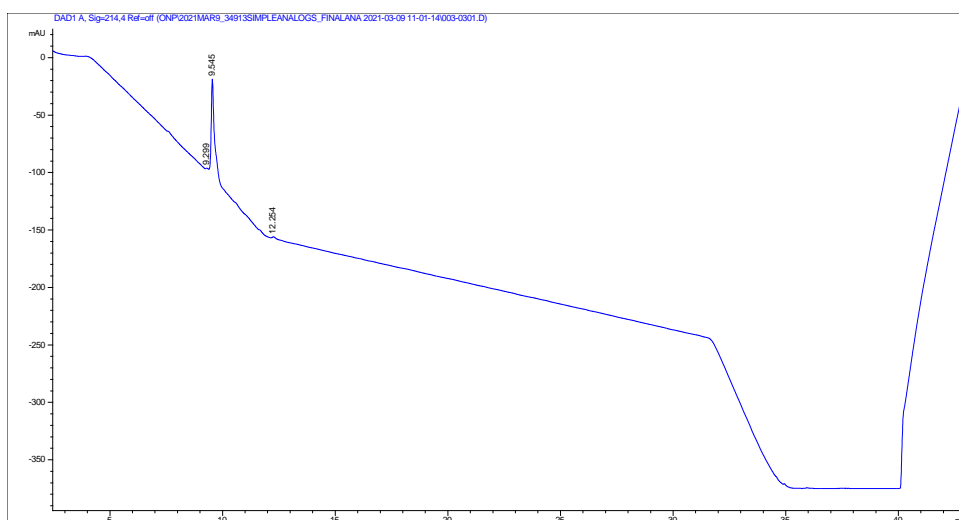
Figure A.3. Structure of analog 34913-4. Structure of simplified lipopeptide analog 34913-4 contains a C-terminal amide and a N-terminal acetylation. Peptide is composed of only D-amino acids.

HPLC:

Full Chromatogram:

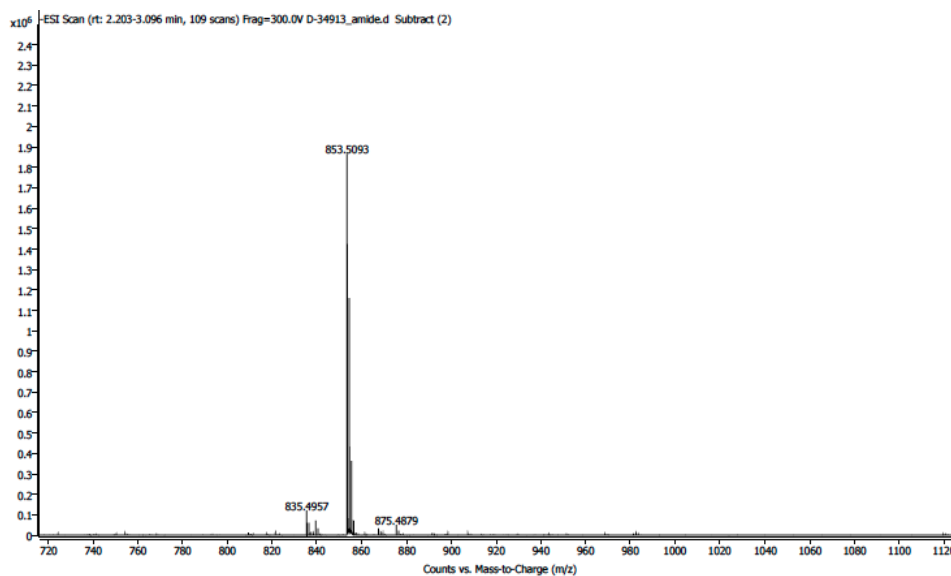
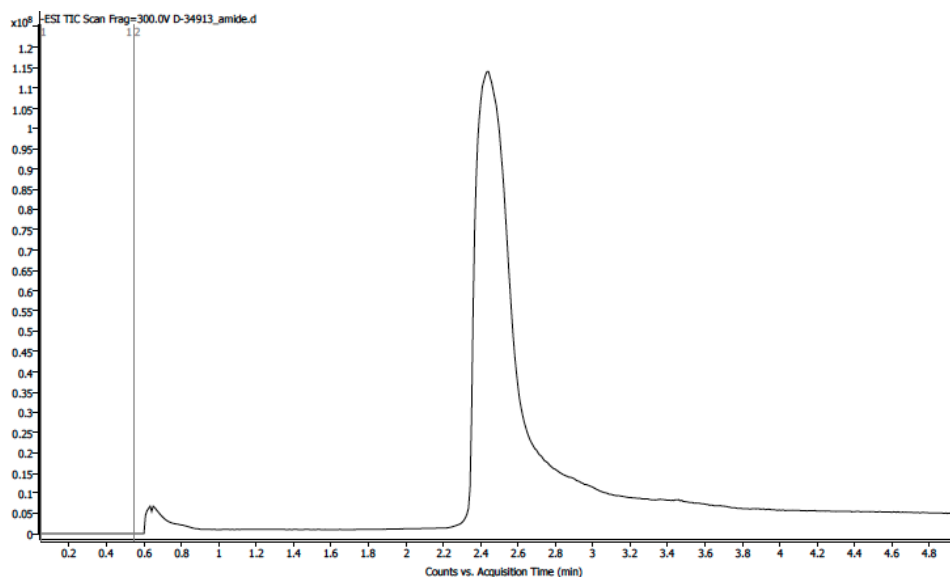


Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-4** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:



LC-MS of **34913-4** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-5

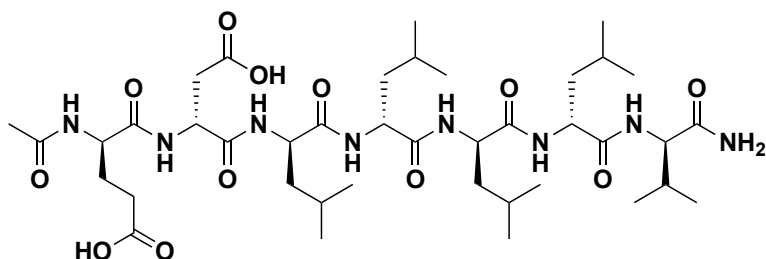
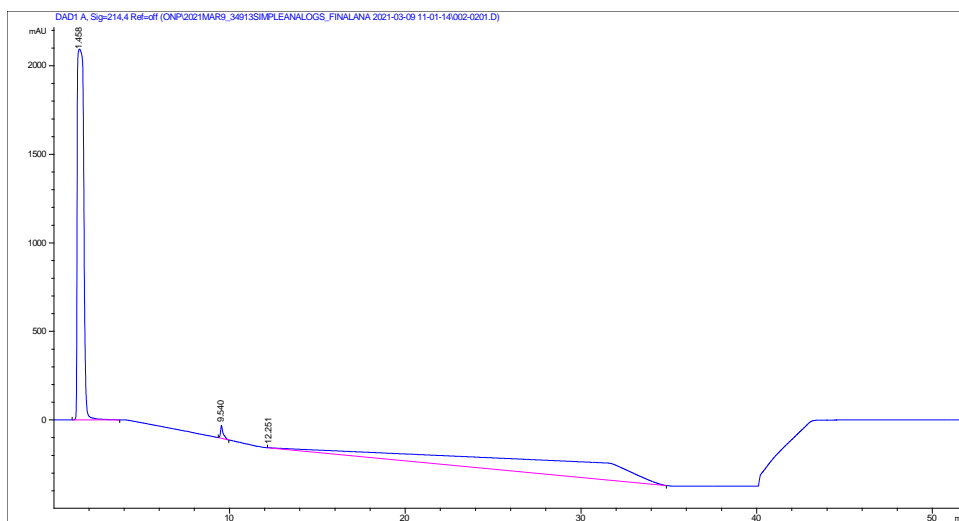


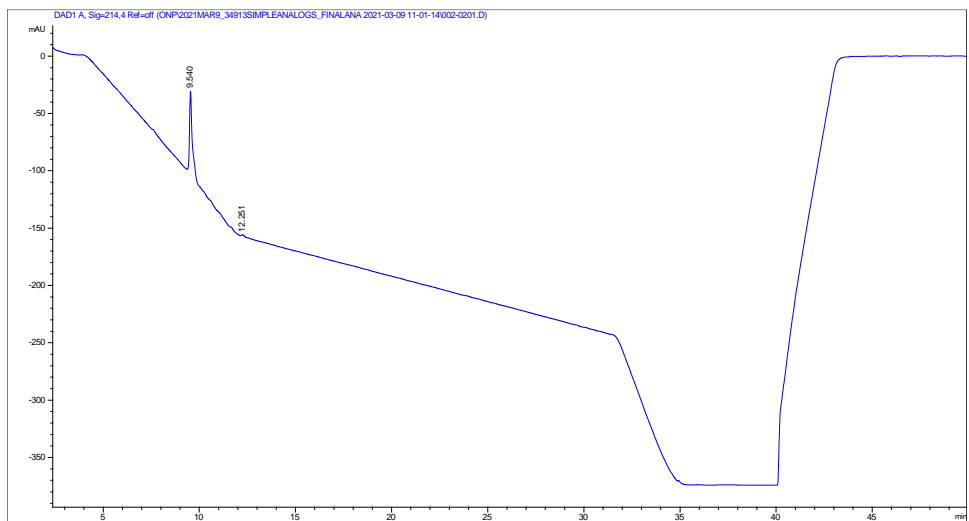
Figure A.4. Structure of analog 34913-5. Structure of simplified lipopeptide analog 34913-5 contains a C-terminal amide and a N-terminal acetylation. Peptide is composed of only L-amino acids.

HPLC:

Full Chromatogram:

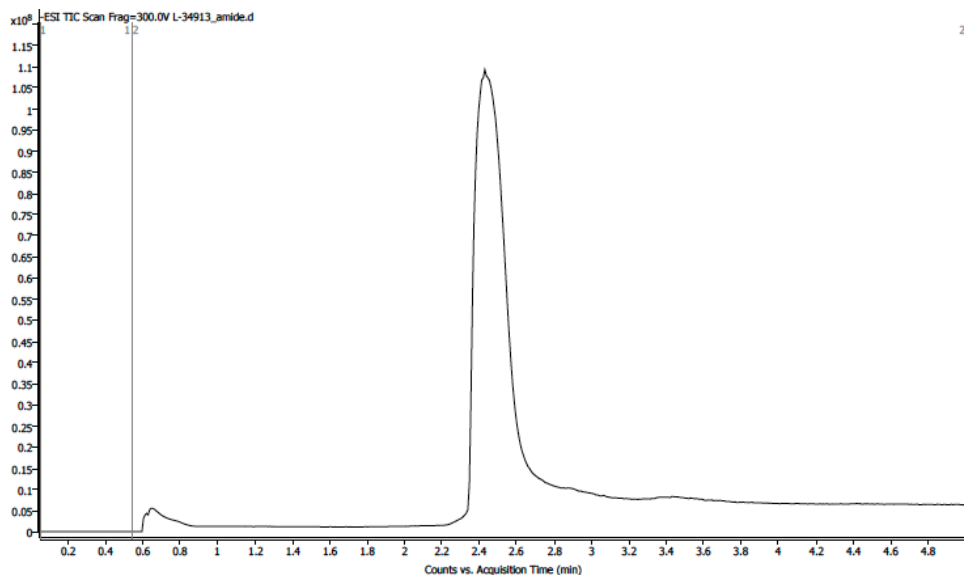


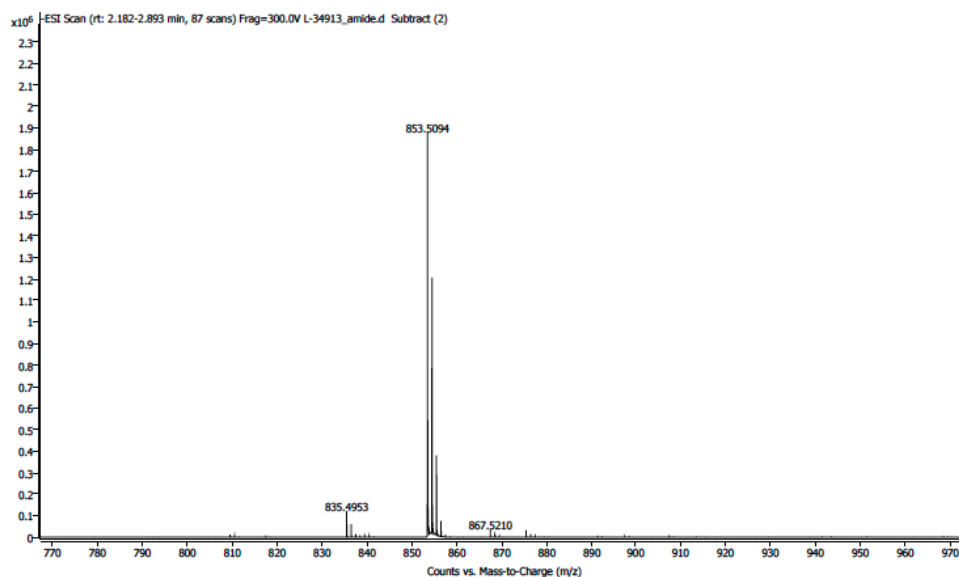
Zoomed in Chromatogram:



Analytical HPLC chromatogram of **34913-5** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS





LC-MS of **34913-5** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-6

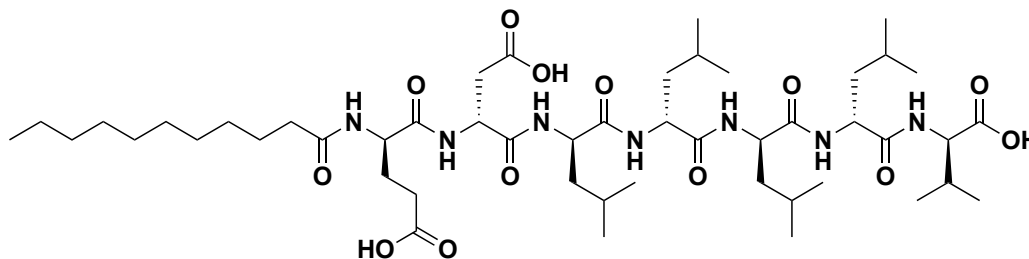
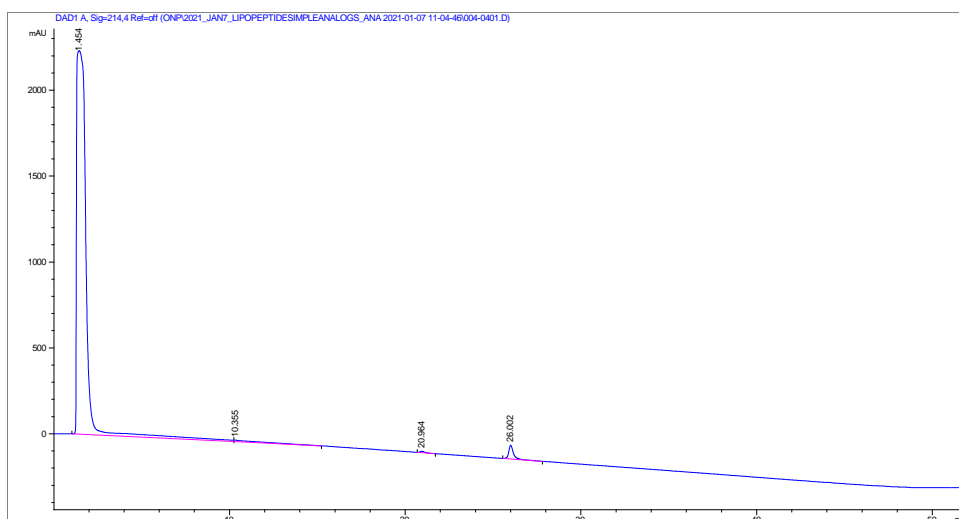


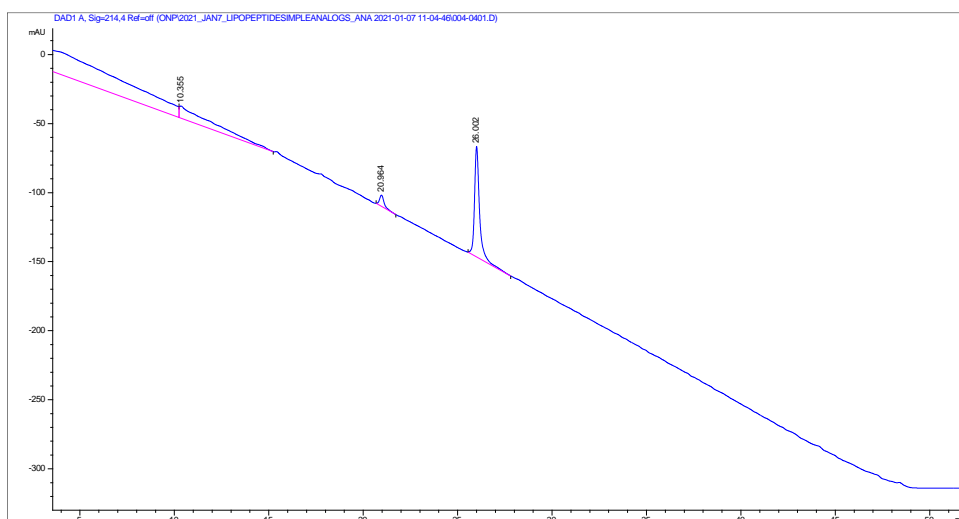
Figure A.5. Structure of analog 34913-6. Structure of simplified lipopeptide analog 34913-6 contains a C-terminal carboxylic acid and a N-terminal attached undecanoic acid. Peptide is composed of only D-amino acids

HPLC:

Full Chromatogram:

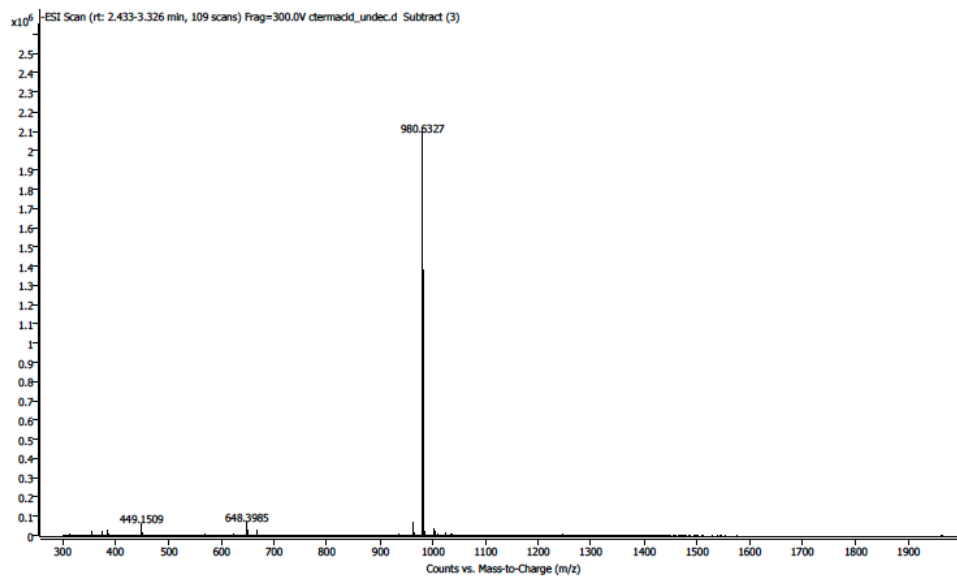
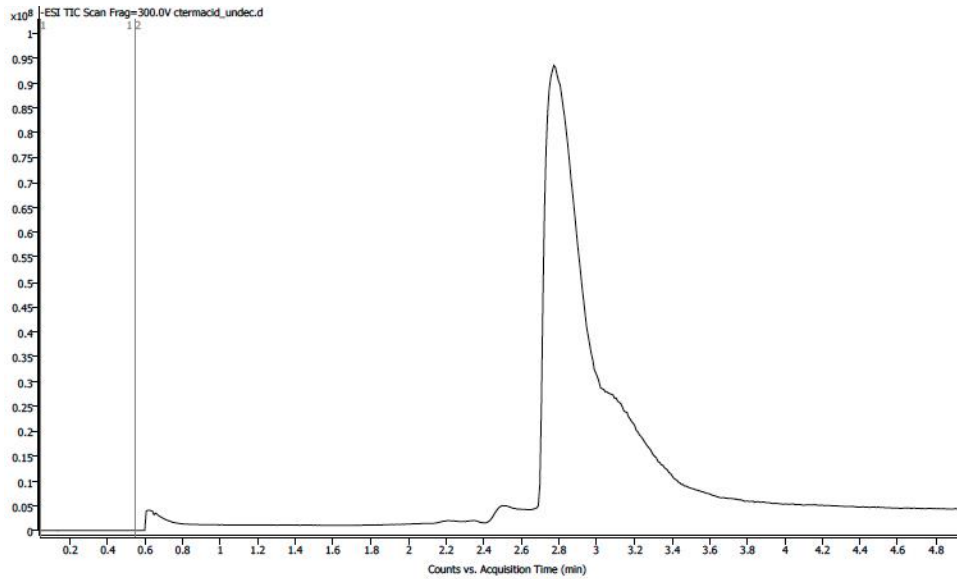


Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-6** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:



LC-MS of **34913-6** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-7

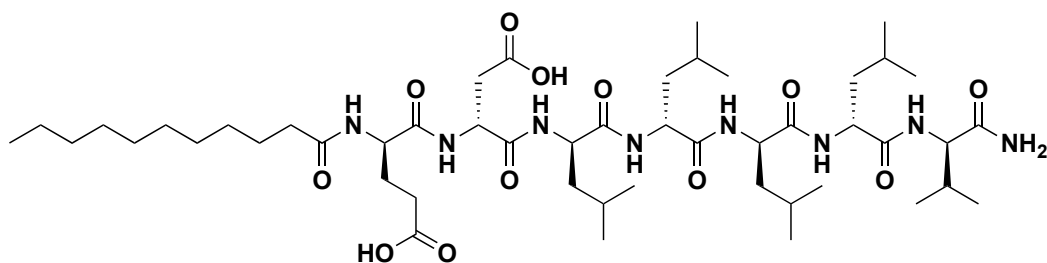
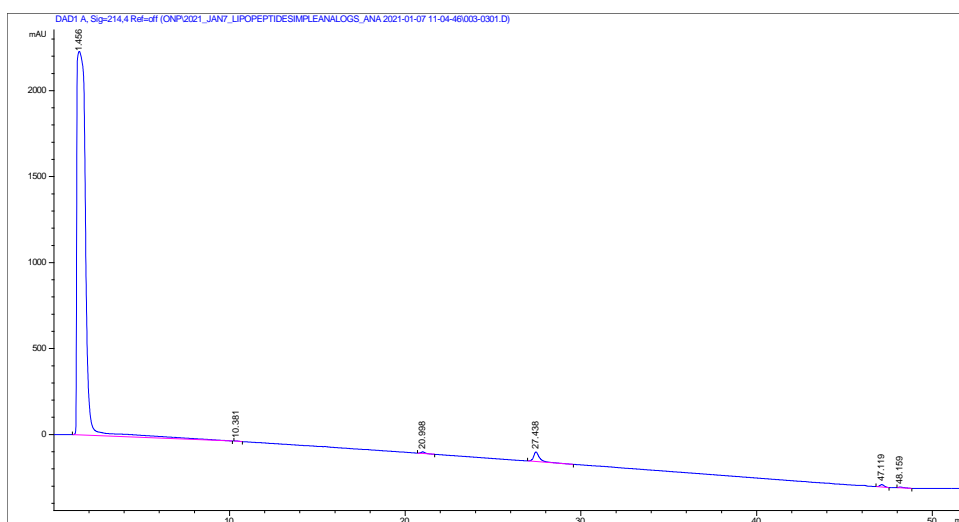


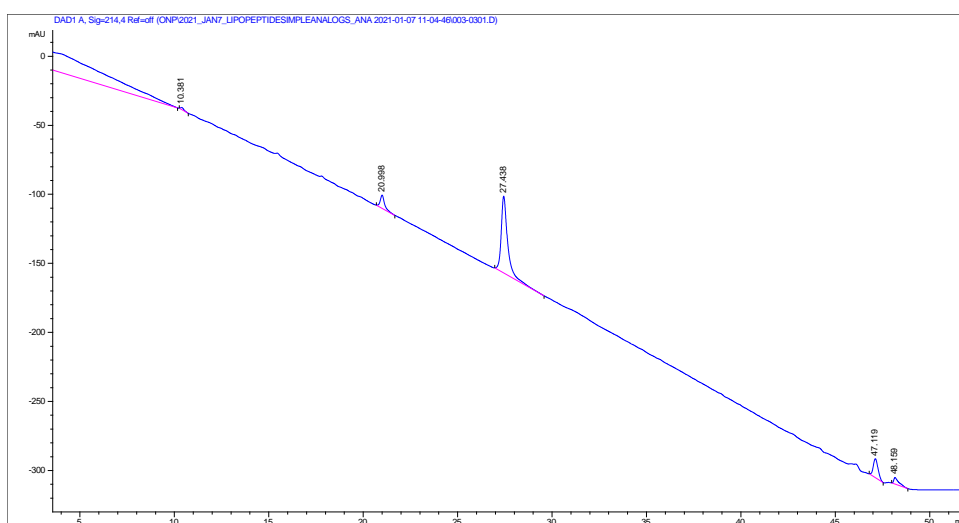
Figure A.6. Structure of analog 34913-7. Structure of simplified lipopeptide analog 34913-7 contains a C-terminal amide and a N-terminal attached undecanoic acid. Peptide is composed of only D-amino acids

HPLC:

Full Chromatogram:

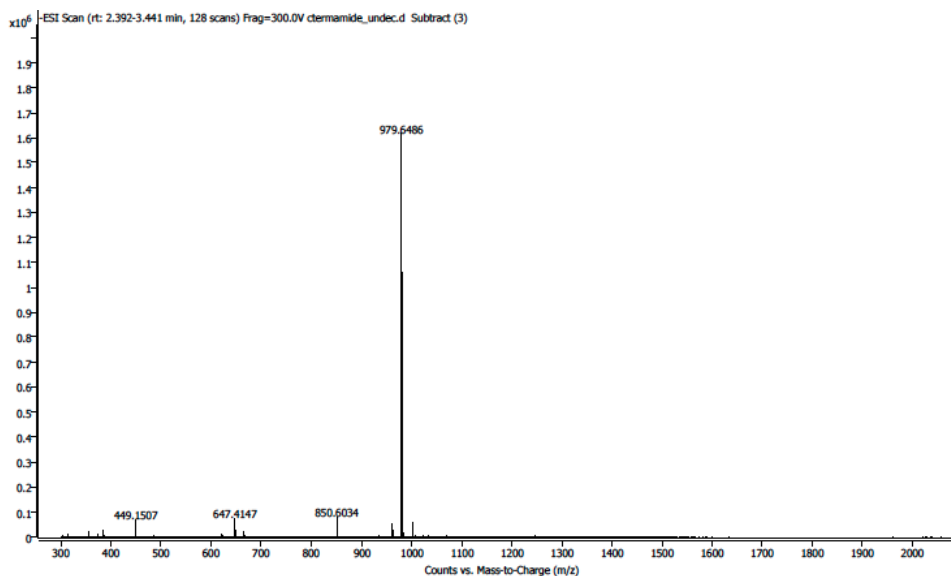
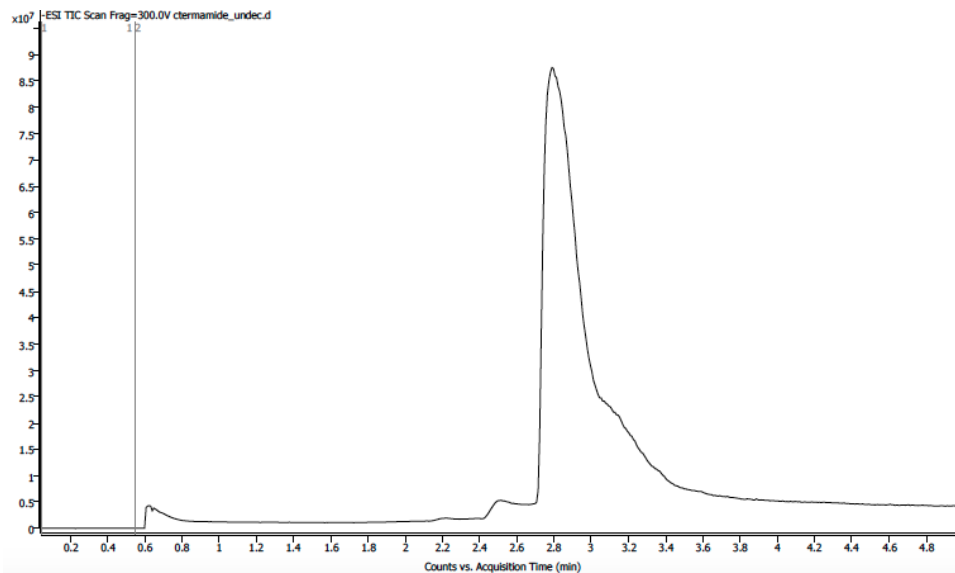


Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-7** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:



LC-MS of **34913-7** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-8

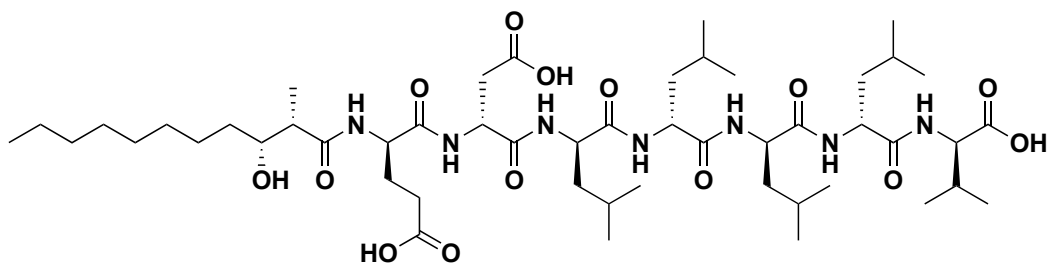
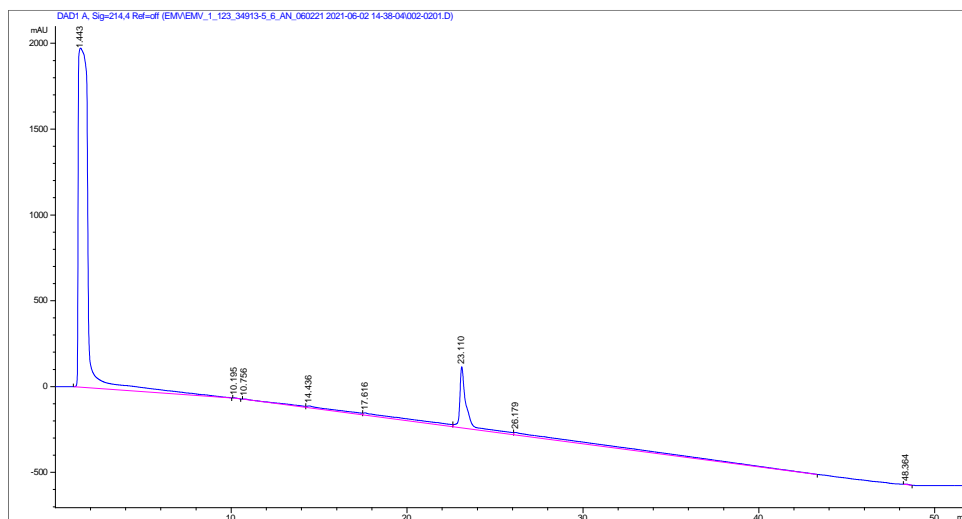


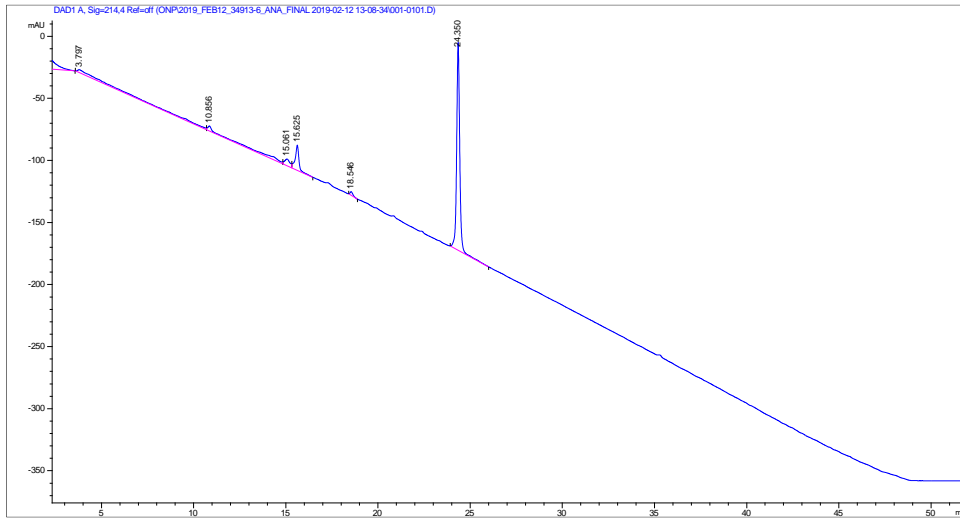
Figure A.7. Structure of analog 34913-8. Structure of simplified lipopeptide analog 34913-8 contains a C-terminal carboxylic acid and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids.

HPLC:

Full Chromatogram:

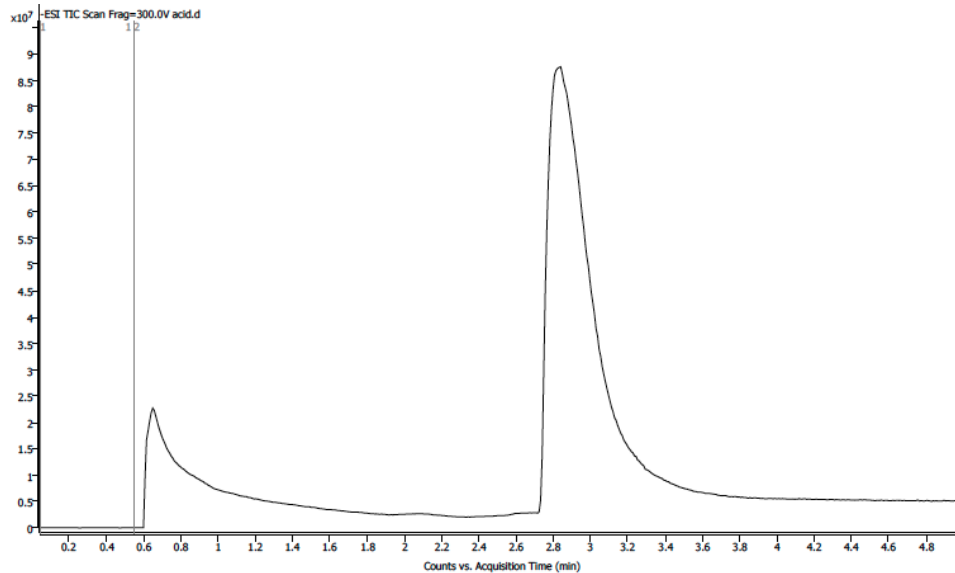


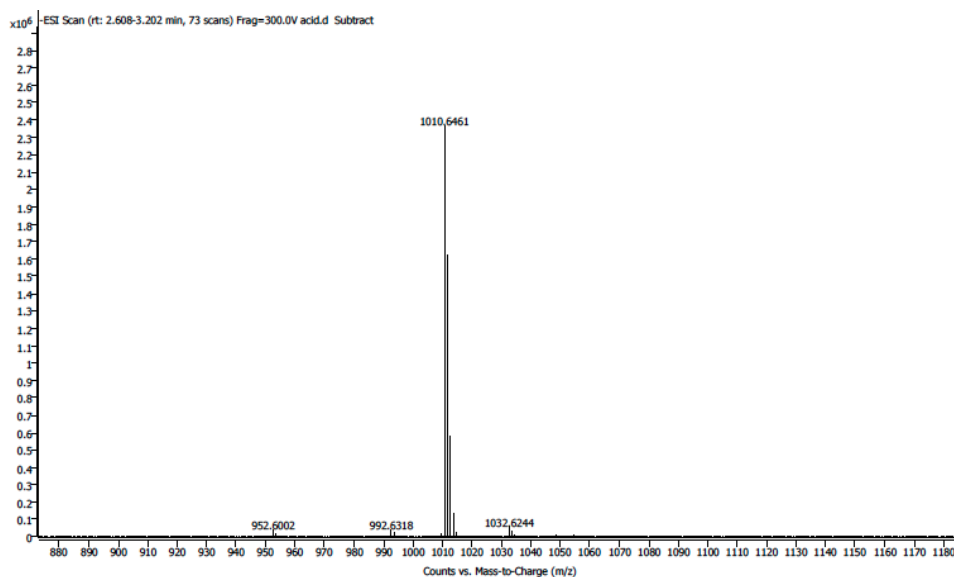
Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-8** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:





LC-MS of **34913-8** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-9

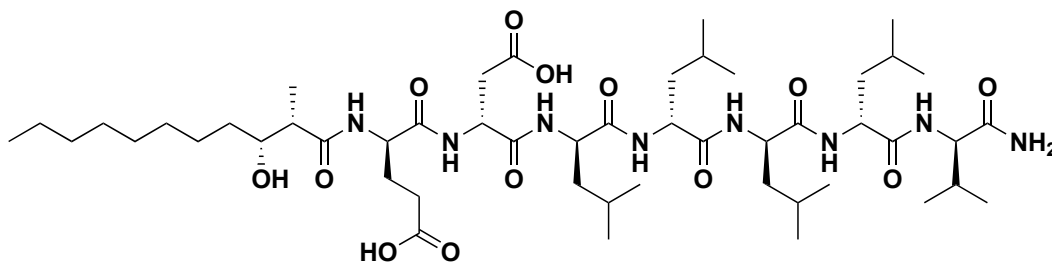
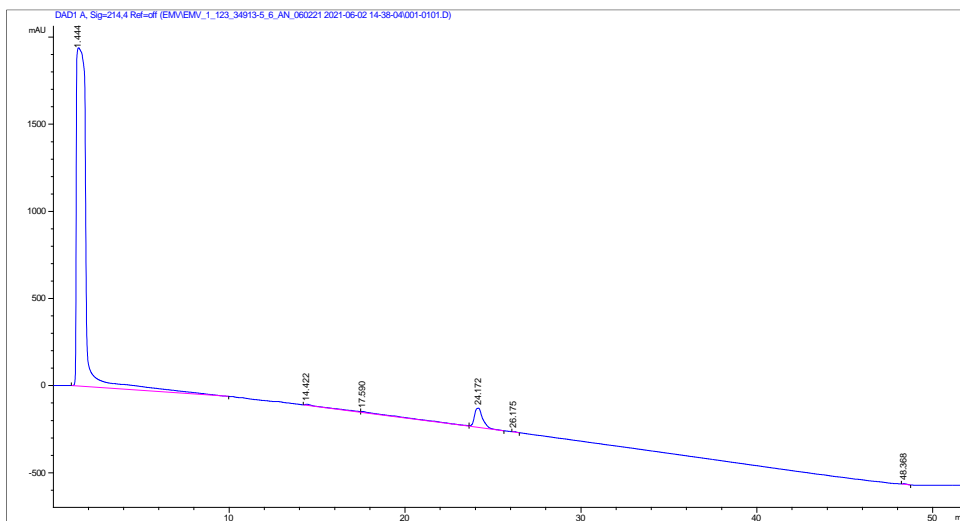


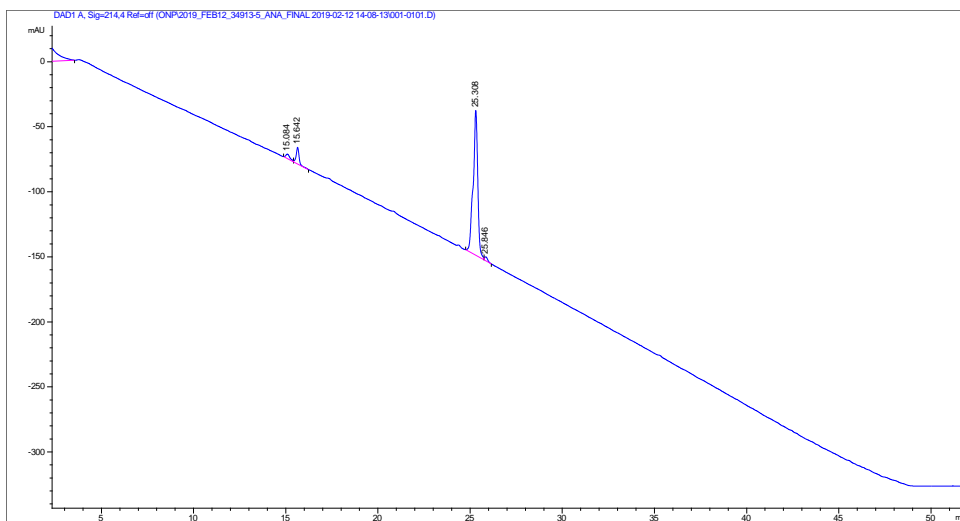
Figure A.8. Structure of analog 34913-9. Structure of lipopeptide analog 34913-9 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2methylundecanoic acid. Peptide is composed of only D-amino acids

HPLC

Full Chromatogram:

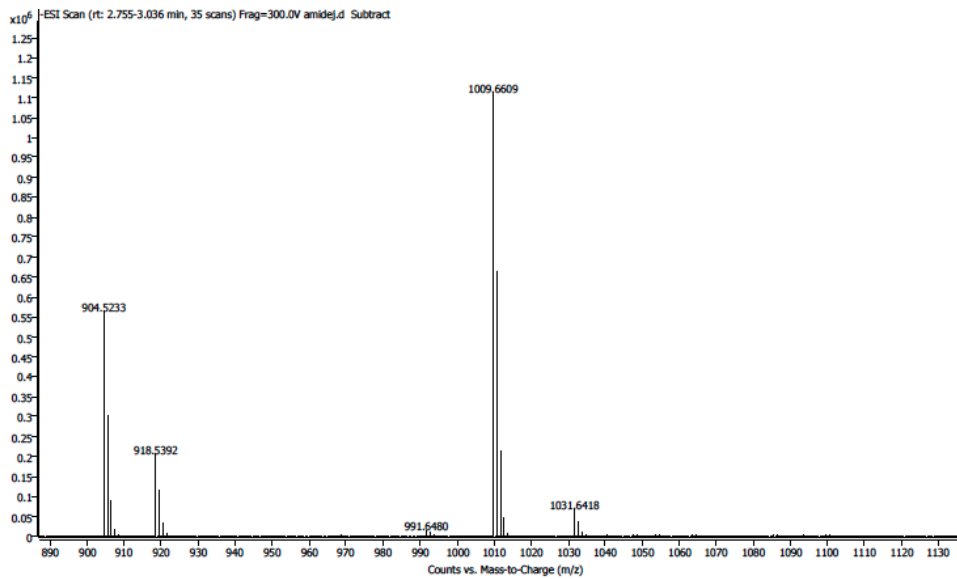
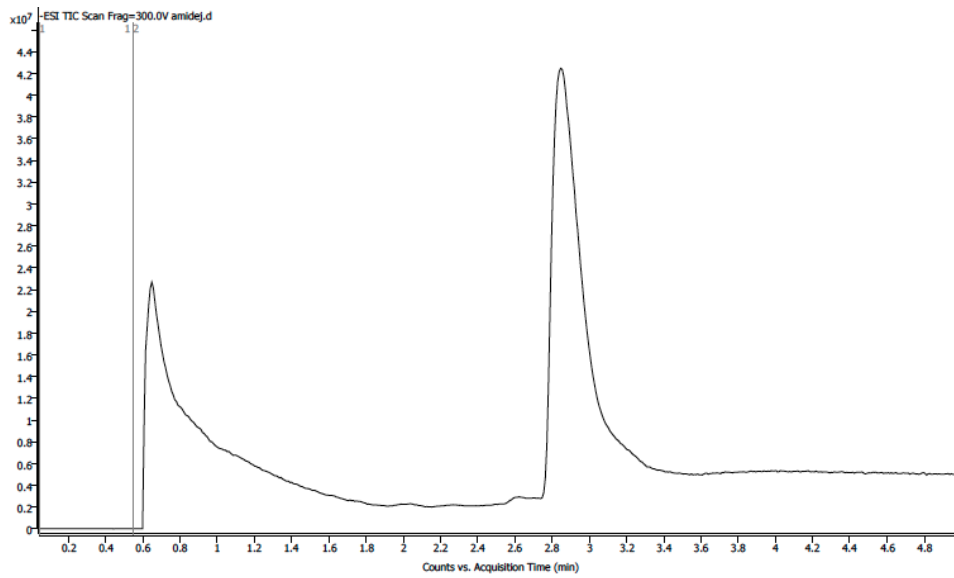


Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-9** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample.

LC-MS:



LC-MS of **34913-9** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala1

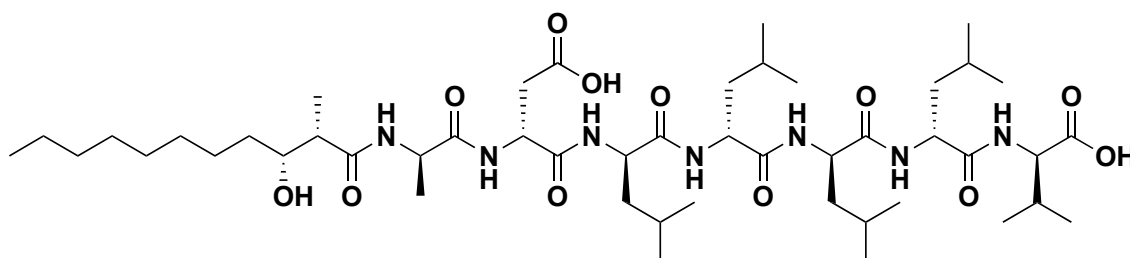
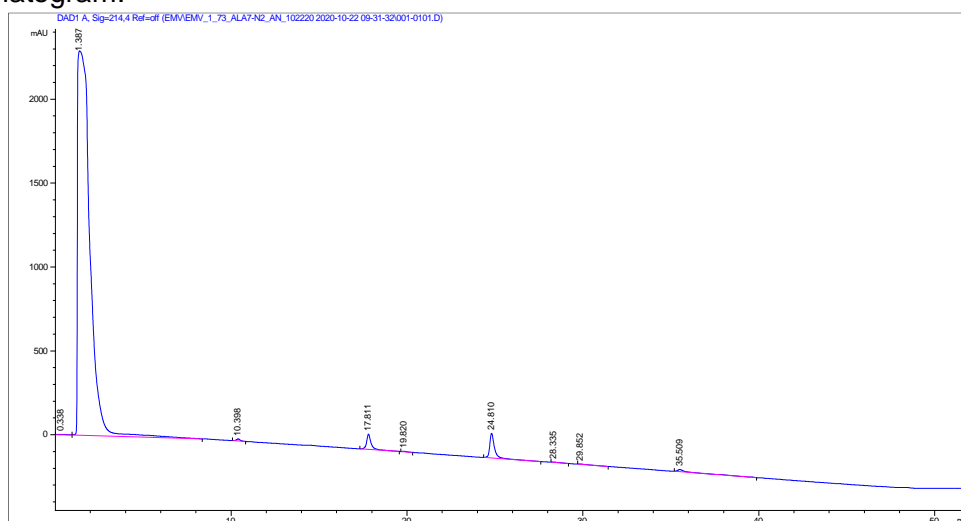


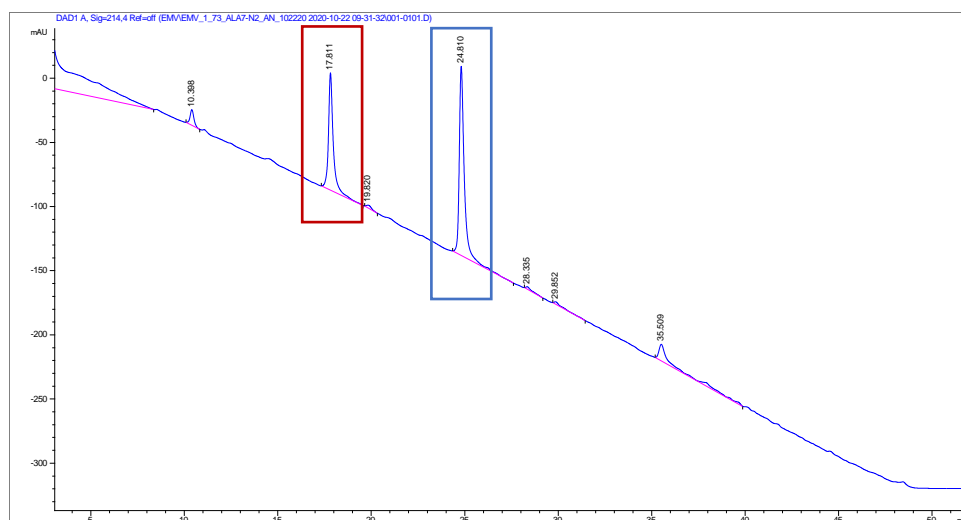
Figure A.9. Structure of analog 34913-Ala1. Structure of lipopeptide analog 34913-Ala contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Glu at residue position 1 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala1** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	0.338	BV	96.32484	2.06589	0.5740	0.013	0.963
2	1.387	VB	1.11622e5	2291.48120	0.6103	0.963	8.366
3	10.398	BB	175.57010	12.21793	0.2156	10.086	10.820
4	17.811	BB	1625.39807	91.42867	0.2554	17.326	19.560
5	19.820	BB	40.47804	2.11279	0.2935	19.646	20.326
6	24.810	BB	2744.55981	147.55103	0.2710	24.353	27.600
7	28.335	BB	40.97061	1.85308	0.3059	28.193	29.180
8	29.852	BB	46.81225	1.78641	0.3597	29.706	31.440
9	35.509	BB	503.13950	12.78136	0.5313	35.193	39.866

Table A.1. Peak integrations of 34913-Ala1. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

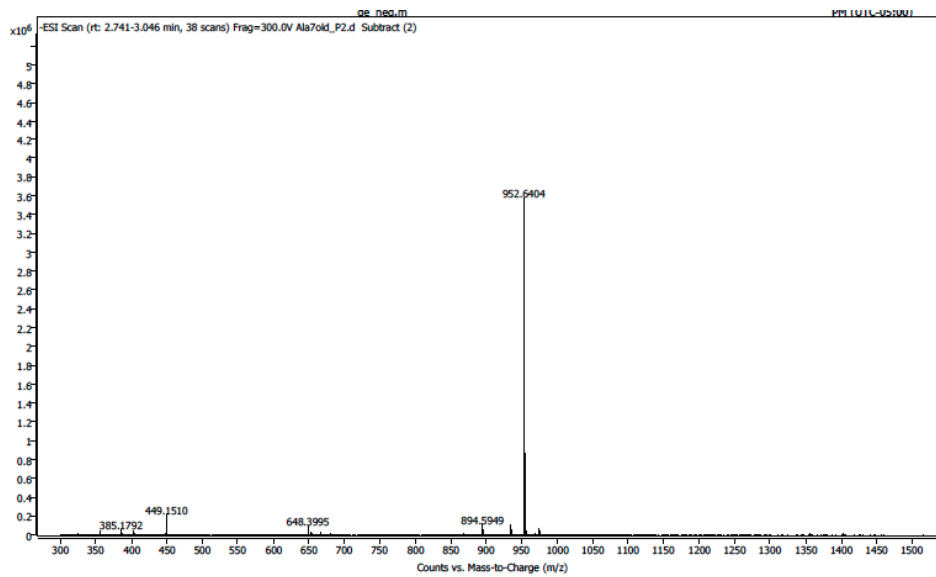
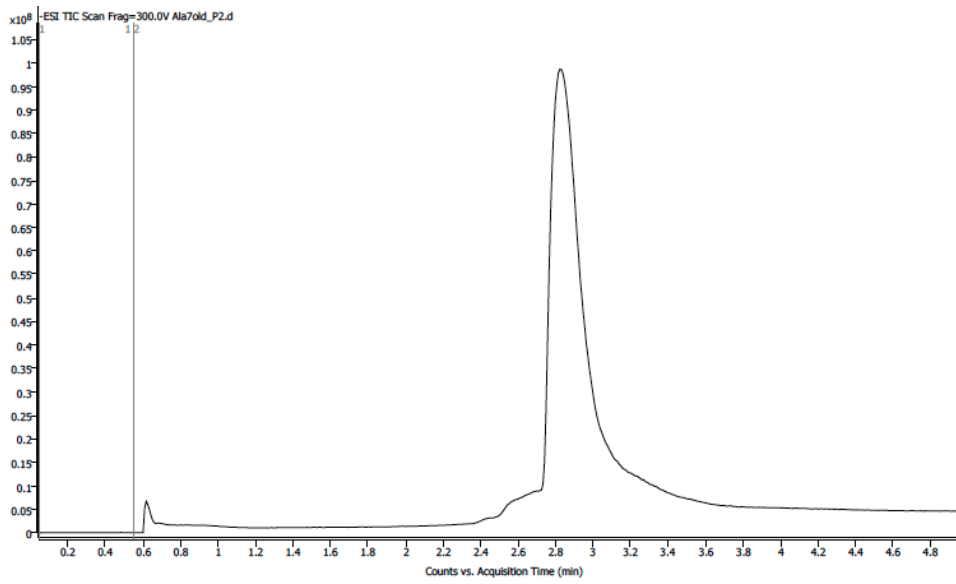
Total Area of 34913-Ala1 + impurity: 4369.96

Area of 34913-Ala1: 2744.56

Approximate % of 34913-Ala1: 62.8%

Approximate % yield of 34913-Ala1 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala1** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala2

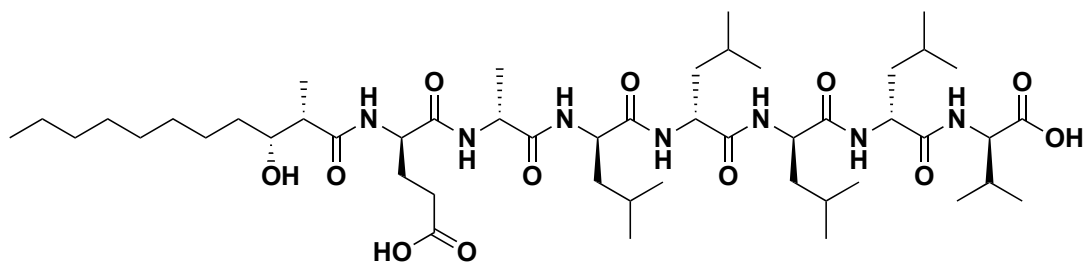
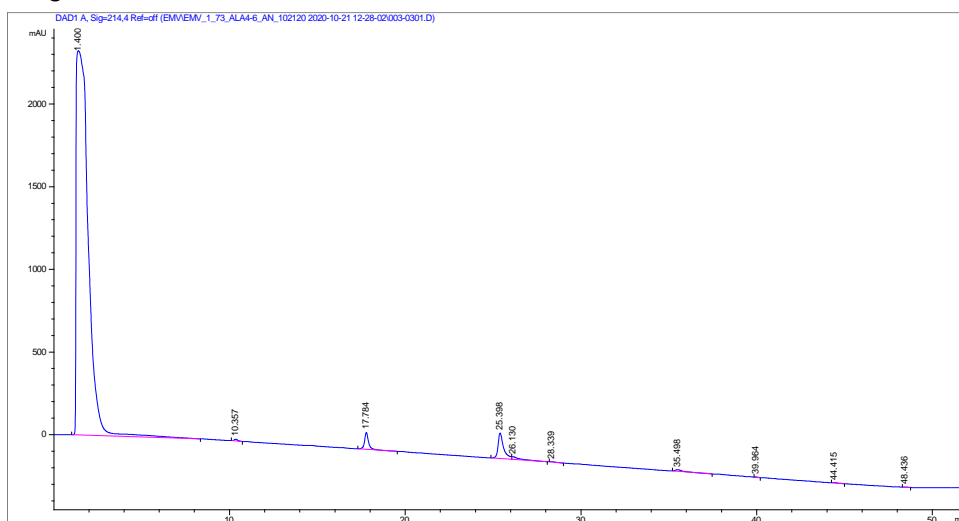


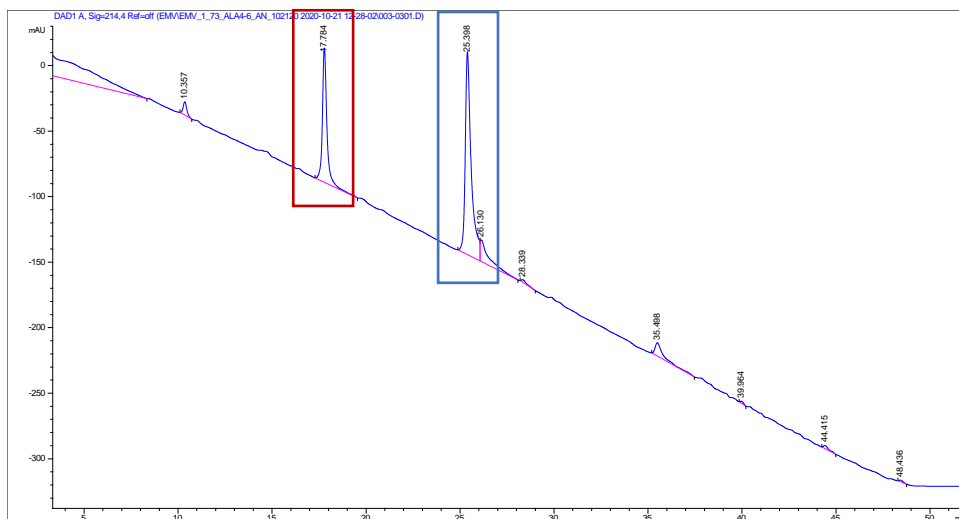
Figure A.10. Structure of analog 34913-Ala2. Structure of simplified lipopeptide analog 34913-Ala2 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Asp at residue position 2 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala2** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

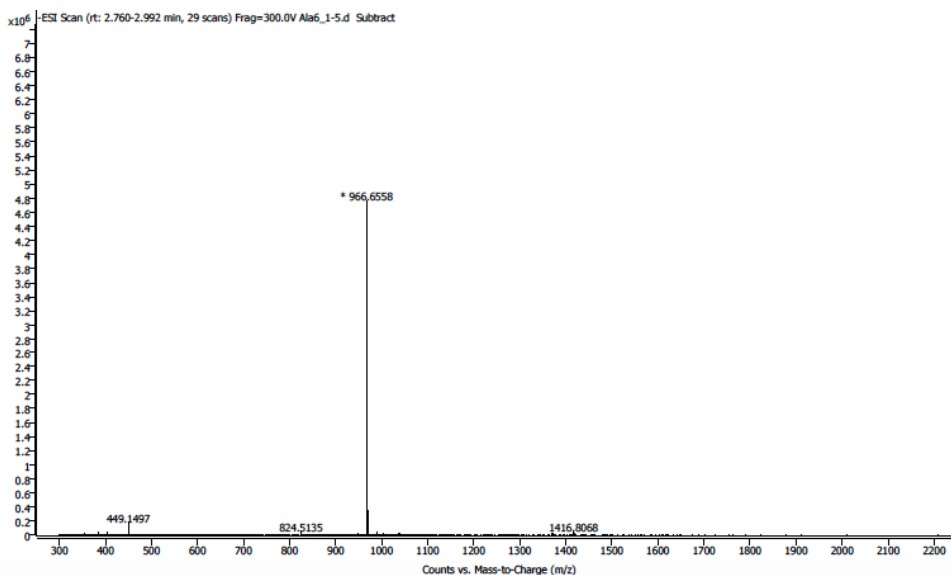
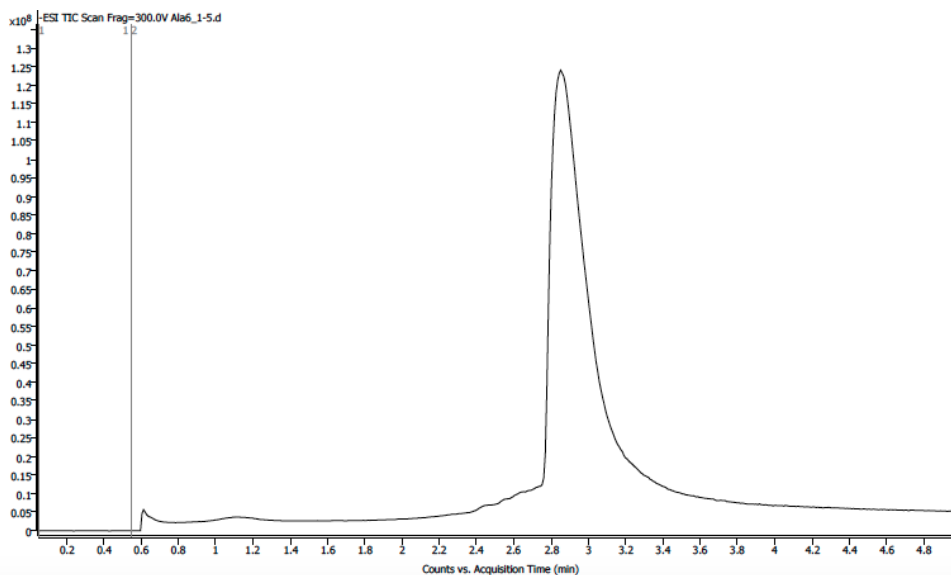
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.400	BB	1.10495e5	2324.28369	0.6046	1.009	8.349
2	10.357	BB	137.32452	10.20084	0.2071	10.109	10.729
3	17.784	BB	1675.31799	102.46528	0.2388	17.296	19.542
4	25.398	BV	3399.16675	154.73364	0.3143	24.882	26.064
5	26.130	VB	498.81854	16.44431	0.3999	26.064	28.076
6	28.339	BB	37.50248	1.99850	0.2689	28.196	29.003
7	35.498	BB	294.29474	10.19863	0.4107	35.202	37.463
8	39.964	BB	19.41242	1.52039	0.2050	39.849	40.209
9	44.415	BB	46.97075	2.42891	0.2836	44.249	44.989
10	48.436	BB	17.37900	1.29852	0.2061	48.289	48.749

Table A.2. Peak integrations of 34913-Ala2. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Ala2 + impurity: 5074.49
Area of 34913-Ala2: 3399.17
Approximate % of 34913-Ala2: 67.0%

Approximate % yield of 34913-Ala2 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala2** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala3

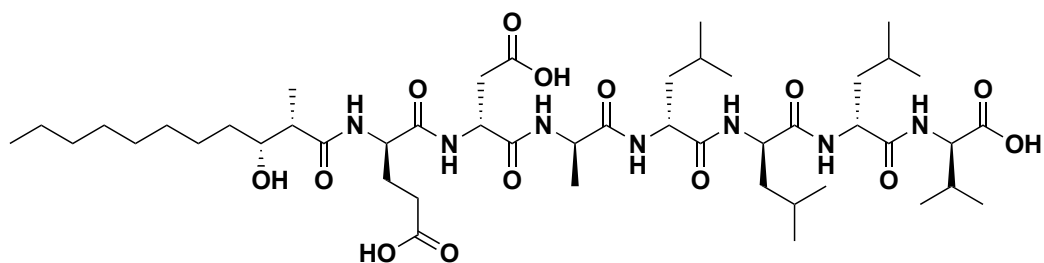
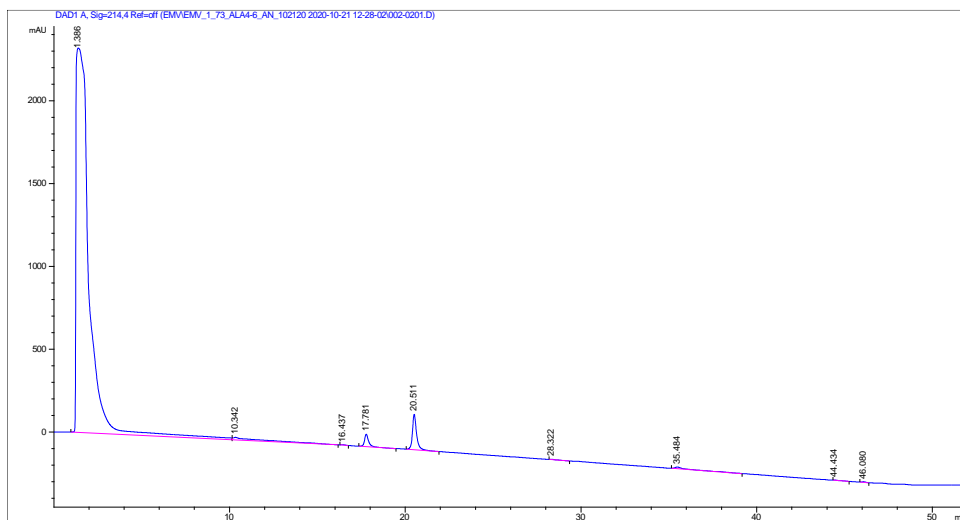


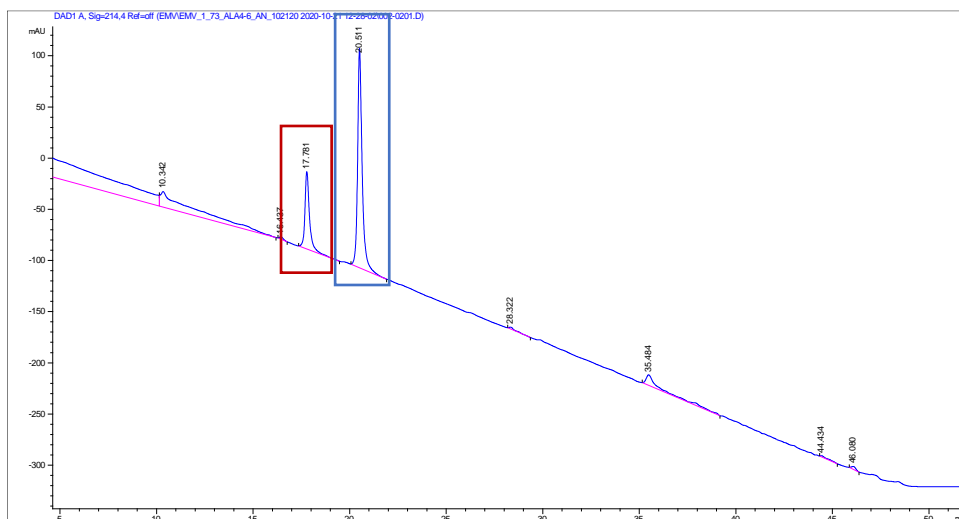
Figure A.11. Structure of analog 34913-Ala3. Structure of simplified lipopeptide analog 34913-Ala3 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Leu at residue position 3 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala3** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((*2S,3R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

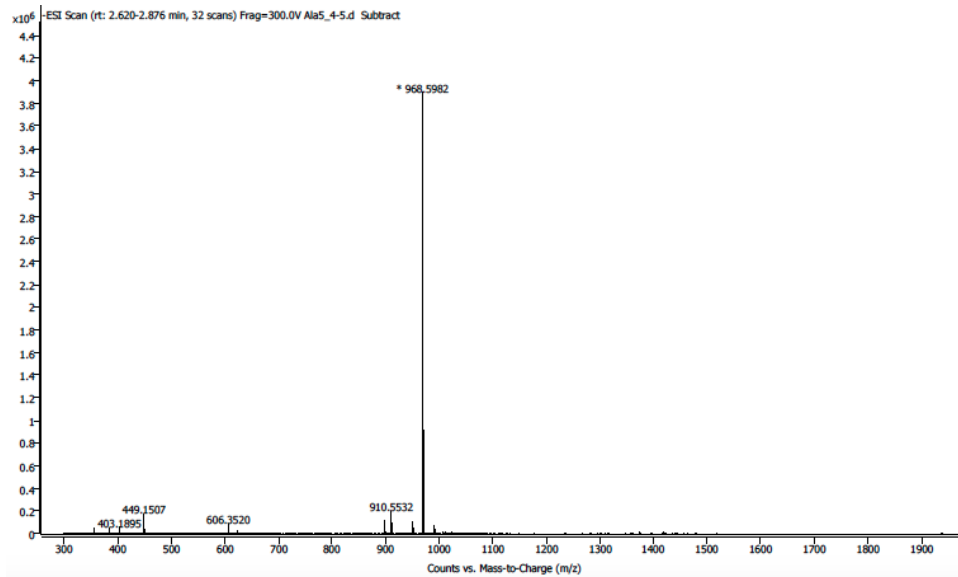
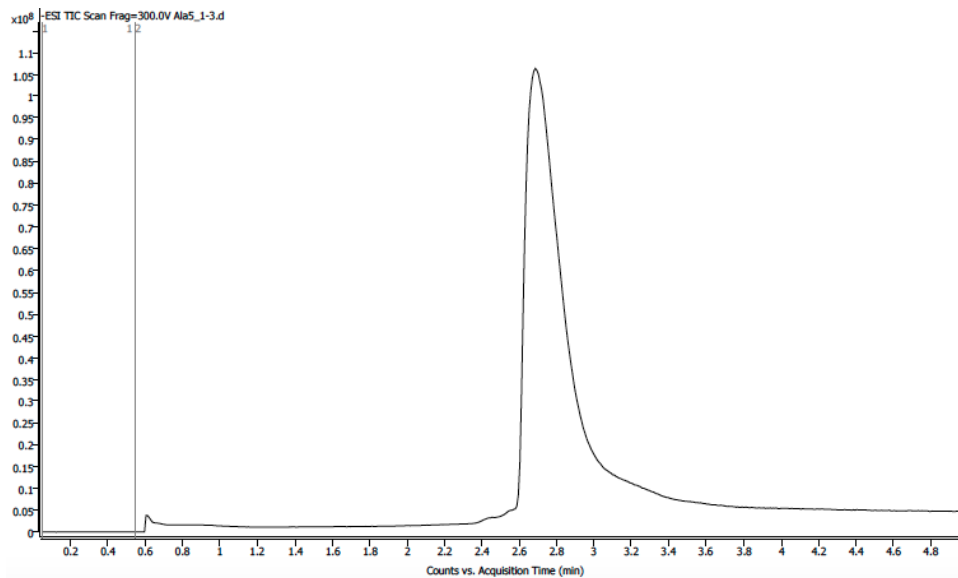
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.386	BV	1.18106e5	2321.31470	0.6356	0.977	10.144
2	10.342	VB	1802.29138	15.13460	1.4373	10.144	16.184
3	16.437	BB	32.36542	2.38376	0.2104	16.284	16.770
4	17.781	BB	1247.00354	75.59955	0.2404	17.357	19.477
5	20.511	BB	3509.67603	214.32384	0.2410	20.064	21.917
6	28.322	BB	34.12189	1.63701	0.2932	28.184	29.357
7	35.484	BB	462.57162	10.40453	0.5907	35.150	39.177
8	44.434	BB	35.76845	1.16584	0.4059	44.338	45.257
9	46.080	BB	41.69216	2.71548	0.2391	45.870	46.384

Table A.3. Peak integrations of 34913-Ala3. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Ala3 + impurity: 4756.68
Area of 34913-Ala3: 3509.68
Approximate % of 34913-Ala3: 73.8%

Approximate % yield of 34913-Ala3 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala3** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala4

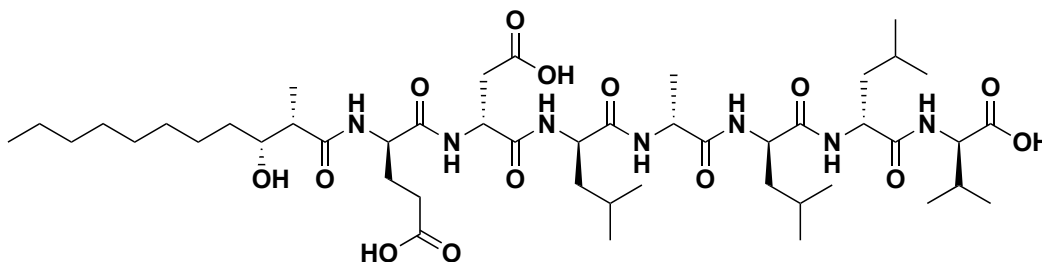
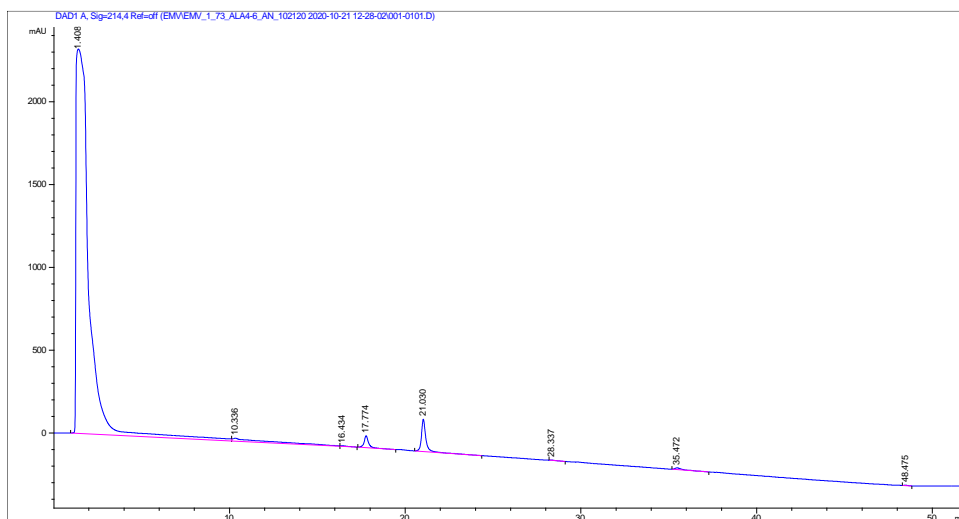


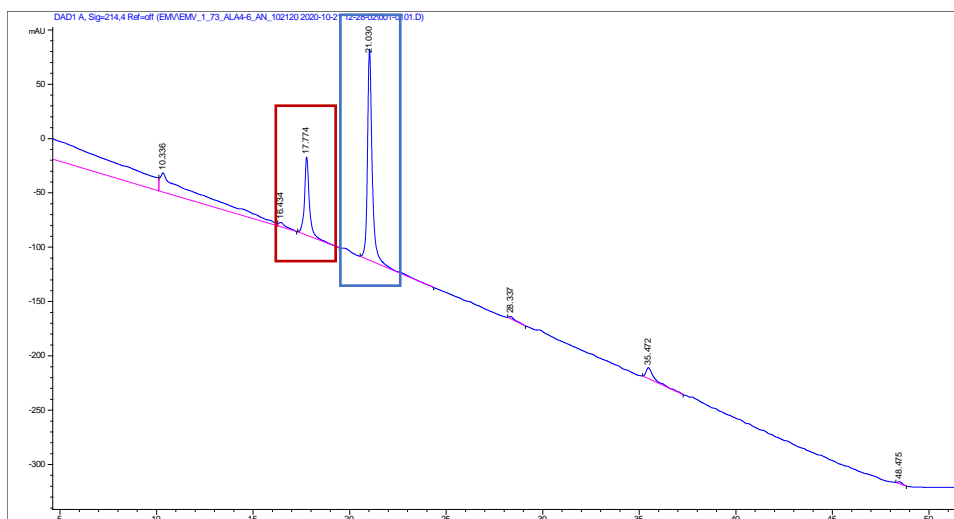
Figure A.12. Structure of analog 34913-Ala4. Structure of simplified lipopeptide analog 34913-Ala4 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Leu at residue position 4 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala4** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

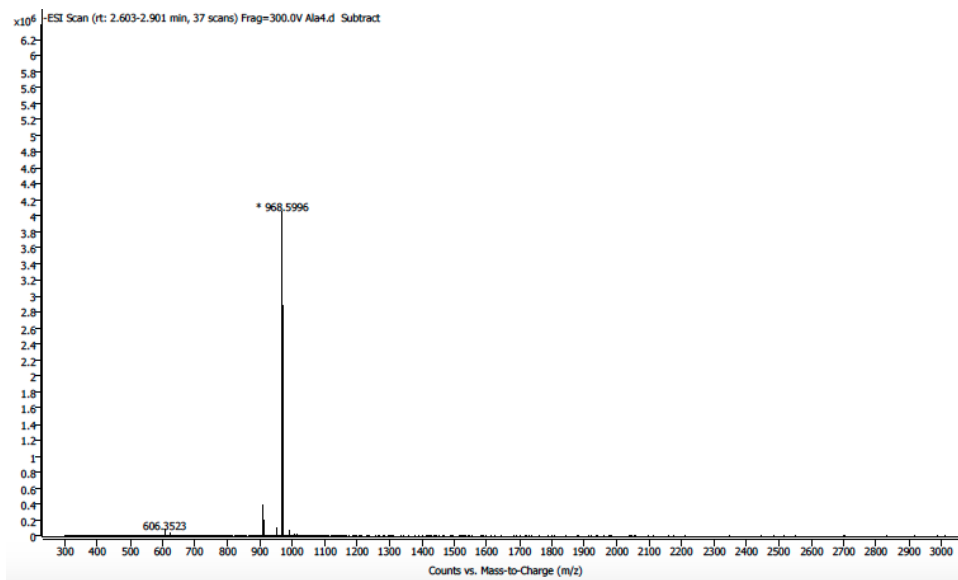
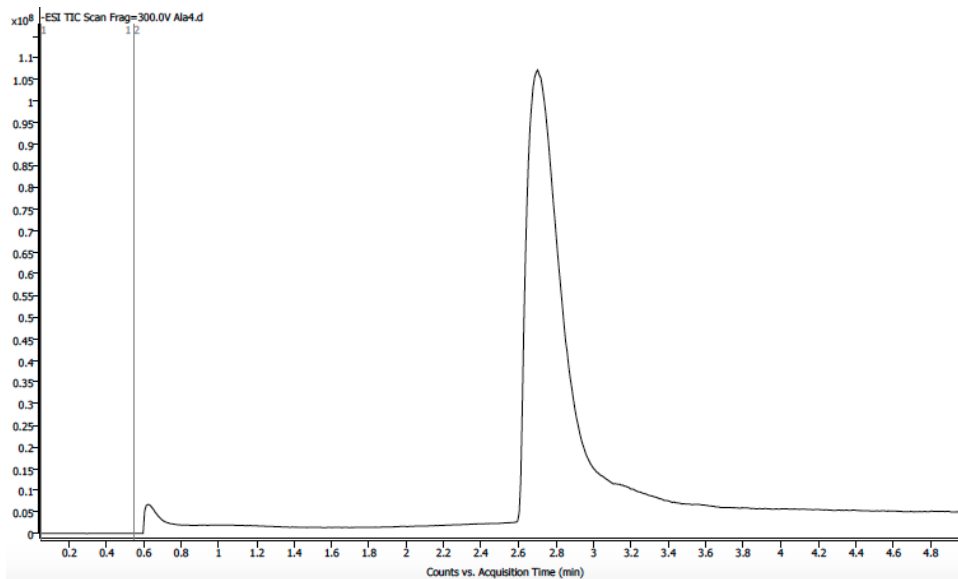
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.408	BV	1.17372e5	2320.76343	0.6500	0.965	10.119
2	10.336	VV	2576.61938	17.64444	1.7517	10.119	16.279
3	16.434	VB	102.91874	4.05054	0.3524	16.279	17.259
4	17.774	BB	1258.41309	71.89717	0.2522	17.312	19.459
5	21.030	BB	3346.90674	194.35297	0.2509	20.545	24.352
6	28.337	BB	38.87243	2.01161	0.2714	28.192	29.099
7	35.472	BB	281.70880	10.14086	0.3941	35.179	37.272
8	48.475	BB	27.52790	1.91591	0.2176	48.285	48.832

Table A.4. Peak integrations of 34913-Ala4. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Ala4 + impurity: 4605.32
Area of 34913-Ala4: 3346.91
Approximate % of 34913-Ala4: 72.7%

Approximate % yield of 34913-Ala4 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala4** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala5

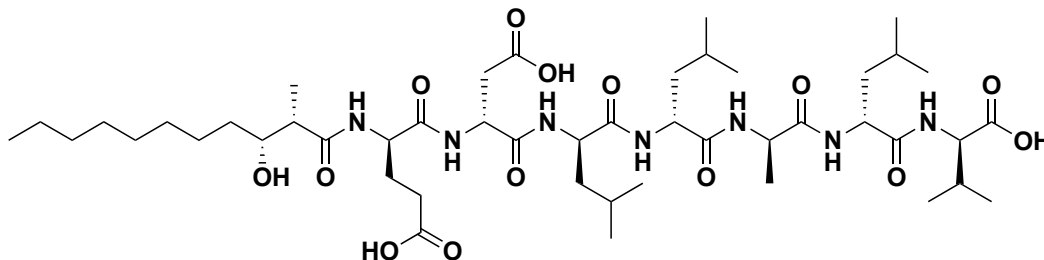
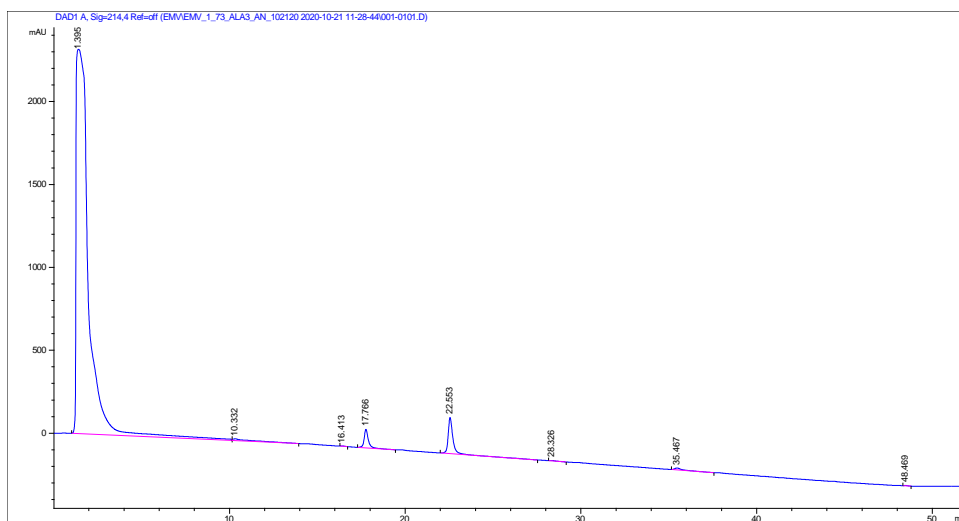


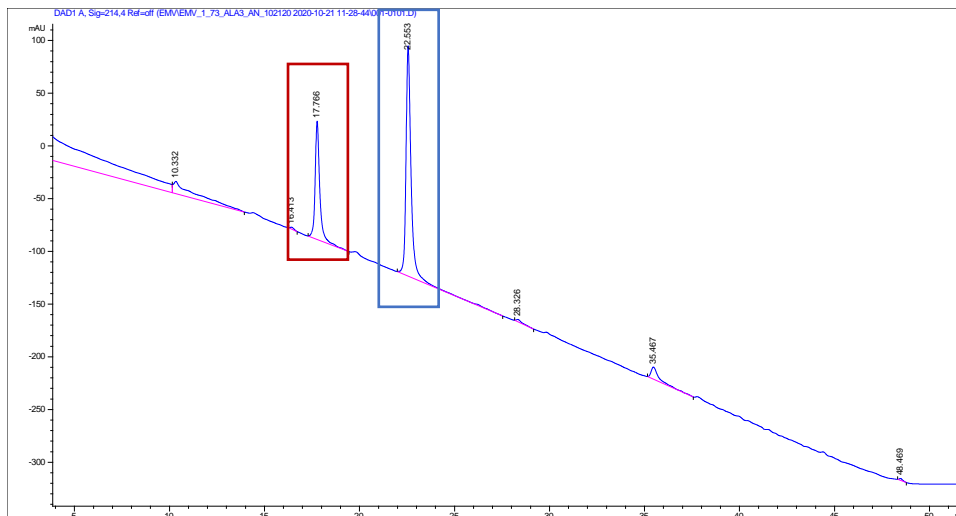
Figure A.13. Structure of analog 34913-Ala5. Structure of simplified lipopeptide analog 34913-Ala5 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Leu at residue position 5 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala5** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((*2S,3R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

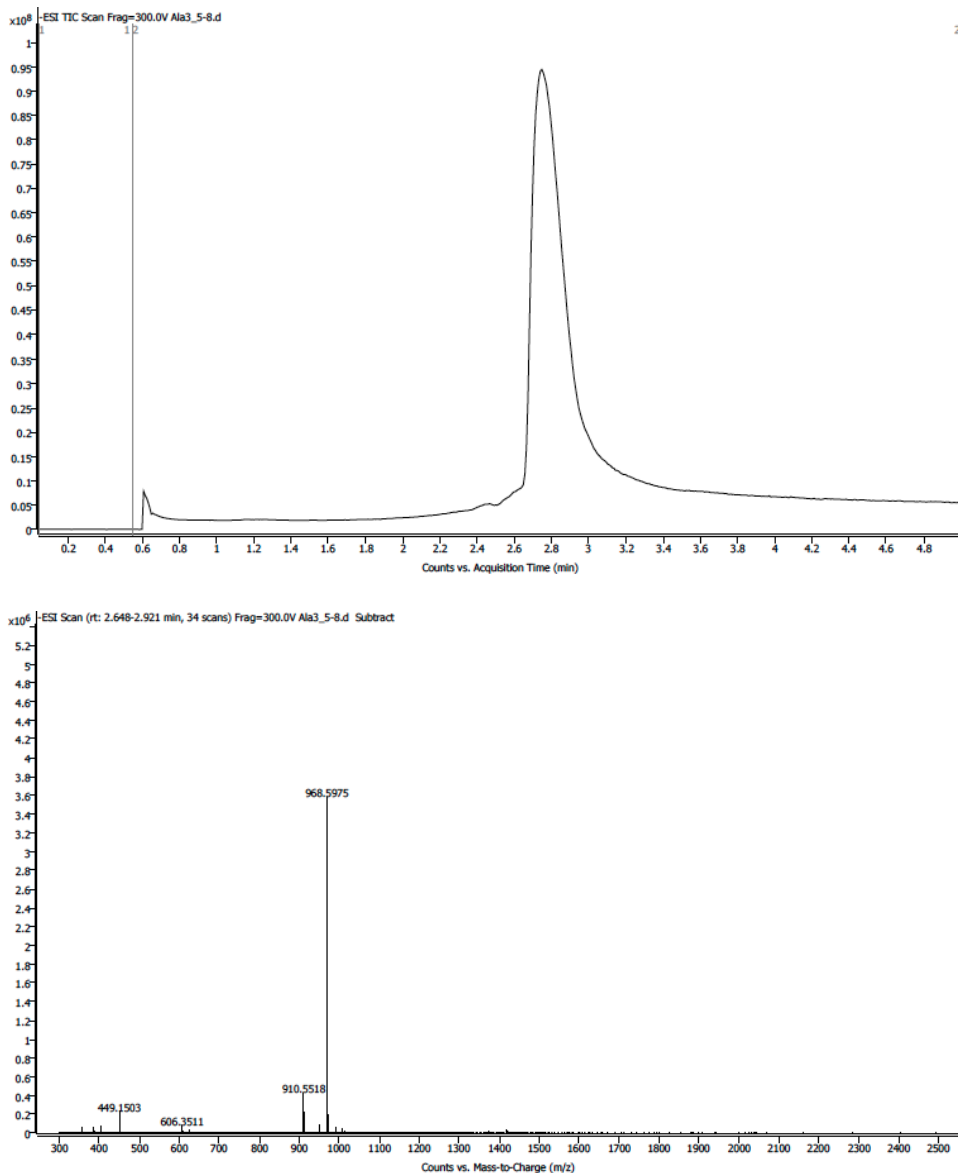
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.395	BV	1.13183e5	2317.45093	0.6178	1.012	10.150
2	10.332	VB	884.16034	11.50993	0.9422	10.150	13.945
3	16.413	BB	22.11676	1.57514	0.2198	16.278	16.718
4	17.766	BB	1883.15149	112.09128	0.2440	17.295	19.452
5	22.553	BB	4049.17920	218.32330	0.2724	21.992	27.538
6	28.326	BB	42.05608	2.15964	0.2792	28.165	29.165
7	35.467	BB	326.42212	11.69078	0.3977	35.152	37.565
8	48.469	BB	21.56048	1.74621	0.1900	48.325	48.785

Table A.5. Peak integrations of 34913-Ala5. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Ala5 + impurity: 5932.33
Area of 34913-Ala5: 4049.18
Approximate % of 34913-Ala5: 68.3%

Approximate % yield of 34913-Ala5 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala5** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala6

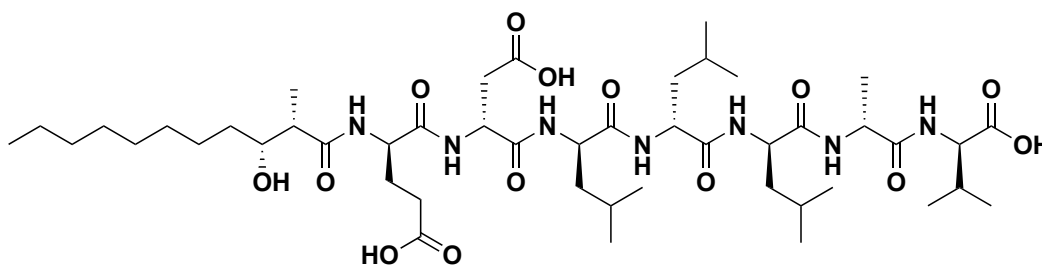
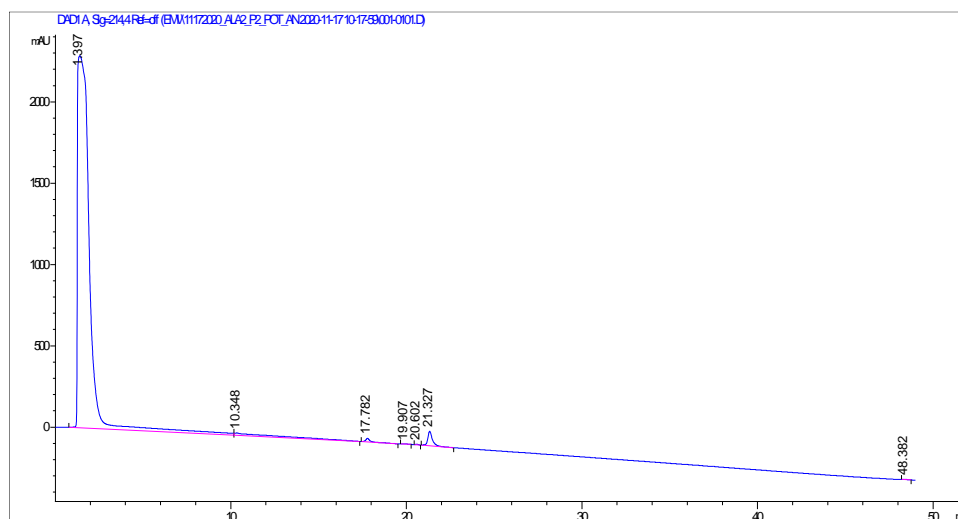


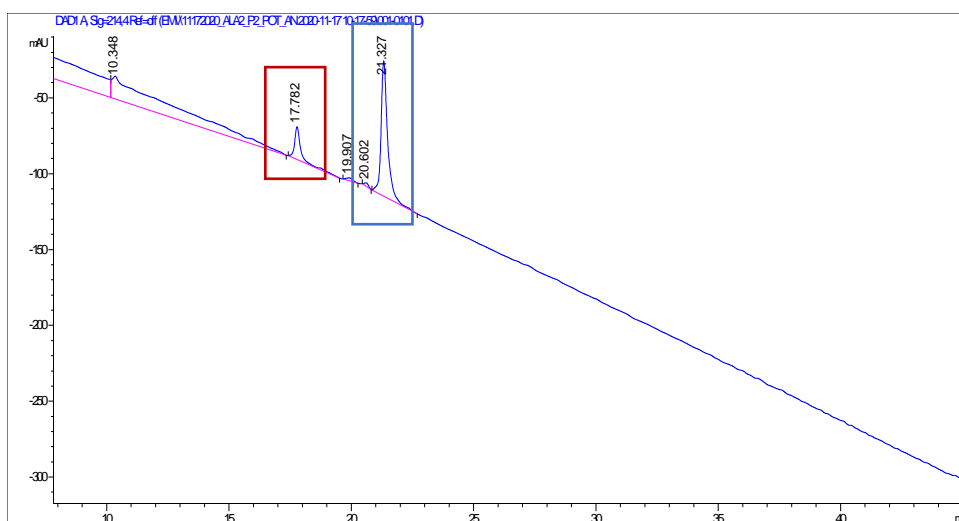
Figure A.14. Structure of analog 34913-Ala6. Structure of simplified lipopeptide analog 34913-Ala6 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Leu at residue position 6 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala6** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

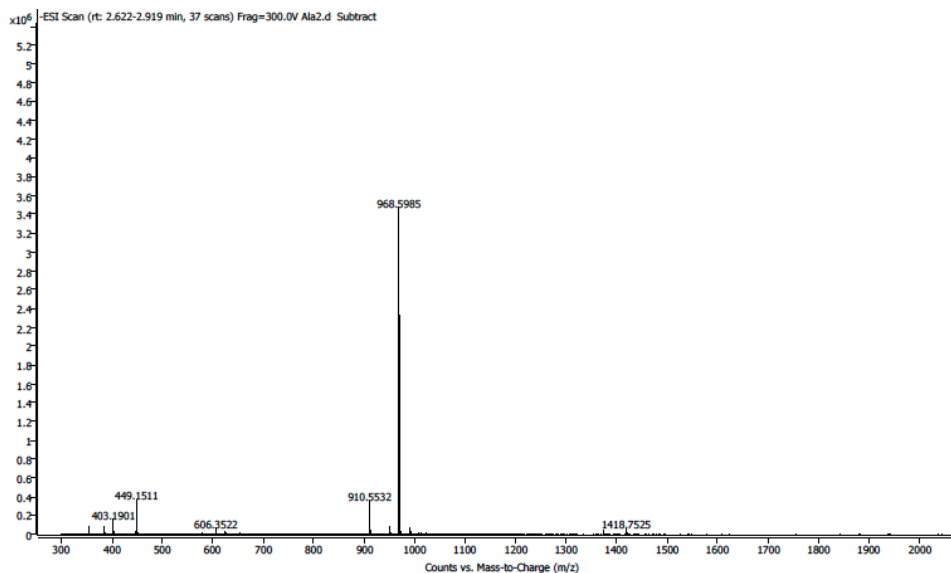
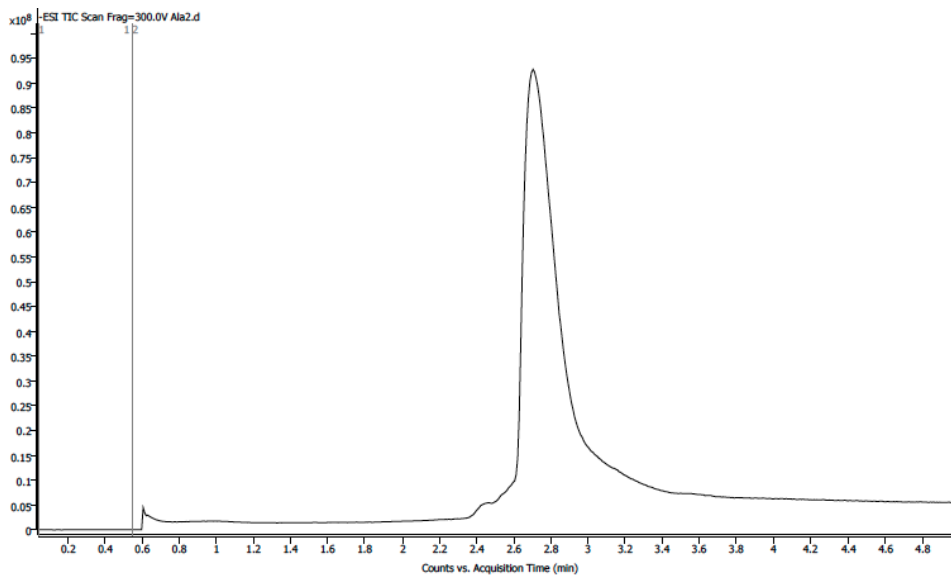
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.397	BV	1.06757e5	2286.08545	0.5928	0.785	10.172
2	10.348	VB	2716.33936	14.99959	2.1568	10.172	17.338
3	17.782	BB	372.32343	21.59072	0.2472	17.425	19.518
4	19.907	BB	36.16009	1.90411	0.2656	19.665	20.271
5	20.602	BB	26.93329	2.31279	0.1859	20.451	20.811
6	21.327	BB	1659.37769	89.24294	0.2729	20.843	22.698
7	48.382	BBA	26.36623	1.86975	0.2145	48.205	48.745

Table A.6. Peak integrations of 34913-Ala6. The peak area based on the analytical HPLC chromatogram as shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Ala6 + impurity: 2031.70
Area of 34913-Ala6: 1659.38
Approximate % of 34913-Ala6: 81.7%

Approximate % yield of 34913-Ala6 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala6** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Ala7

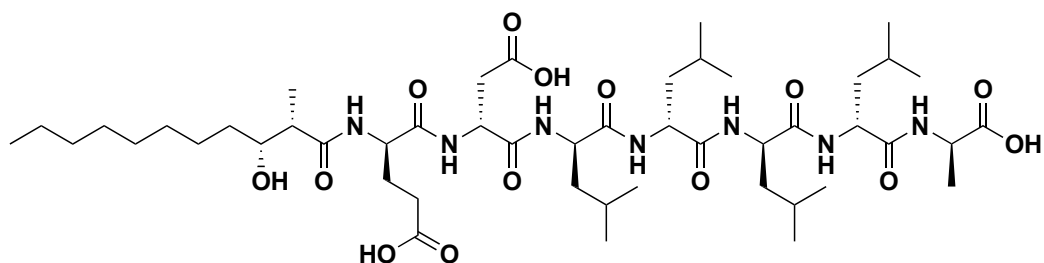
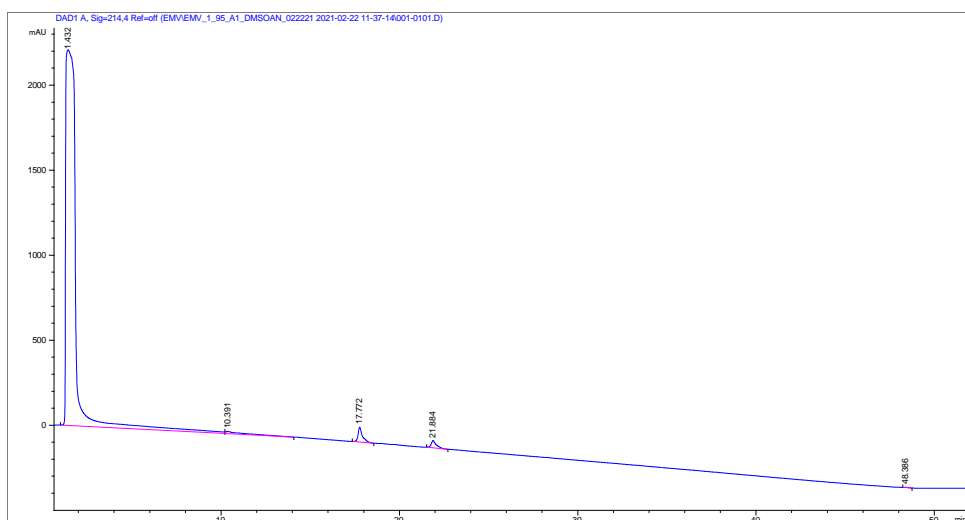


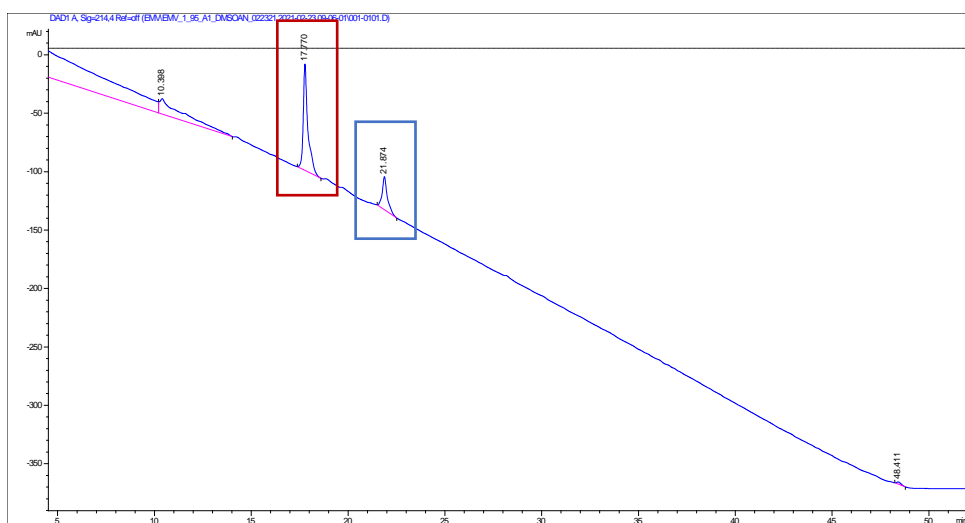
Figure A.15. Structure of analog 34913-Ala7. Structure of simplified lipopeptide analog 34913-Ala7 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Val at residue position 7 substituted with D-Ala.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Ala7** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.424	BV	8.48968e4	2183.59546	0.5076	0.856	10.216
2	10.398	VB	1119.27905	13.11155	1.0379	10.216	14.016
3	17.770	BB	1515.67554	90.99804	0.2343	17.383	18.596
4	21.874	BB	521.03528	28.27304	0.2570	21.503	22.516
5	48.411	BB	26.63870	1.77619	0.2210	48.243	48.790

Table A.7. Peak integrations of 34913-Ala7. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

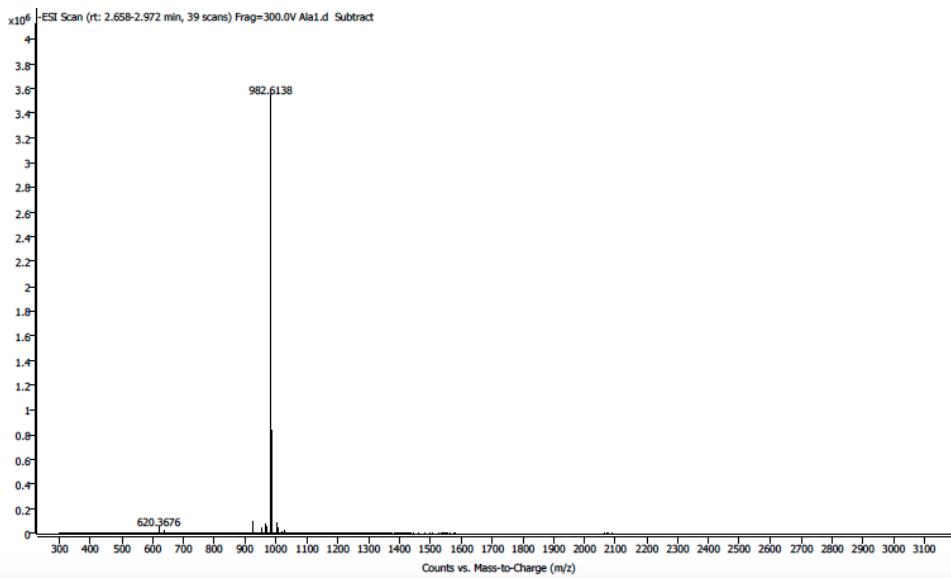
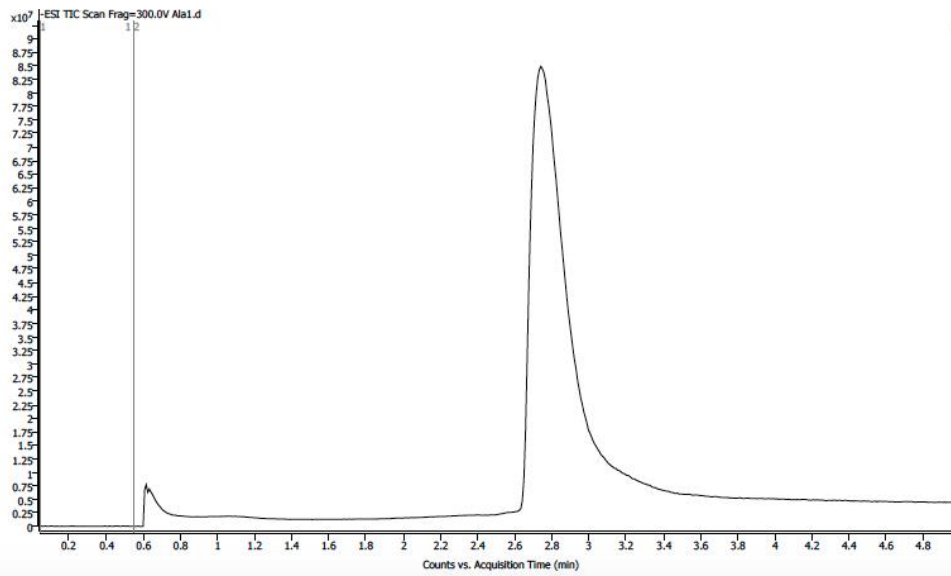
Total Area of 34913-Ala7 + impurity: 2036.72

Area of 34913-Ala7: 521.04

Approximate % of 34913-Ala7: 25.6%

Due to the low purity of 34913-Ala7, this lipopeptide was excluded from figures in Chapter 3. Approximate % yield of 34913-Ala7 is calculated based off the assumption that the molar absorptivities are similar. Definitive calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Ala7** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Gln1

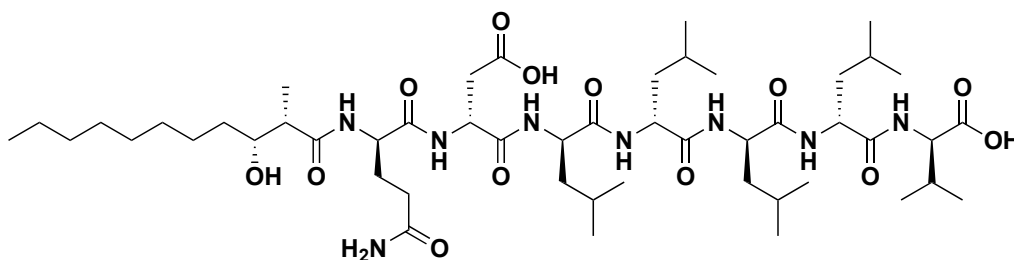
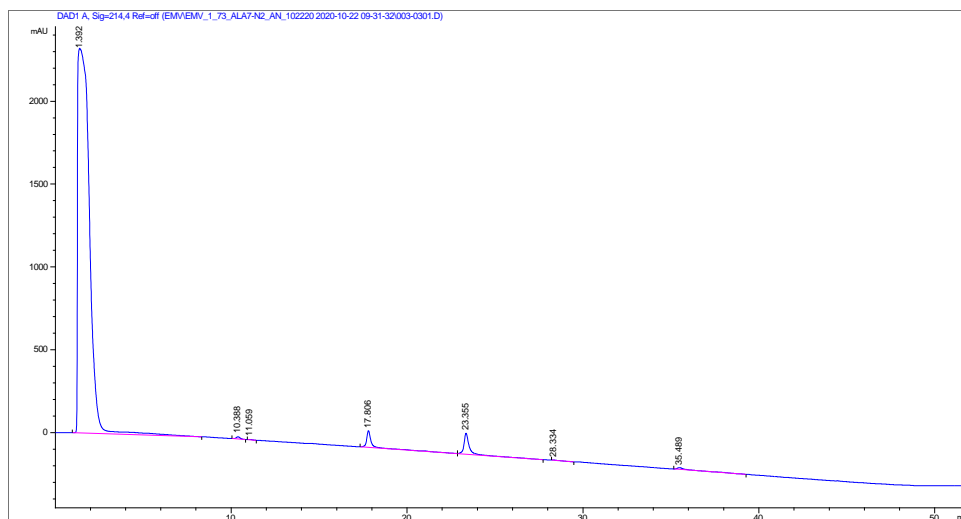


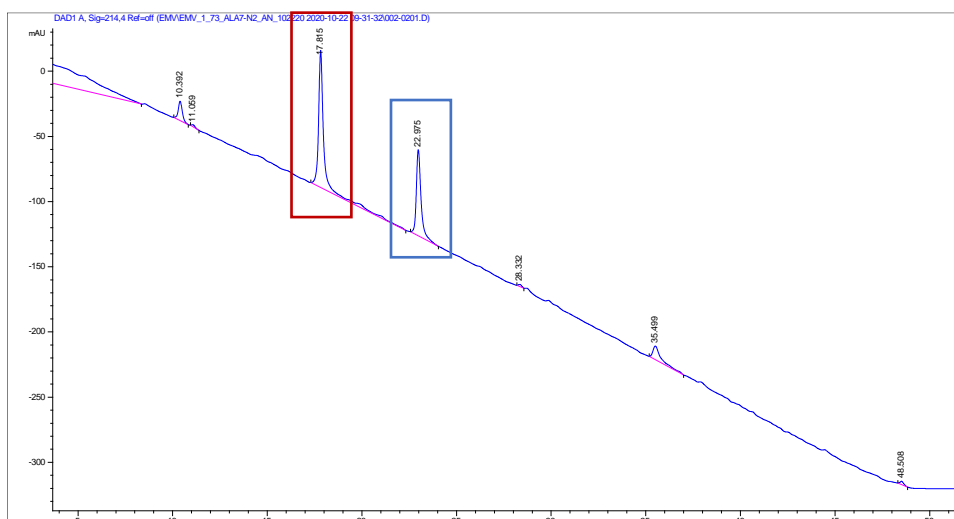
Figure A.16. Structure of analog 34913-Gln1. Structure of simplified lipopeptide analog 34913-Gln1 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Glu at residue position 1 substituted with D-Gln.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Gln1** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

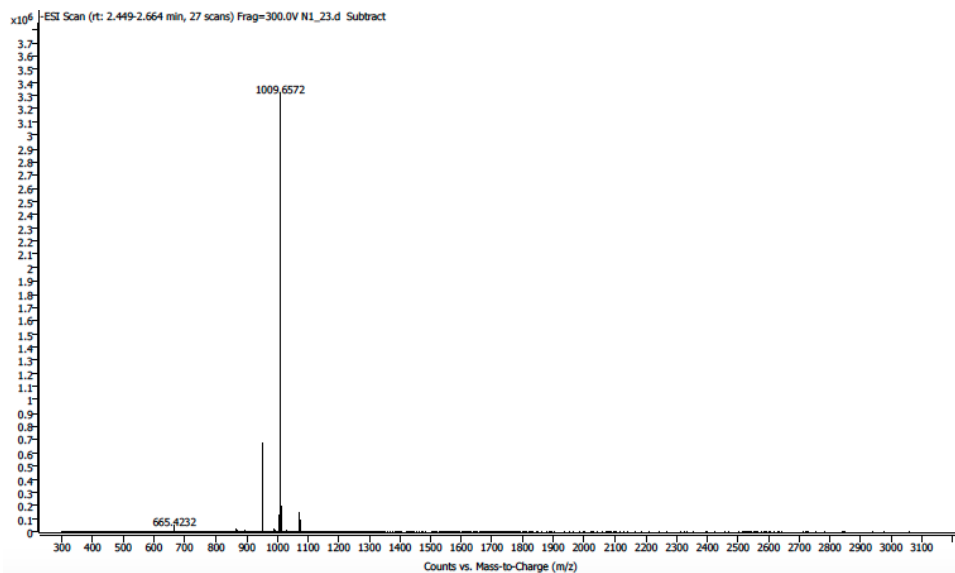
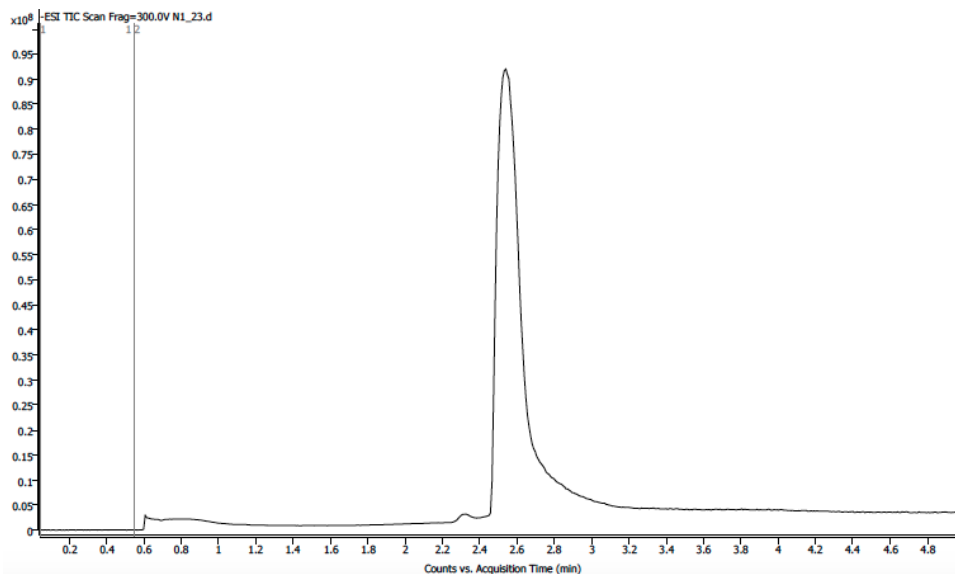
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.392	BB	1.07186e5	2321.28564	0.5847	0.977	8.323
2	10.388	BB	202.13309	13.83399	0.2205	10.050	10.823
3	11.059	BB	22.83535	1.72407	0.2065	10.923	11.430
4	17.806	BV	2033.17627	101.72644	0.2812	17.330	22.875
5	23.355	VB	2675.83545	126.93591	0.2979	22.875	27.730
6	28.334	BB	35.41012	1.41972	0.3430	28.217	29.477
7	35.489	BB	462.50812	10.73733	0.5725	35.170	39.277

Table A.8. Peak integrations of 34913-Gln1. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Gln1 + impurity: 4709.02
Area of 34913-Gln1: 2675.84
Approximate % of 34913-Gln1: 56.8%

Approximate % yield of 34913-Gln1 is calculated based off the assumption that the molar absorptivities are similar. Definitive calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Gln1** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Asn2

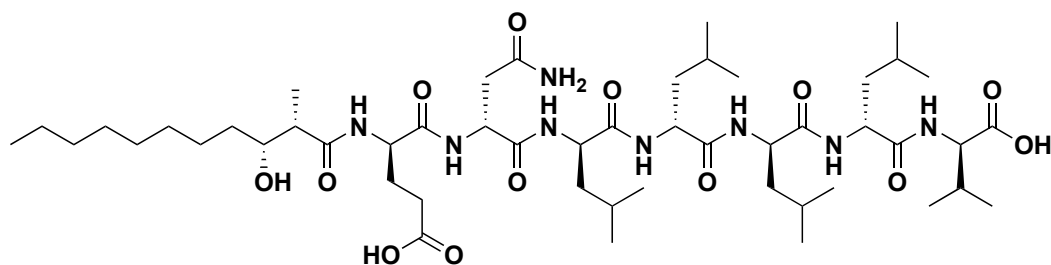
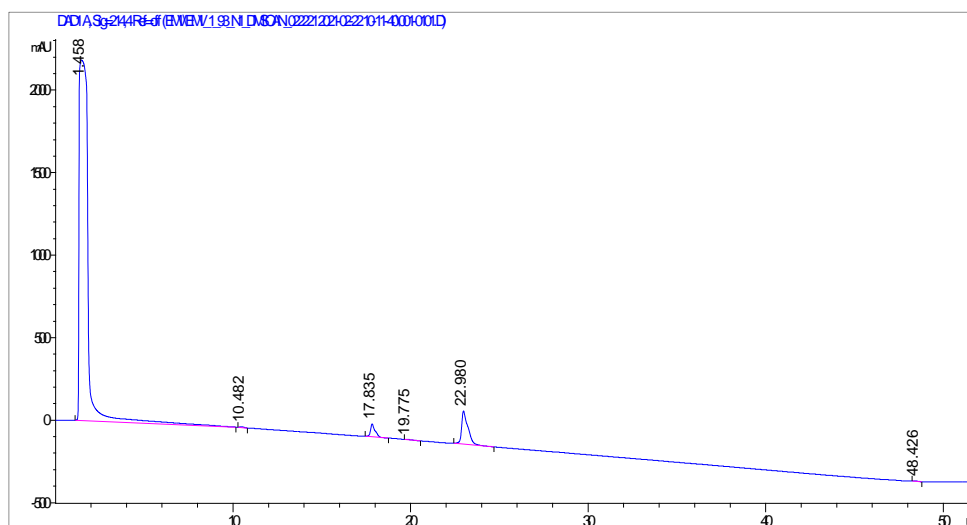


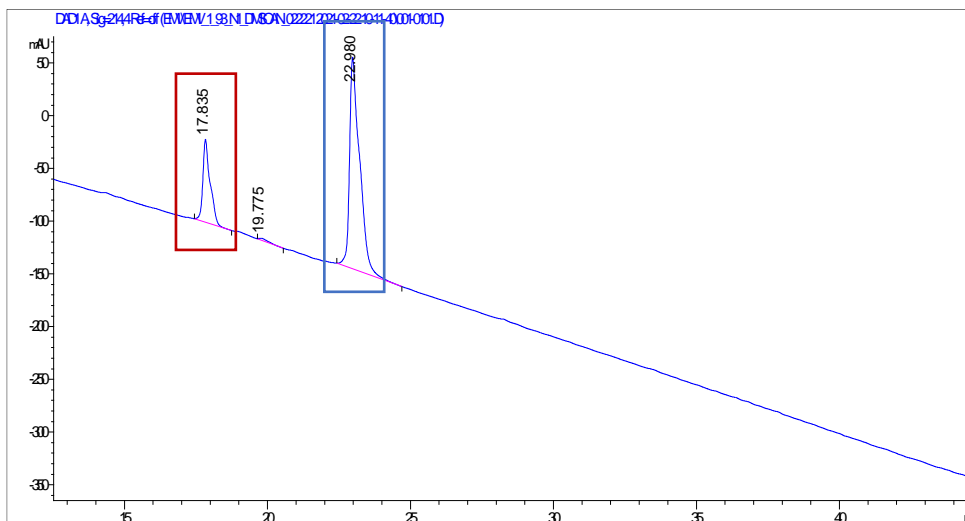
Figure A.17. Structure of analog 34913-Asn2. Structure of simplified lipopeptide analog 34913-Asn2 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Asp at residue position 2 substituted with D-Asn.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Asn2** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

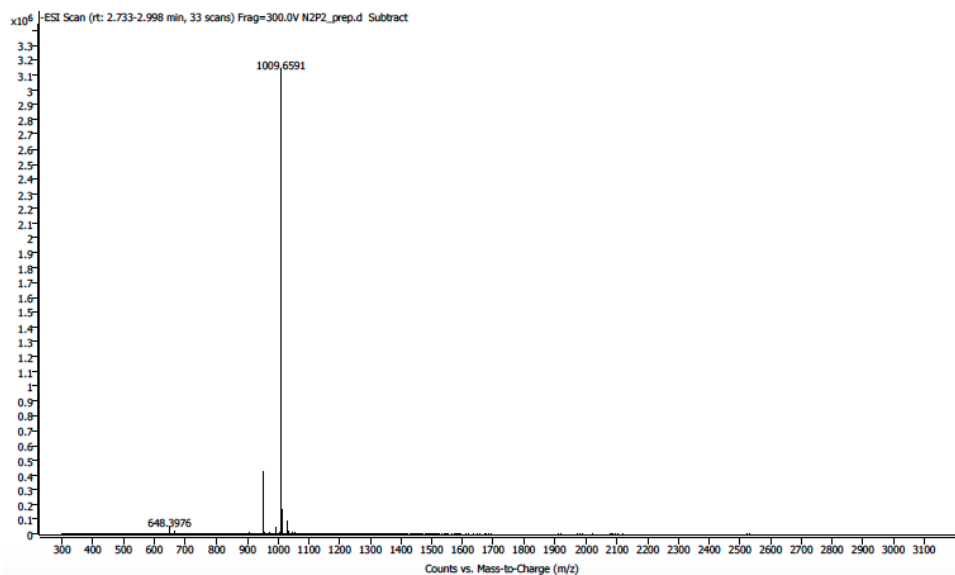
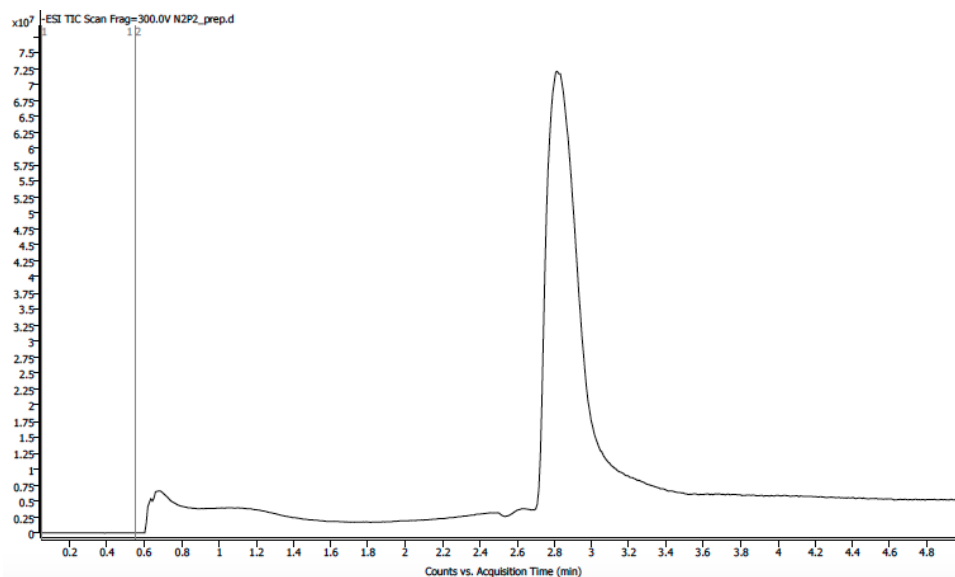
Peak #	Time [min]	Type	Area [mAU*s]	Height [mAU]	Width [min]	Start [min]	End [min]
1	1.458	BB	7.53011e4	2184.61548	0.4541	1.102	10.162
2	10.482	BB	65.08808	4.70058	0.2135	10.282	10.808
3	17.835	BB	1380.39014	78.57597	0.2470	17.448	18.748
4	19.775	BB	39.55629	1.72083	0.3122	19.653	20.555
5	22.980	BB	4940.75098	200.83226	0.3390	22.428	24.702
6	48.426	BB	34.39751	2.34236	0.2213	48.248	48.795

Table A.9. Peak integrations of 34913-Asn2. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

Total Area of 34913-Asn2 + impurity: 6321.14
 Area of 34913-Asn2: 4940.75
 Approximate % Purity of 34913-Asn2: 78.2%

Approximate % yield of 34913-Asn1 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Asn2** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

34913-Gln1Asn2

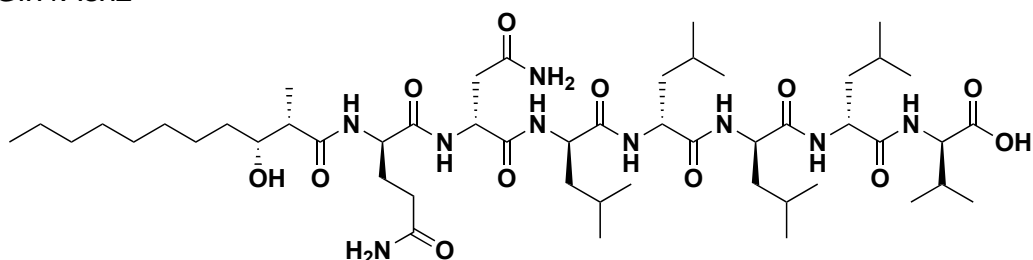
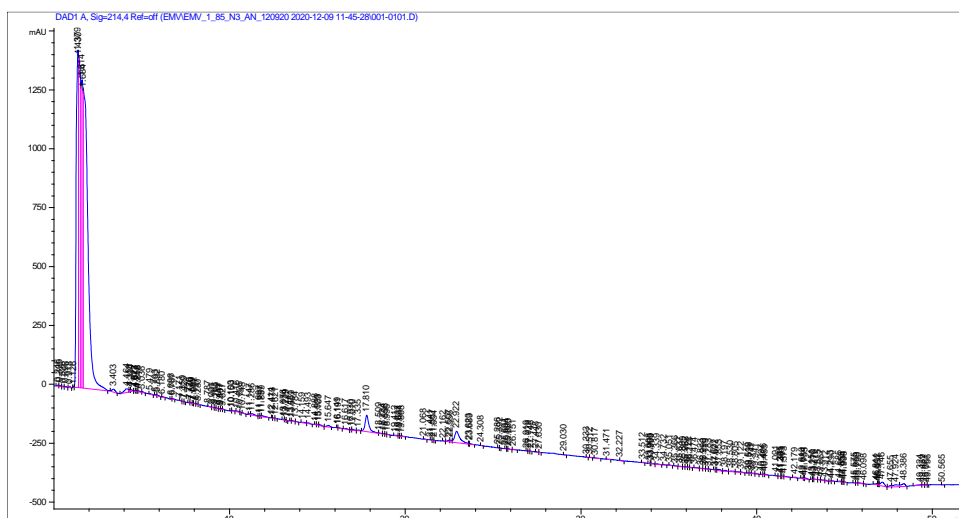


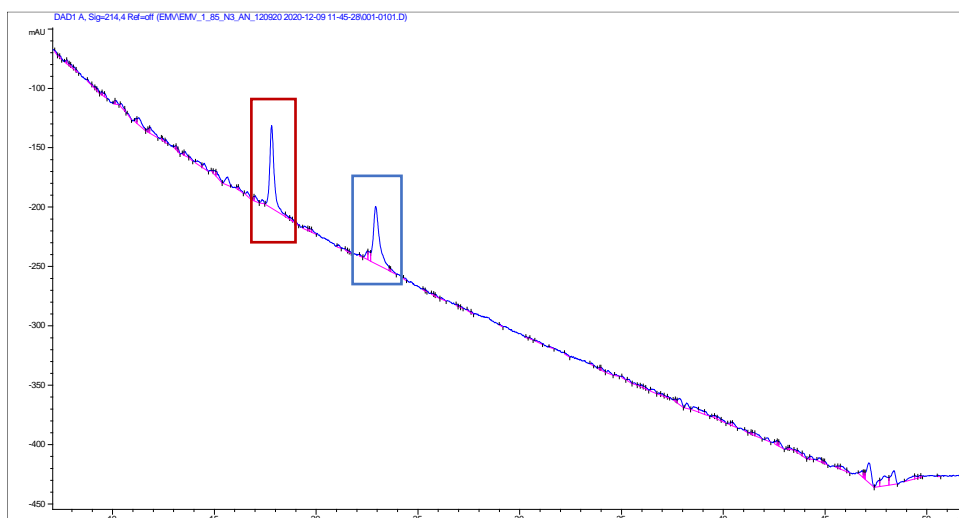
Figure A.18. Structure of analog 34913-Gln1Asn2. Structure of simplified lipopeptide analog 34913-Gln1Asn2 contains a C-terminal amide and a N-terminal attached (2S, 3R)-3-hydroxy-2-methylundecanoic acid. Peptide is composed of only D-amino acids, with the D-Glu at residue position 1 substituted with D-Gln, and the D-Asp at the residue position 2 substituted with D-Asn.

HPLC:

Full Chromatogram:



Zoomed In Chromatogram:



Analytical HPLC chromatogram of **34913-Gln1Asn2** monitored at 214 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-100% acetonitrile over 45 mins. The large peak within the first 5 minutes of the analytical run corresponds to the solvent, DMSO, that was used to dissolve the lipopeptide sample. The peak outlined in red is the impurity, (*R*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, was identified using mass spectrometry. This impurity is from the crude sample of lipophilic tail that was coupled to the N-terminus of each substitution peptide.

Peak Integration Table:

72	17.335	BB	26.89188	3.08642	0.1137	17.212	17.470
73	17.810	VV	1048.24084	69.47610	0.2200	17.505	18.490
74	18.509	VV	20.20171	2.27445	0.1116	18.490	18.703
75	18.734	VV	14.69089	2.28970	0.0869	18.703	18.838
76	18.856	VB	5.98795	1.66394	0.0540	18.838	18.950
77	18.996	BV	9.71192	1.48162	0.0845	18.952	19.145
78	19.412	VV	29.51740	2.42924	0.1498	19.339	19.588
79	19.627	VV	10.13025	2.17873	0.0775	19.588	19.674
80	19.694	VV	14.54758	2.38372	0.0772	19.674	19.778
81	19.808	VV	19.66406	2.60461	0.1001	19.778	20.014
82	21.068	VB	16.36227	2.09926	0.1009	21.030	21.236
83	21.441	VV	6.94470	1.27508	0.0675	21.423	21.534
84	21.557	VV	5.68079	1.38725	0.0598	21.534	21.623
85	21.654	VV	5.67099	1.43302	0.0562	21.623	21.718
86	22.162	VB	7.93497	1.50452	0.0675	22.124	22.290
87	22.507	BV	52.12470	7.22364	0.0922	22.298	22.538
88	22.556	VV	62.77707	7.35386	0.1056	22.538	22.682
89	22.922	VV	1021.24866	48.25405	0.2889	22.682	23.598

Table A.10. Peak integrations of 34913-Gln1Asn2. The peak area based on the analytical HPLC chromatogram us shown here and is used to calculate approximate % purity of sample.

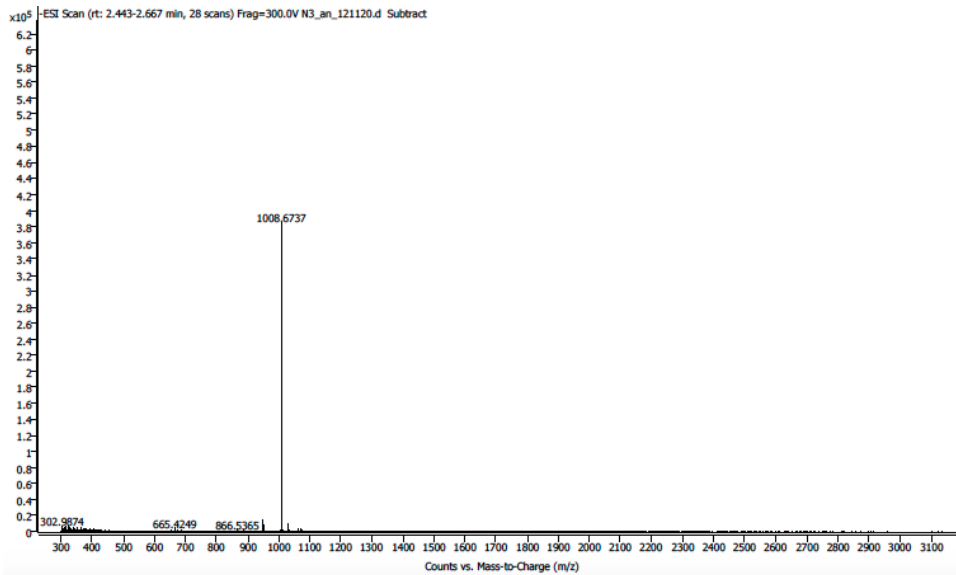
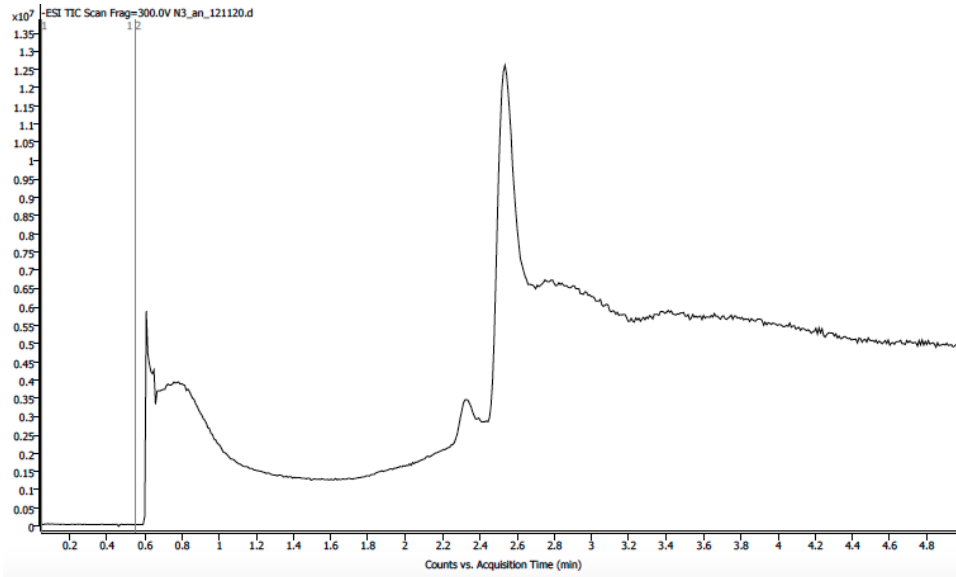
Total Area of 34913-Gln1Asn2 + impurity: 2069.49

Area of 34913-Gln1Asn2: 1021.25

Approximate % Purity of 34913-Gln1Asn2: 49.3%

Approximate % yield of 34913-Gln1Asn2 is calculated based off the assumption that the molar absorptivities are similar. Definite calculations of % yield will require further experimentation.

LC-MS:



LC-MS of **34913-Gln1Asn2** using LC-MS qTOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

Impurity of substitution analogs:
(R)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one

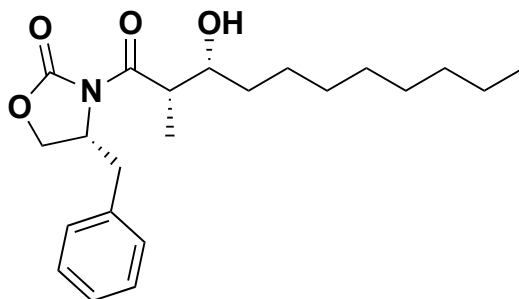
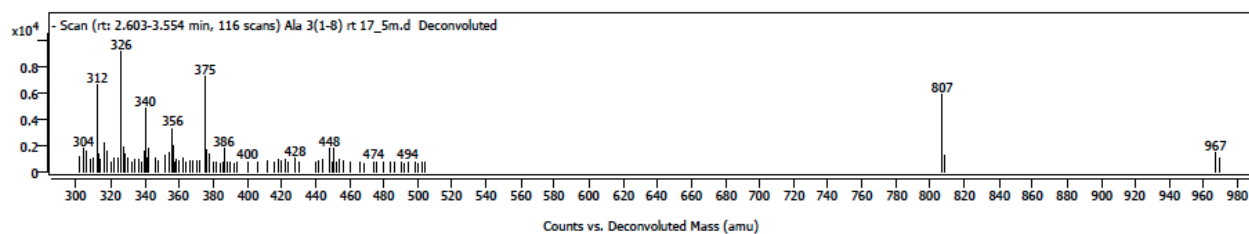


Figure A.19. Structure of substitution analog impurity. Shown is the structure of *(R)*-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, the impurity present in the substitution analogs.

LC MS:



Expected Mass (g/mol)	Observed Mass (g/mol)
375.24	375.2532

Table A.11. Mass of impurity in substitution analogs. Shown is the expected and observed mass of impurity, *(R)*-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylundecanoyl)oxazolidin-2-one, present in the substitution analogs.