

Stony Brook University



OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

**Structure Determination, Enzymatic Characterization, and Inhibitor Discovery of Select
Protein Targets of Bacterial Pathogens**

A Dissertation Presented

by

Brian E. McGillick

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

August 2015

Stony Brook University

The Graduate School

Brian McGillick

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

**Subramanyam Swaminathan – Dissertation Advisor
Senior Scientist, Brookhaven National Laboratory**

**Peter J. Tonge - Chairperson of Defense
Professor of Chemistry, Stony Brook University**

**Robert C. Rizzo – Committee Member of Defense
Professor of Applied Mathematics and Statistics, Stony Brook University**

**F. William Studier – Outside Member of Defense
Senior Scientist Emeritus, Brookhaven National Laboratory**

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

**Structure Determination, Enzymatic Characterization, and Inhibitor Discovery of Select
Protein Targets of Bacterial Pathogens**

by

Brian McGillick

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2015

Our interests are in using structure based drug design to target bacterial agents and toxins that have the potential for intentional release into the environment and therefore pose a severe threat to public health and safety. Specifically, we focus on the pathogenic bacteria *Francisella tularensis* and *Yersinia pestis*, as well as the *Clostridium botulinum* neurotoxins (BoNTs).

Our strategy for *F. tularensis* and *Y. pestis* was to target the bacterial system for fatty acid biosynthesis (FAS-II). The FAS-II system contains several enzymes whose sequence and structure is highly conserved across a vast array of pathogens. This coupled with the low homology to the equivalent system in humans makes this pathway an excellent target for antimicrobial drug development in general, as well as the development of broad-spectrum antibiotics. To this end, we have cloned, expressed, and purified the FAS-II enzyme, β -hydroxyacyl-Acyl Carrier Protein Dehydratase (FabZ), from both *F. tularensis* (FtFabZ) and *Y.*

pestis (YpFabZ). We also performed an enzymatic characterization of both proteins as well as several mutant forms of YpFabZ in order to elucidate the structural features important for substrate and inhibitor binding. Additionally, we have discovered two novel compounds, Mangostin and Stictic Acid, which demonstrate potent inhibition of both YpFabZ and FtFabZ. The scaffolds of these two compounds represent excellent starting points for further design of broad-spectrum antimicrobial drugs.

The BoNTs are zinc metalloproteases that cleave and inactivate proteins critical for neurotransmission. There are seven serotypes of the BoNTs, several of which are toxic to humans. In order to adequately screen potential drugs targeting the BoNTs we developed a novel fluorescence assay that utilizes a recombinant GFP-bound substrate and is capable of detecting enzymatic cleavage by all seven BoNT serotypes. Using this new assay, in conjunction with a virtual screen of 1.4 million potential inhibitor compounds, we discovered a novel inhibitor of BoNT serotype E.

Table of Contents

List of Figures.....	viii
List of Tables	ix
List of Schemes.....	x
List of Abbreviations	xi
Acknowledgements	xv

Chapter 1: Bacterial Protein Targets for Structure Based Drug Design..... 1

Structure based drug design	1
Fatty acid biosynthesis type II (FAS II) pathway	3
β -hydroxyacyl-ACP dehydratase (FabZ)	6
Known inhibitors of β -hydroxyacyl-ACP dehydratase (FabZ)	9
<i>Clostridium botulinum</i> neurotoxins	12
The Structure and catalytic mechanism of Botulinum neurotoxin light chain.....	14
BoNT light chain interactions with substrate	16
Known inhibitors of the Botulinum neurotoxin light chain.....	17
BoNT/A inhibitors	17
BoNT/B inhibitors	18
BoNT/E inhibitors.....	19
BoNT/F inhibitors.....	19
Current assays of BoNT light chain endopeptidase activity	22
Research project overview	25

Chapter 2: β -Hydroxyacyl-Acyl Carrier Protein Dehydratase (FabZ) from *Francisella tularensis* and *Yersinia pestis*: Structure Determination, Enzymatic Characterization, and Cross Inhibition Studies..... 27

Introduction	27
Methods	29
Cloning, expression and purification	29
Crystallization and structure determination	30
Enzymatic characterization	33
Docking studies	33
Results and Discussion	34
Structure of YpFabZ	34
Structure of FtFabZ	38
Biochemical properties of FtFabZ and YpFabZ	41
YpFabZ (H19A) mutant study	43

Inhibitor screening and evaluation.....	44
Mangostin docking study.....	45
Stictic Acid docking study.....	46
Inhibition studies of known FabZ inhibitors	47
Conclusion and Future Directions	54
Design and synthesis of novel FabZ inhibitors.....	54
Inhibitors targeting the ACP binding groove of FabZ.....	56
Chapter 3: A Recombinant Fluorescent Substrate With Cleavage Sites of All Seven BoNT Serotypes For High Throughput Screening of Inhibitor Compounds	58
Introduction	58
Methods	59
Cloning, expression, and purification of repcon	59
Cloning, expression, and purification of BoNT/A-F light chains	62
Repcon solution studies	62
Fluorescence based continuous assay	63
Results.....	63
In-solution Repcon digest	63
Repcon fluorescence assay	65
Fluorescence assay reproducibility	66
Testing of known inhibitors using Repcon fluorescence assay	68
Discussion	70
Conclusion and Future Directions	71
Chapter 4: Lead Identification Targeting Botulinum Neurotoxin Serotype E Using Computational and Biochemical Methods	74
Introduction	74
Methods	76
SNAP-Etide assay.....	76
Repcon fluorescence assay	77
Results and Discussion	77
Virtual screen.....	77
SNAP-Etide screen	78
Repcon screen	79
Structural analysis of compound C562-1101.....	81
Conclusions and Future Directions.....	85
Appendix A: Botulinum Neurotoxin Serotype C (BoNT/C): Substrate Recognition and Cleavage Site Specificity.....	86
Introduction	86
Results.....	92
Cloning, expression, and purification of BoNT/C and SNAP-25 constructs.....	92
Crystallographic studies.....	94
Appendix B: Cloning, Purification, and Inhibition Studies of β-hydroxyacyl-ACP dehydrase from <i>Burkholderia pseudomallei</i> (BpFabZ) and Additional Inhibitor Studies....	98

Results	98
Cloning of β -hydroxyacyl-ACP dehydrase from <i>Burkholderia pseudomallei</i>	98
Purification of β -hydroxyacyl-ACP dehydrase from <i>Burkholderia pseudomallei</i>	98
Preliminary inhibition data for select compounds against YpFabZ, FtFabZ, and BpFabZ	100
 Bibliography	 103

List of Figures

Figure 1.1 Iterative process of structure based drug design.....	3
Figure 1.2 Fatty acid synthesis pathway	5
Figure 1.3 Structure of the FabZ dimer.....	8
Figure 1.4 Dual Pose of FabZ Inhibitors.....	10
Figure 1.5 The chemical structures of known FabZ inhibitors	11
Figure 1.6 The structure of BoNT/A light chain	14
Figure 1.7 The structure of the BoNT/A-SNAP-25 complex	17
Figure 1.8 The chemical structures of known small molecule inhibitors of BoNT/A	21
Figure 1.9 The chemical structures of known BoNT/B and BoNT/E inhibitors	22
Figure 2.1 The structural arrangement of YpFabZ	35
Figure 2.2 The YpFabZ active site	37
Figure 2.3 Structural alignment of YpFabZ and FtFabZ	40
Figure 2.4 Kinetic analysis of YpFabZ and FtFabZ	42
Figure 2.5 The effect of NaCl concentration on YpFabZ activity	43
Figure 2.6 Kinetic analysis of YpFabZ H19A mutant.....	44
Figure 2.7 Double reciprocal plots for Mangostin and Stictic Acid inhibitors	45
Figure 2.8 The DOCK poses of Mangostin and Stictic Acid.....	47
Figure 2.9 Sequence alignment of FabZ isoforms	49
Figure 2.10 Structural alignment of YpFabZ and FabA	51
Figure 2.11 Hypothetical binding mode of conjugate inhibitor of YpFabZ	55
Figure 2.12 Library of FabZ conjugate inhibitors	56
Figure 2.13 FabZ interactions with ACP and the ACP binding groove.....	57
Figure 3.1 Sequence and domain architecture of Repcon substrate.....	64
Figure 3.2 In-solution digest of Repcon	65
Figure 3.3 Fluorescence signal obtained from Repcon digest	66
Figure 3.4 Repcon fluorescence assay reproducibility	67
Figure 3.5 Repcon fluorescence signal dependence on BoNT enzyme concentration	68
Figure 3.6 Inhibition data for known BoNT/E and BoNT/F inhibitors	69
Figure 4.1 The structure of peptide inhibitor, RIME, in complex with BoNT/E.....	76
Figure 4.2 The chemical structures and % inhibition of compounds from SNAP-Etide screen....	79
Figure 4.3 Raw BoNT/E inhibitor screening results.....	80
Figure 4.4 IC ₅₀ curves for compounds NSC-77053 and C562-1101	81
Figure 4.5 DOCK pose of compound C562-110 in BoNT/E active site.....	83
Figure 4.6 Footprint plot of compound C562-1101	84
Figure A.1 Schematic representation of SNAP-25	88
Figure A.2 Schematic representation of Syntaxin	89
Figure A.3 Proposed mechanism of BoNT/C dual substrate specificity.....	91
Figure A.4 The structure of BoNT/C light chain with cadmium metal ion	97
Figure B.1 Refolding strategy for BpFabZ.....	100

List of Tables

Table 1.1 Target molecule and cleavage site location of the BoNT serotypes	13
Table 1.2 Peptide and peptidomimetic BoNT inhibitors	20
Table 2.1 Crystal data and refinement statistics for YpFabZ and FtFabZ	32
Table 2.2 Kinetic values for YpFabZ and FtFabZ	42
Table 2.3 Inhibitory activities of select compounds against YpFabZ and FtFabZ	53
Table A.1 Primers for design of BoNT/C and SNAP-25 constructs.....	94
Table A.2 Summary of BoNT/C light chain crystallographic data.....	96
Table B.1 Primers used for BpFabZ cloning	98
Table B.2 Inhibitory activities of select compounds against YpFabZ, FtfabZ, and BpFabZ	102

List of Schemes

Scheme 1.1 The conversion of β -hydroxyacyl-ACP to trans-2-enoyl-ACP	7
Scheme 1.2 Enzymatic mechanism of FabZ	9
Scheme 1.3 Enzymatic mechanism of BoNT proteolysis	15
Scheme 1.4 The reaction of sulfhydryl groups with maleimide	25

List of Abbreviations

AccABCD	Acetyl-CoA carboxylase
ACP	Acyl carrier protein
acyl-ACP	Acylated acyl carrier protein
BoNT	Botulinum neurotoxin
BoNT/A	Botulinum neurotoxin serotype A
BoNT/B	Botulinum neurotoxin serotype B
BoNT/C	Botulinum neurotoxin serotype C
BoNT/D	Botulinum neurotoxin serotype D
BoNT/E	Botulinum neurotoxin serotype E
BoNT/F	Botulinum neurotoxin serotype F
BoNT/G	Botulinum neurotoxin serotype G
<i>B. pseudomallei</i>	<i>Burkholderia pseudomallei</i>
BtFabZ	FabZ from <i>Burkholderia thailandensis</i>
CjFabZ	FabZ from <i>Campylobacter jejuni</i>
CoA	Coenzyme A
CFP	Cyan fluorescent protein
DMSO	Dimethyl Sulfoxide
<i>DAB</i>	2,4-diaminobutanoic acid
DABCYL	4-((4-(dimethylamino)phenyl)azo) benzoic acid
<i>DNP-DAB</i>	4-(2,4-dinitrophenylamino)-2- amino-butanoic acid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>

ELISA	Enzyme-linked immunosorbent assay
<i>F. tularensis</i>	<i>Francisella tularensis</i>
FabA	β -hydroxylacyl-ACP dehydratase
FabD	Malonyl-CoA:ACP transacylase
FabF	β -ketoacyl-ACP synthase II
FabG	β -ketoacyl-ACP reductase
FabH	β -ketoacyl-ACP synthase III
FabI	Enoyl-ACP reductase
FabZ	β -hydroxylacyl-ACP dehydratase
FAS	Fatty acid synthase
FAS-I	Eukaryotic fatty acid biosynthesis
FAS-II	Bacterial fatty acid biosynthesis
FITC	fluorescein isothiocyanate
FPS	Footprint similarity scoring
FRET	fluorescence resonance energy transfer
FtFabZ	FabZ from <i>Francisella tularensis</i>
GFP	Green Fluorescent Protein
HC	Heavy chain
HpFabZ	FabZ from <i>Helicobacter pylori</i>
HPLC	High performance liquid chromatography
IC ₅₀	Half maximal inhibitory concentration
IPTG	Isopropyl-1-thio- β -D-galactopyranoside
k _{cat}	Turnover number

k_f	Equilibrium constant, forward reaction
k_r	Equilibrium constant, reverse reaction
K_i	Inhibition constant
K_M	Michaelis-Menton constant for the substrate
LC	Light chain
LB	Luria-Bertani Agar
NmFabZ	FabZ from <i>Neisseria meningitides</i>
NMR	Nuclear Magnetic Resonance
NTA	Nitrilotriacetic acid
OD ₆₀₀	Optical density at 600 nm
PaFaBZ	FabZ from <i>Pseudomonas aeruginosa</i>
PEG	Polyethylene glycol
PfFabZ	FabZ from <i>Plasmodium Falciparum</i>
PCR	Polymerase chain reaction
PDB	Protein Data Bank
RFU	Relative Fluorescence Units
SDS	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNAP-25	Synaptosomal-associated protein 25
SNARE	Soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptors
VAMP	Vesicle associated membrane protein
V_i	Initial velocity
V_{max}	Maximum velocity
wt	wild type

YFP Yellow fluorescent protein

Y. Pestis *Yersinia Pestis*

YpFabZ FabZ from *Yersinia Pestis*

Acknowledgments

I would like to express my sincere gratitude to my advisor, Dr. Subramanyam Swaminathan, for his endless patience and support throughout my graduate studies. The time I spent in his lab has been the most memorable and rewarding experience of my life, having been given the rare opportunity to learn and grow as a scientist with the perfect blend of freedom and guidance. He has become not just a professional mentor, but a personal mentor as well, and it is my hope that he understands the depth of my gratitude. Dr. Swaminathan is a scientist of rare intelligence and integrity whose example I will continue to strive for throughout my scientific career.

I would like to thank Dr. Kumaran and Dr. Eswaramoorthy who have become both colleagues and friends during my time in the Swaminathan Lab. They have spent countless hours imparting as much of their skill and expertise onto me as possible, a sometimes arduous task for which I am eternally thankful. I could not have possibly completed my graduate studies without their continuous direction, humor, and positive attitude. It is my hope that they will continue to be lifelong friends and mentors.

I would like to thank the chair of my dissertation committee, Dr. Peter Tonge, for his expert guidance throughout the formidable years of my graduate studies. Only a year ago it seemed unlikely that I would be able to defend on time, but his efforts and advice have made it possible. I would also like to specifically thank Dr. Tonge for helping me to solve a scientific problem that I had been stuck on for several months and threatened to derail my entire project. Without his help I would still be doing endless V_i calculations.

I would like to thank my committee member, Dr. Robert Rizzo, who was also my masters degree advisor. When I joined Dr. Rizzo's lab, I had only a passing interest in research, but by the time I left I could think of no other career path I wanted to pursue more. Also, Dr. Rizzo's efforts on my behalf are almost solely responsible for my acceptance into the MSTP program, a fact which I will never forget and will always be grateful for.

I would like to thank my committee member, Dr. Bill Studier, for always being available to answer any and all questions I had when we passed one another in the halls. I know he was often busy and had better things to do, but it never stopped him from giving me all the time I needed. Whether he realizes it or not, Dr. Studier's advice helped shape several of my projects and his meticulous approach to science is something I try to emulate.

I would like to thank all of our collaborators in the Rizzo, Tonge, Ojima, and Ichchenko Labs. It has been an honor to work on so many interesting projects with these talented and dynamic scientists. I would also like to specifically thank Weixuan Yu for volunteering her time to help me with my project, which had a tremendous impact on my success.

I would like to thank Dr. Paul Freimuth, who has basically been an unofficial member of my thesis committee based on the number of hours he has spent helping me with my projects.

Dr. Freimuth's guidance, particularly on matters of protein expression and purification, are greatly appreciated.

I would like to thank my undergraduate student, Casey Vieni, who was a major asset to all of my projects. Casey has been an unwavering hard worker and is already a talented researcher.

I would like to thank the Stony Brook Medical Scientist Training Program and its chair, Dr. Mike Frohman, for their years of support and guidance.

I would also like to thank my beautiful and talented girlfriend, Jamie, not only for putting up with me still being in school, but for putting up with me in general. My success is due in no small part to her presence in my life and finding her will always be my greatest discovery.

Last, but certainly not least, I would like to thank my parents for their constant and unconditional support throughout my entire life. I know it hasn't always been easy. I promise one day soon you will be able to tell someone I am somewhere other than, "still in school". I'd also like to thank my brothers, Mike and Tom, who are my two best friends.

Chapter 1: Bacterial Protein Targets for Structure Based Drug Design

Structure based Drug design

The use of protein structures for the development of novel drugs is a relatively new enterprise with the first examples having been published only a few decades ago.¹ Today, although it remains an imperfect process, structure based drug design has become the bedrock of the drug discovery process both in academia and in industry. Recent advances in x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy have allowed these fields, which were once primarily accessible only to physicists and mathematicians, to become truly interdisciplinary. As a result, we have witnessed the rapid expansion of the known body of protein structures, and with it the availability of potentially suitable drug targets.

Proper selection of a good target is the first and most important step in the discovery and design of new drugs.² In the case of antimicrobial drug targets, there are several key elements to consider. First, it is imperative that the drug target be essential, and ideally carries out a unique function in the microbe being considered. This will help to ensure that inhibition of the drug target will lead to the death of the pathogen. Additionally, it is crucial that the drug target not share strong homology with any host proteins to limit potential toxic side effects. If significant homology exists, it will be necessary to exploit any key differences between the host and pathogen protein. Lastly, the target protein should have a well-defined active site that can be bound by a small molecule or short peptide inhibitor. Satisfaction of these key elements to target selection greatly increases the chance of a successful drug discovery effort.

Once an appropriate target is selected, the process of structure based drug design is iterative, typically beginning with the successful isolation of the target molecule in solution (Fig. 1.1). For bacterial targets, this is most commonly accomplished through molecular cloning of a recombinant gene, followed by high-level expression and purification of the protein of interest. Once the protein is isolated, its structure is determined using x-ray crystallography, NMR, or homology modeling. After structure determination, the search for an initial lead commences with the goal of finding a compound that demonstrates at least micromolar inhibition *in vitro*. Initial leads can be discovered via simple high throughput (HT) screening of compounds selected from a database based on their size and shape complementarity to the active site of the target molecule. More commonly a virtual screen is performed prior to any biochemical testing, particularly in an academic setting where access to instruments to automate the HT screening process is limited. Virtual screening is conducted via molecular docking whereby thousands of candidate compounds are positioned algorithmically into the target molecule's active site and scored based on their electrostatic and steric interactions.³ The highest ranked compounds, based on the lowest binding energy calculated from a scoring function, can then be purchased or synthesized and tested in a biochemical assay. Ideally, compounds identified through the initial screen will be co-crystallized with the protein target allowing key interactions to be visualized, thus providing insights into which functional groups can be optimized to increase potency. Several cycles of drug optimization followed by co-crystallization may be necessary before a compound with nanomolar or better inhibition is discovered. At this stage, further testing in cell based

assays to check for potency and bioavailability will be necessary. Any compounds that satisfactorily meet these benchmarks are ready for eventual clinical testing.

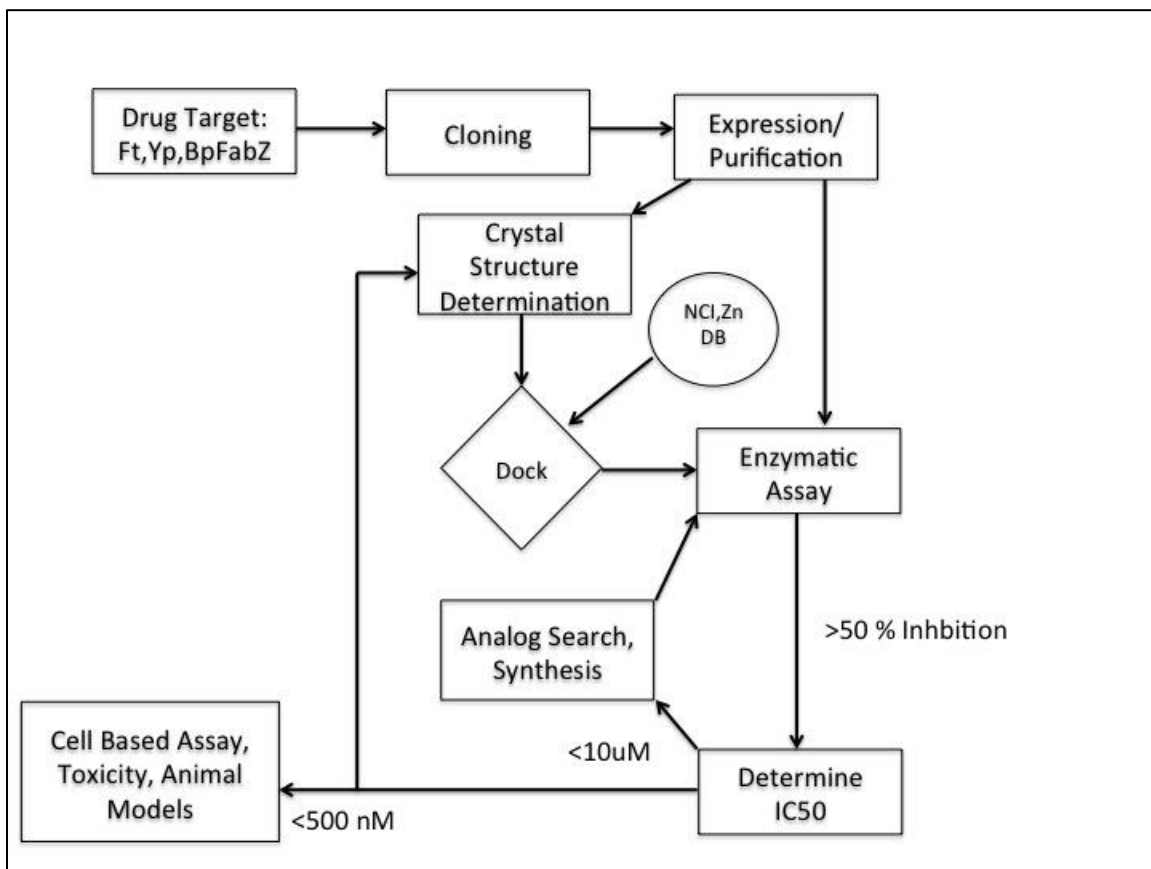


Figure 1.1: Iterative process of structure based drug design.

In this work, we focus on two specific bacterial protein targets for drug discovery efforts: The fatty acid biosynthesis pathway enzyme, β -hydroxyacyl-ACP dehydratase (*FabZ*), and the *Clostridium botulinum* neurotoxins (BoNTs).

Fatty Acid Biosynthesis Type II (FAS II) Pathway

The fatty acid biosynthesis pathway is responsible for creating the essential fatty acid precursors for the phospholipid bilayer of cell membranes. In bacteria, fatty acid biosynthesis is carried out by a series of individual enzymes, each encoded by a separate

gene and collectively known as Type II FAS. This contrasts with the Type I FAS system present in humans, which is comprised of a single multifunctional polypeptide containing all of the initiation and chain elongation enzymes of fatty acid biosynthesis.^{4,5} The actual chemical mechanisms underlying saturated fatty acid synthesis are nearly identical in both systems. However, the major differences between these two FAS systems in enzyme sequence, structure, and organization suggest that bacterial fatty acid biosynthesis can be successfully targeted without interfering with the host process.^{6,7} Furthermore, many of the specific enzymes of the Type II FAS system contain significant sequence and structural homology across a wide range of bacterial pathogens, making this pathway an excellent target for developing broad-spectrum antibiotics.⁸

In bacteria, the fatty acid biosynthesis pathway can be understood through a description of the process in *E. Coli*.⁹ Generally, the process consists of initiation followed by four sequential steps of chain elongation, each carried out by an individual enzyme, as summarized in Figure 1.2. Initiation begins with the heterotetrameric enzyme acetyl-CoA carboxylase (AccABCD) catalyzing the biotin dependent carboxylation of acetyl-CoA to malonyl-CoA.¹⁰ Next, the malonyl group of malonyl-CoA is transferred to the sulfhydryl of acyl carrier protein (ACP) by malonyl-CoA:ACP transacylase (FabD) forming malonyl-ACP.¹¹ The first cycle of chain elongation is initiated by the condensation of acetyl-CoA with malonyl-ACP by β -ketoacyl-ACP synthase III (FabH).¹² The second step of the elongation cycle is carried out by β -ketoacyl reductase (FabG) forming β -hydroxyacyl-ACP.¹³ Next, the β -hydroxyacyl-ACP intermediate is dehydrated by β -hydroxyacyl-ACP dehydratases, FabZ or FabA, to form trans-2-enoyl-ACP.¹⁴ Finally, the double bond of trans-2-enoyl is reduced by enoyl-ACP reductase

(FabI). The acyl-ACP product of the FabI reaction feeds back into the cycle and β -ketoacyl-ACP synthase II (FabF) initiates subsequent rounds of elongation. The elongation cycle repeats to increase the chain length until chain termination catalyzed by thioesterase cleavage of the acyl chain produces free fatty acid.⁶

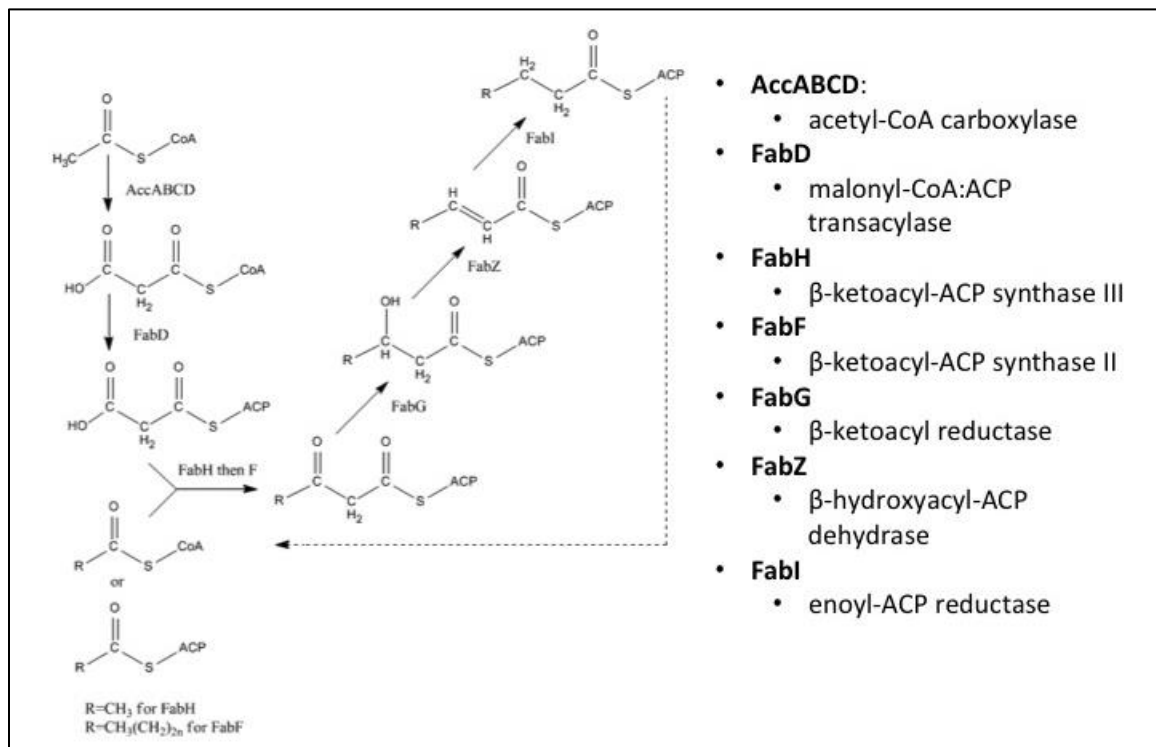
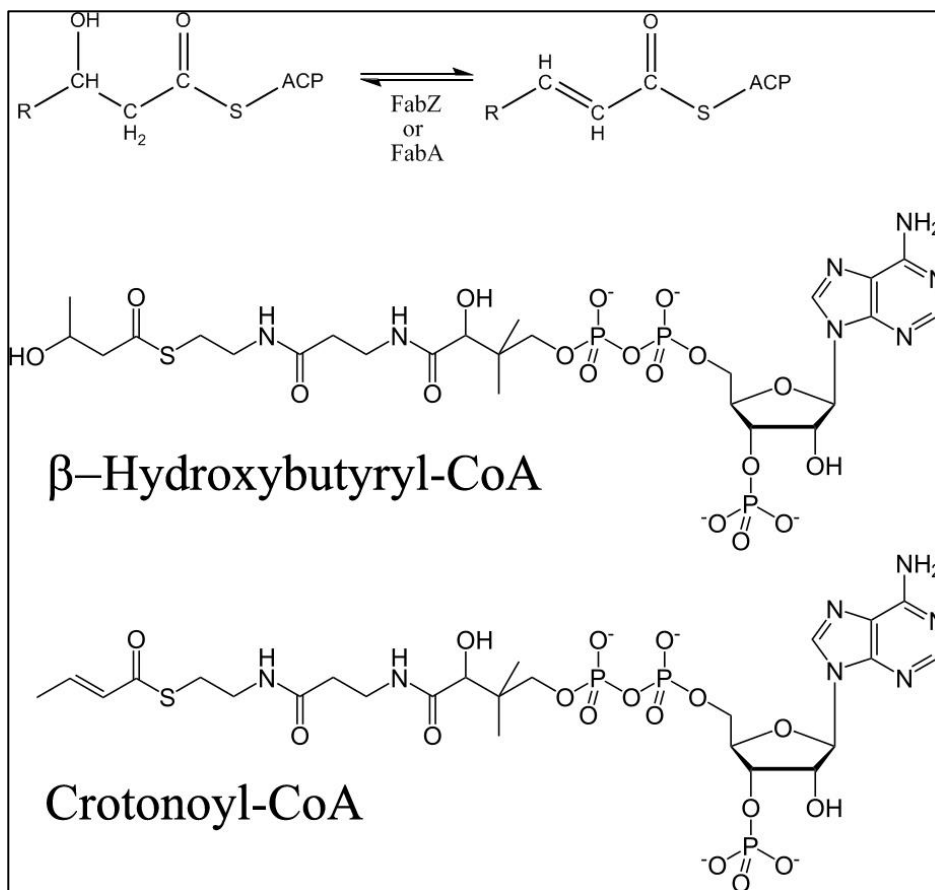


Figure 1.2: Fatty Acid Synthesis Pathway: Initially, acetyl-CoA is carboxylated by AccABCD forming malonyl-CoA. Next, malonyl-CoA is transferred to ACP by malonyl-CoA:ACP transacylase (FabD). The first cycle of chain elongation is initiated by the condensation of acetyl-CoA with malonyl-ACP by β -ketoacyl-ACP synthase III (FabH). The second step of the elongation cycle is carried out by β -ketoacyl reductase (FabG) forming β -hydroxyacyl-ACP. Next, the β -hydroxyacyl-ACP intermediate is dehydrated by FabZ to form trans-2-enoyl-ACP. Finally, the double bond of trans-2-enoyl is reduced by enoyl-ACP reductase (FabI). The acyl-ACP product of the FabI reaction feeds back into the cycle and FabF initiates subsequent rounds of elongation. Figure adapted from Campbell et. al.⁶

β -hydroxyacyl-ACP dehydratase (FabZ)

In the third step of the elongation cycle of fatty acid biosynthesis, β -hydroxyacyl-ACP is dehydrated by one of two β -hydroxyacyl-ACP dehydratases (FabA or FabZ) to form an α,β -unsaturated acyl-ACP (scheme 1.1). Additionally, these dehydratases have been shown to be active against the substrate analogs crotonoyl-CoA and β -hydroxybutyryl-CoA, however at a much higher K_m and lower V_{max} values compared to the natural ACP bound substrate. Unlike the situation in vivo, the reaction equilibrium with these analogs is towards the reverse reaction favoring the formation of β -hydroxybutyryl-CoA, with a ratio of rate constants (k_f/k_r) of approximately 1:7.¹⁵ The forward reaction is likely favored under physiological conditions due to the presence of FabI, which removes the FabZ product and shuttles it to the next enzyme in the FAS cycle, thus shifting the equilibrium towards the forward reaction. Supporting this are FabZ-FabI coupled assays, where a marked reduction in the formation of β -hydroxybutyryl-CoA, concurrent with an increase in butyryl-CoA, is observed.¹⁵

Although both FabA and FabZ are capable of catalyzing the dehydration step of the FAS cycle, they are not equally good drug targets. The primary role of FabA is in unsaturated fatty acid synthesis and is only found in Gram-negative bacteria, while FabZ is the primary dehydratase in the formation of both saturated and unsaturated fatty acids. This, coupled with the ubiquitous distribution of FabZ across both Gram-positive and Gram-negative pathogenic bacteria, makes this protein especially suited for antimicrobial drug discovery efforts.⁶



Scheme 1.1: β -hydroxyacyl-ACP is the natural substrate of FabZ, which is converted to trans-2-enoyl-ACP *in vivo*. Also shown are the substrate analogs β -hydroxybutyryl-CoA and crotonoyl-CoA.

The first FabZ structure to be solved was from *Plasmodium falciparum* (PfFabZ), which revealed an active site tunnel formed from the dimerization of two individual FabZ monomers (Fig 1.3).¹⁶ Within the tunnel reside the catalytic histidine and glutamic acid residues, which are completely conserved across FabZs from different organisms. Additionally, a conserved glycine, located near the catalytic residues, is believed to play a role in FabZ enzymatic activity.

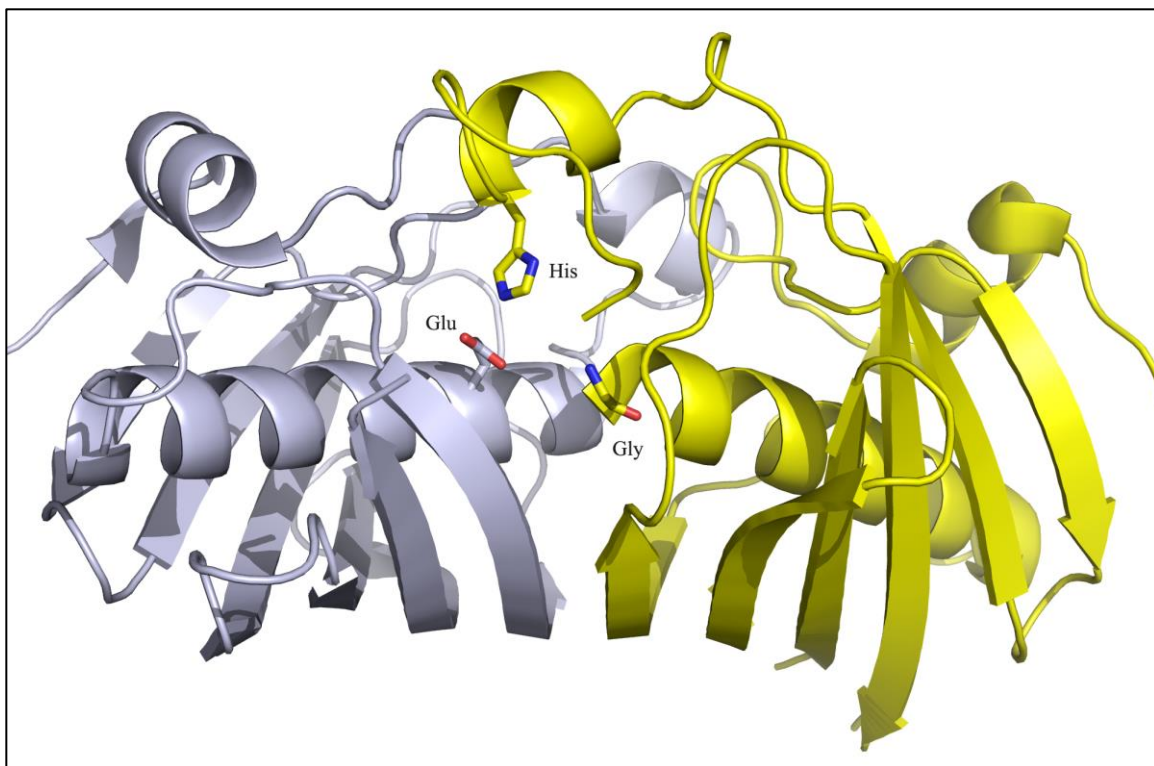
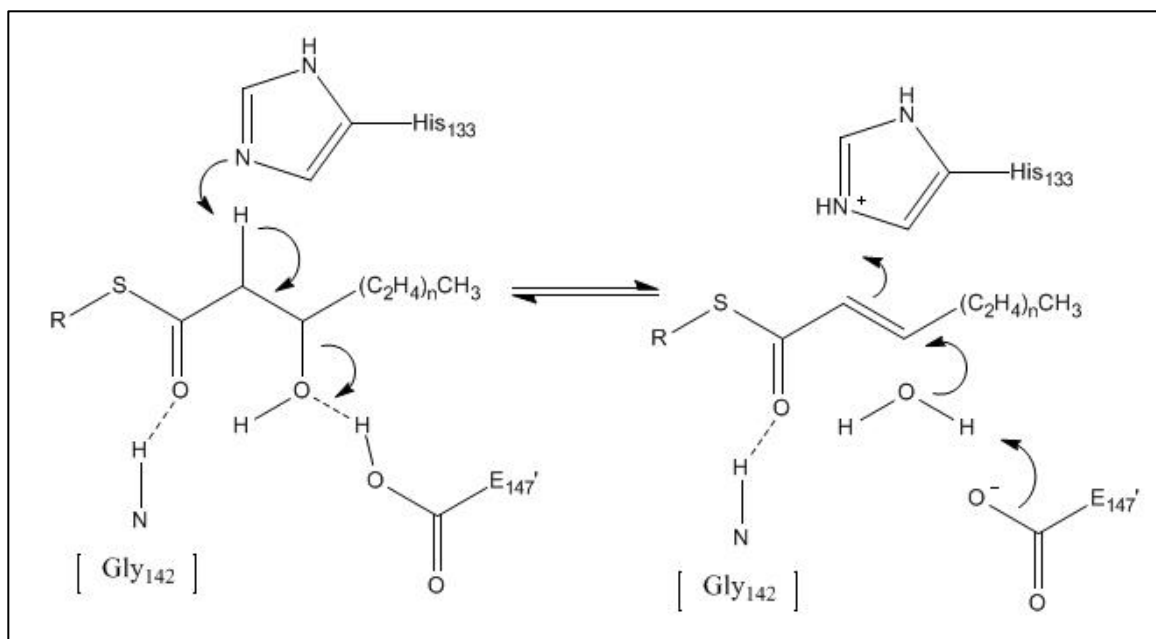


Figure 1.3: Ribbon diagram of the FabZ dimer. Individual monomers are colored silver and yellow, respectively. Residues that participate in the enzymatic mechanism are shown in stick representation.

The proposed enzymatic mechanism of FabZ for *E. coli* is shown in Scheme 1.2. In this mechanism, a hydrogen bond forms between the thioester carbonyl oxygen of the substrate, and the backbone nitrogen of a conserved glycine located in the active site. This serves to positively polarize the carbonyl bond. Additionally the free electron pair of the N_{ϵ_2} atom of the catalytic histidine positively polarizes the α -C-H bond. Similarly, the β -C-OH bond is positively polarized by the hydrogen bond formed between the β -hydroxyl group and the carboxylic acid group of the catalytic glutamic acid.^{14, 17}



Scheme 1.2: The proposed FabZ enzymatic mechanism. Figure adapted from Kostrewa et al.¹⁷

Known Inhibitors of β -hydroxyacyl-ACP dehydratase (FabZ)

The first FabZ inhibitors discovered were synthetic compounds, named NAS91 and NAS21, targeted against the *P. falciparum* variant of the enzyme (PfFabZ).^{15, 18} These inhibitors were identified from a large screen of synthetic compounds, which were designed to have a wide range of functional groups, while also taking into account their molecular geometry to ensure their suitability as a potential target of FabZ. NAS91 and NAS21 were found to have K_i values of 1.31 μM and 1.46 μM and were shown to inhibit Plasmodium growth in vitro. Also, NAS91 contains a 4-chlorophenol group similar to the well-characterized antibacterial, triclosan, which inhibits FabI.¹⁹ However, NAS91 did not show any activity against PfFabI and similarly triclosan does not appear to inhibit FabZ.

The remaining known inhibitors were identified against FabZ from *Helicobacter pylori* (*HpFabZ*). These include two synthetic compounds, originally designated Compound 1 and Compound 2²⁰, as well as the natural compounds Quercetin²¹, Apigenin²¹, Emodin²², and Juglone.²³ These compounds have K_i values ranging from 1.9 μM to 14.9 μM and all have been successfully crystallized with *HpFabZ*. With the exception of Compound 2, all of these inhibitors demonstrated two different binding poses in the crystal structure. One pose is located at the entrance to the active site tunnel and is referred to as a “gate blocking” pose (Fig 1.4a). The other pose is located deep in the active site tunnel generally locating near the active site residues. This pose is referred to as “channel blocking” (Fig. 1.4b). The chemical structures and inhibition data of all known FabZ inhibitors are summarized in Figure 1.5.

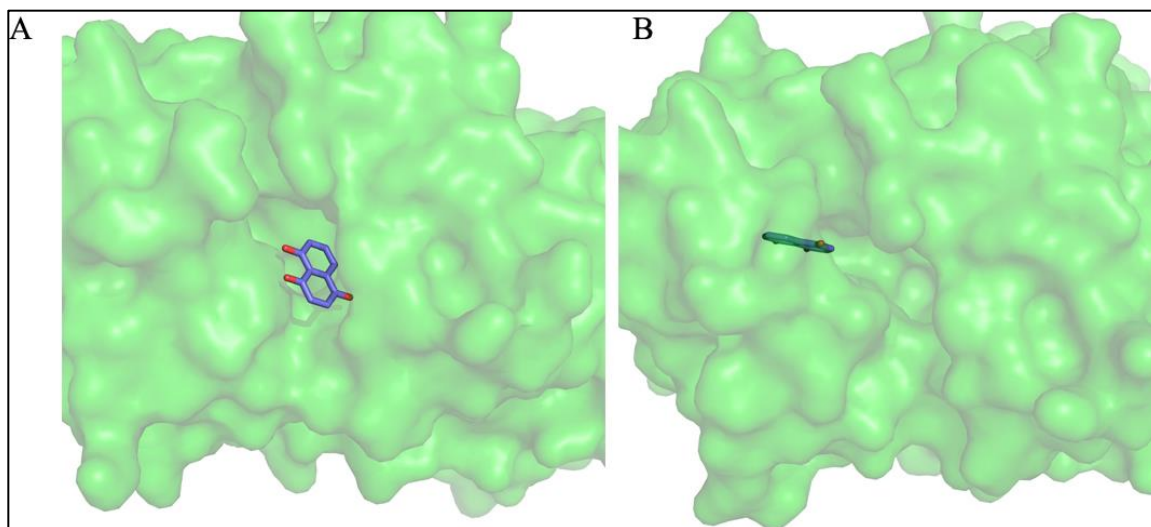


Figure 1.4: Example dual pose for the *HpFabZ* inhibitor, Juglone. (A) Gate blocking pose where Juglone locates at the entrance to the active site tunnel. (B) Channel blocking pose where Juglone occupies a position deeper in the active site tunnel near the catalytic residues.

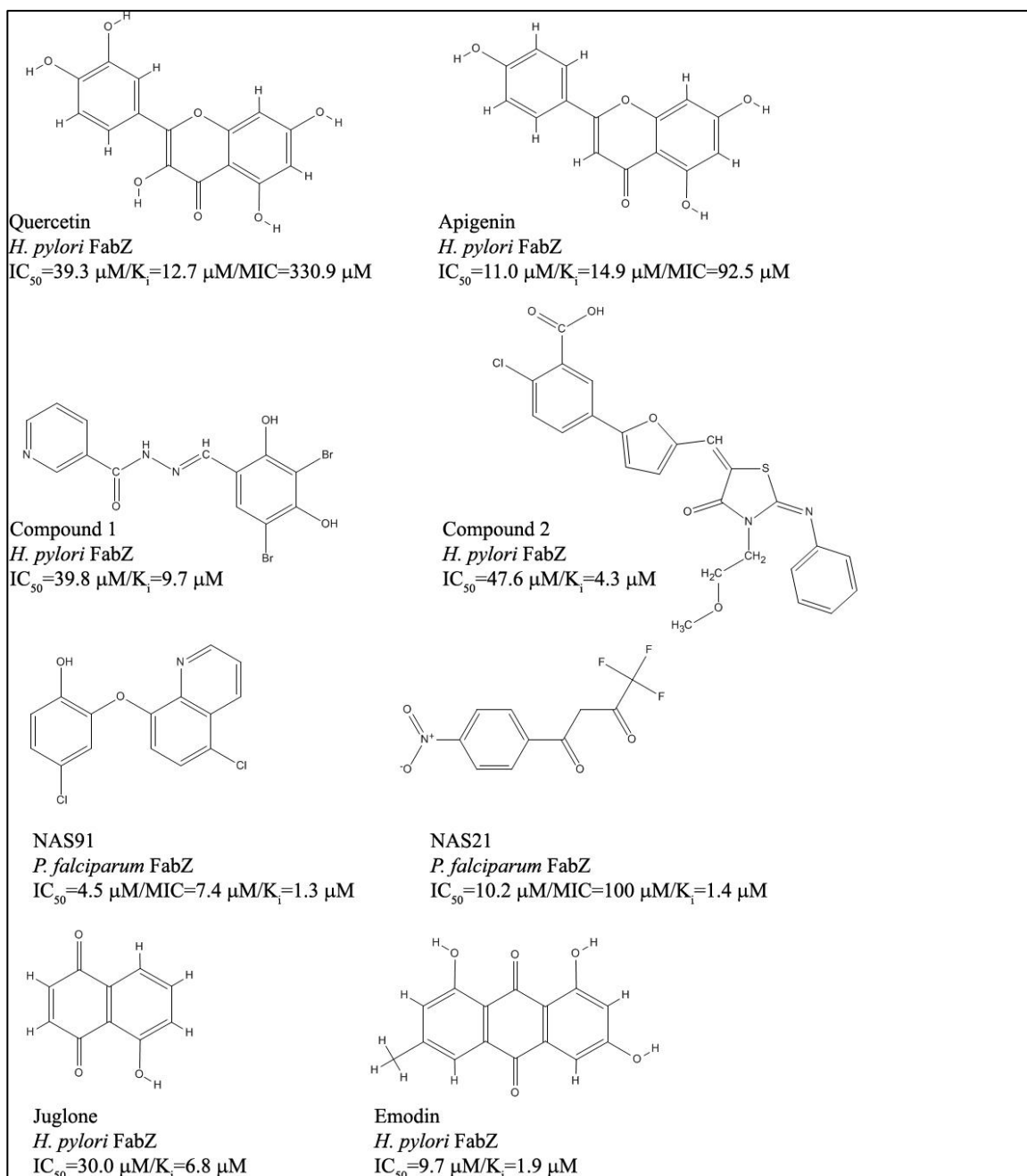


Figure 1.5: The chemical structures and inhibition data of all known FabZ inhibitors.

Clostridium Botulinum Neurotoxins

Botulinum neurotoxins (BoNT) are among the most potent toxins that exist in nature and are found in soil and untreated water throughout the world.²⁴ There are several variants of disease caused by the BoNTs, depending on the mode of exposure. The most common illness caused by BoNTs is food borne botulism resulting from the ingestion of preformed toxin found in spoiled food. A more unusual form of botulism is inhalation botulism, caused by exposure to aerosolized toxin. This form of botulism has only been observed in lab workers; however, it illustrates the disturbing possibility that Botulinum toxins could be easily weaponized and released into the environment causing widespread illness in humans and livestock. Due to its potentially lethal effects, the development of a safe and effective inhibitor against the BoNTs is imperative. Unexpectedly, Botulinum neurotoxin has prospective therapeutic applications as well. Currently, these toxins are used for the treatment of several illnesses including strabismus and blepharospasm, in addition to its more commonly known use in cosmetic applications.²⁵

The Botulinum neurotoxins are produced as a single inactive chain with a molecular mass of 150kDa. This single chain is later cleaved by bacterial and host tissue proteases into a 100 kDa heavy chain (HC) and a 50kDa light chain (LC). The HC is responsible for host cell binding, internalization, and translocation of the toxin, while the LC represents the enzymatic portion of the BoNT, which acts to cleave specific bonds in proteins of the neuronal SNARE complex (Soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptors).²⁴ The SNARE complex itself is composed of three separate proteins: VAMP, Syntaxin, and SNAP-25.²⁶ As a unit it mediates vesicle-membrane fusion at synaptic terminals and subsequent neurotransmitter release into the

synaptic cleft.²⁷ There are seven antigenically distinct forms of BoNT-LCs (serotypes BoNT/A-BoNT/G), all sharing significant sequence and structural homology; however there is wide variation in protein and cleavage site specificity between these different subtypes. Specifically, SNAP-25 is the substrate for serotypes A, C, and E, while VAMP is the substrate for serotypes B, D, F, and G. Serotype C is unique in that it has dual substrate specificity and is capable of cleaving Syntaxin in addition to SNAP-25. Each serotype cleaves a single distinct peptide bond within its respective substrate, with no overlap between any of the serotypes (Table 1.1).²⁸

Table 1.1: Target molecule and cleavage site of all seven BoNT serotypes.

BoNT Serotype	Target	Cleavage Site
BoNT/A	SNAP-25	Gln197-Arg198
BoNT/B	VAMP	Gln76-Phe77
BoNT/C	SNAP-25, Syntaxin	Arg198-Ala199, Lys253-Ala254
BoNT/D	VAMP	Lys59-Leu60
BoNT/E	SNAP-25	Arg180-Ile181
BoNT/F	VAMP	Gln58-Lys59
BoNT/G	VAMP	Ala81-Ala82

Current treatments for systemic botulism are limited, consisting mainly of antisera, which if given prior to complete paralysis, can prevent worsening symptoms and shorten overall recovery time. Other than antisera, treatments for botulism are primarily supportive and may include placing a patient on a ventilator if respiratory failure occurs. Other treatments depend on the route of exposure and may include induced vomiting for food born exposure, or surgical care to remove bacteria from an infected wound.²⁹ Clearly, the treatment options for botulism are limited, highlighting the need for new effective drug therapies.

The Structure and Catalytic Mechanism of Botulinum Neurotoxin Light Chain

The structures for all seven BoNT-LC serotypes have been solved over the last several years. These structures mainly consist of a mix of α -helix and β -strand secondary structure (Fig. 1.6a). The BoNT-LC has a single active site buried deep within the enzyme, located approximately 20 Å from the surface. This active site has an overall negative surface charge and is accessible via a large channel that is about 12 Å × 15 Å × 35 Å wide (Fig. 1.6b).³⁰ The primary structural differences between the various serotypes of BoNT-LC are at flexible solvent exposed loops, but the overall structure of each enzyme and the arrangement of their active sites are very similar.

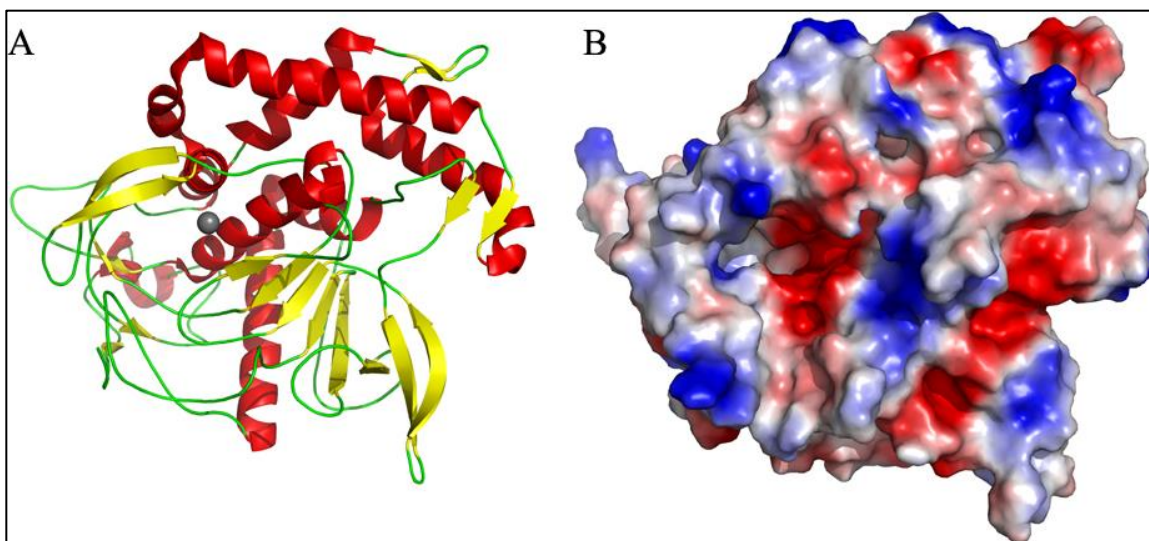
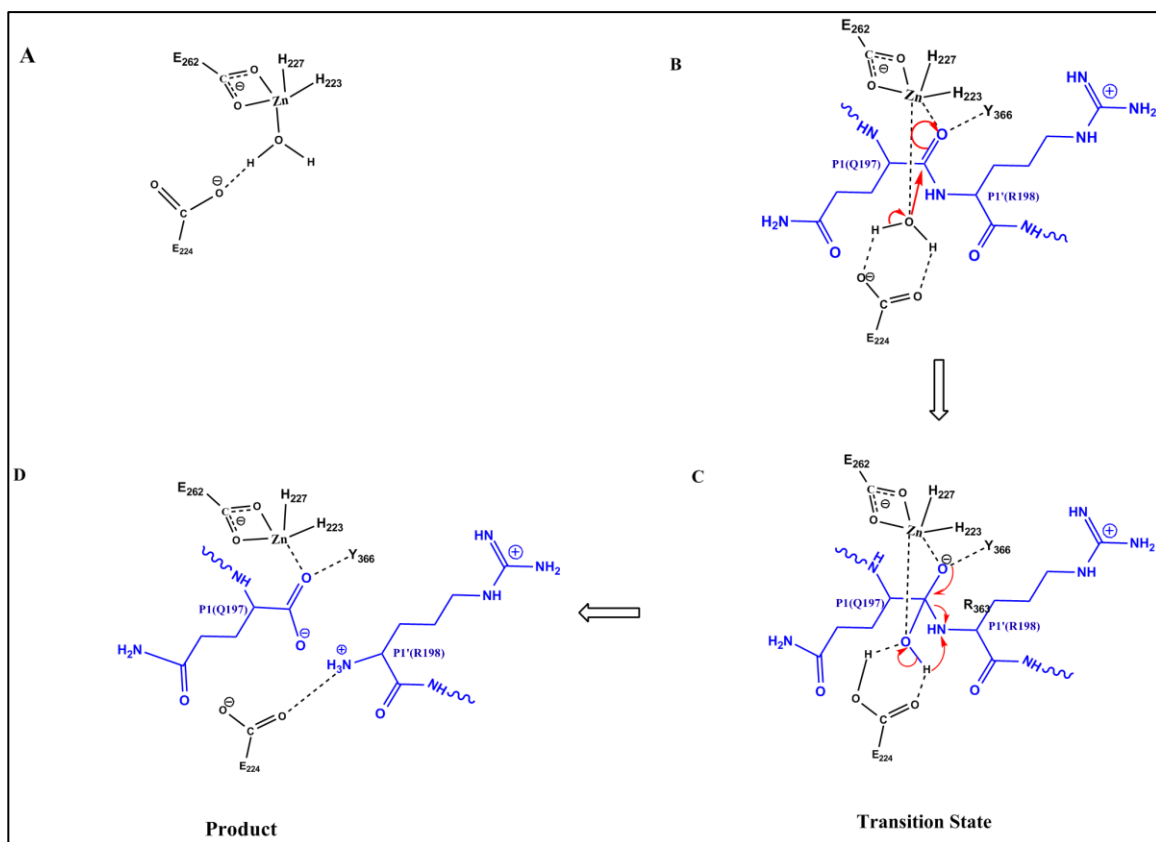


Figure 1.6: Structure of BoNT/A-LC. (A) Ribbon diagram of BoNT/A-LC. α -helices are colored red, β -strands are colored yellow, and loop regions are colored green. The active site zinc metal ion is shown as a sphere and colored grey. (B) Electrostatic surface view of BoNT/A, with molecule oriented to display the active site. Electronegative regions are shown in red. Electropositive regions are shown in blue. PDB ID: 2QNO.

The BoNT active site contains a zinc metal ion and a water molecule, both of which are crucial to its proteolytic activity. The zinc ion is coordinated by a conserved HExxH zinc-binding motif, which is common to all the BoNT serotypes. The active site

of native BoNT/A is shown in Scheme 1.3a, where His227, His223, Glu262, as well as the active site water molecule, coordinate the zinc ion.³¹



Scheme 1.3: Enzymatic mechanism of BoNT proteolysis. Residues of the BoNT enzyme are colored in black, while residues of the SNAP-25 substrate are colored in blue. (A) Native active site. (B) Substrate bound state. (C) Transition state. (D) Product Formation. Adapted from Kumaran et al.³¹

The catalytic mechanism of the BoNTs can be understood through a description of the process in BoNT/A, which has been more extensively characterized than the other serotypes (Scheme 1.3). The process begins with Glu224, which, acting as a general base, removes a proton from the active site nucleophilic water. This event primes the water molecule for attack on the carbonyl carbon of the scissile bond, resulting in the formation of a tetrahedral transition state. It is believed that the zinc ion and Tyr366 stabilize this

transition state. The process completes through a shuttling of electrons, aided by Glu224 forming a stable leaving amino group.³¹

BoNT light chain interactions with substrate

The most complete structural data available for BoNT-LC with its natural substrate is of an inactive double mutant of BoNT/A in complex with a truncated form of SNAP-25 consisting of residues 141-204 (Fig. 1.7).³² This BoNT/A-SNAP-25 structure revealed an extensive network of interactions extending far from the catalytic active site, with SNAP-25 wrapping around most of the surface of BoNT/A. This observation led to the hypothesis that substrate recognition begins on the portion of SNAP-25 distal to the active site within an area on BoNT/A termed the α -exosite. This initial interaction is believed to induce successive contacts between BoNT/A and SNAP-25 along a path towards the active site. This has the effect of anchoring the SNAP-25 substrate so that it is spatially oriented in the correct position for enzymatic cleavage. These observations help to explain how the individual BoNT serotypes maintain such high specificity, while simultaneously possessing very similar active site architectures.

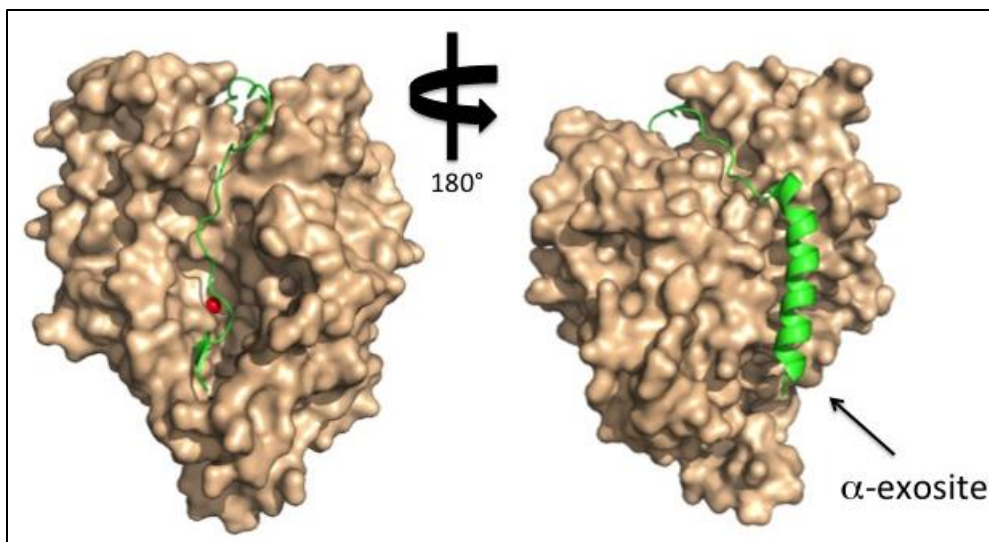


Figure 1.7: The structure of BoNT/A bound to SNAP-25 residues 141-204. BoNT/A is shown in surface representation. SNAP-25 is shown in green cartoon representation. The active site zinc ion is shown as a red sphere. The position of the BoNT/A exosite is also labeled. PDB ID: 1XTG

Known Inhibitors of the Botulinum Neurotoxin Light Chain

BoNT/A Inhibitors. In the case of BoNT/A, it was discovered that the minimum length segment of SNAP-25 that remains cleavable by the enzyme is DEANQ/RATK, with the scissile bond located between Q197 and R198.²² This observation led to exhaustive searches for peptide and peptidomimetic compounds of shorter sequence length, which would presumably be uncleavable and compete with the full-length substrate. A number of such inhibitors have been discovered, many of which have been successfully co-crystallized with BoNT/A (Table 1.2). The earliest of these inhibitors was of the sequence CRATKML with a K_i of 2.0 μM .³³ Subsequent attempts at optimizing this compound revealed that the arginine at the P1' position was critical to potency and could not be altered. However, replacement of the N-terminal cysteine to various sulfhydryl-containing compounds was successful, with 2-mercapto-3-phenylpropionyl proving to be the most effective substitution, lowering the K_i to 330 nM.^{34, 35} This finding demonstrated

that improvements to peptide-based inhibitors were possible and many of the subsequently discovered BoNT/A inhibitors flowed from this initial work. Building on this, Zuniga et al. performed an exhaustive optimization of the native QRATKML sequence, synthesizing multiple peptidomimetic inhibitors of BoNT/A, the most potent of which was *DNP-DAB-RWT-DAB-ML* (*DNP-DAB* is 4-(2,4-dinitrophenylamino)-2-amino-butanoic acid and *DAB* is 2,4-diaminobutanoic acid) with a K_i of 41 nM.³⁶ Additionally, tetrapeptide inhibitors corresponding to the P1-P1'-P2'-P3' positions of SNAP-25 have also been tested with various success, the strongest being RRGF with an IC_{50} of 900 nM.³⁷ Peptides shorter than four amino acids have demonstrated a significant reduction in potency and have not resulted in successful inhibitors.^{37, 38}

Due to certain drawbacks, such as their large size, it is often difficult to transform peptide-based inhibitors into usable drugs. For this reason, much work has been done to also identify small molecule inhibitors of BoNT/A (Fig. 1.8). Several compounds have been discovered over the years, most of which have IC_{50} and/or K_i values in the low micromolar range. The most potent of these small molecule inhibitors was a 4-amino-7-chloroquinoline based inhibitor developed by Videnovic *et al.*, with an IC_{50} of 100 nM.³⁹ This compound also demonstrated strong protection of SNAP-25 in cell based assays.

BoNT/B Inhibitors. To date, there are three known inhibitors of BoNT/B (Table 1.2 and Fig. 1.9). The first compound discovered was a small molecule, 7-*N*-phenylcarbamoylamino-4-chloro-3-propyloxyisocoumarin, with an IC_{50} of 27.6 μ M.⁴⁰ A far more potent BoNT/B inhibitor was later synthesized based on the rational design of several pseudotriptides containing β -amino thiols. This inhibitor was designed by Anne et al. and has a K_i of 20 nM.^{41, 42} Lastly, the natural peptide, Buforin, has been shown to

inhibit BoNT/B with an IC_{50} of 1.0 μM .⁴³ None of these three inhibitors have been successfully co-crystallized with BoNT/B.

BoNT/E Inhibitors. There are two known inhibitors of BoNT/E; one is a small molecule inhibitor and the other is the substrate based peptide inhibitor, RIME (Fig. 1.9). The peptide RIME represents the SNAP-25 sequence from residue 180 to 183, spanning the scissile bond. The structure of RIME in complex with BoNT/E has been solved and this structure has become useful as a receptor for virtual screening of small molecule inhibitors of BoNT/E.⁴⁴ One such screen was performed on 1990 molecules of the NCI diversity set using Autodock. This screen proved fruitful, yielding several compounds showing activity at a concentration of 250 μM . The most potent of these inhibitors was 2-(9H-fluoren-2-ylcarbonyl)benzoic acid) with a K_i of 1.29 μM .⁴⁵

BoNT/F Inhibitors. Several substrate-based inhibitors of BoNT/F exist, but conceptually they are of similar design (Table 1.2). Work done previously on substrate based inhibitors of BoNT/A had shown that replacing the P1 residue with a zinc binding sulfhydryl group could dramatically increase its inhibitory potency. A similar strategy was used with these BoNT/F inhibitors which all contain a D-cysteine substitution at VAMP position Glu58, the P1 position for BoNT/F. Multiple permutations of the substrate length were then tested and it was determined that VAMP₁₇₋₅₈ (Q58/D-C) and VAMP₂₂₋₅₈ (Q58/D-C), were the most potent inhibitors, both with K_i values of approximately 1.0 nM.^{46, 47}

Table 1.2: Peptide and Peptidomimetic inhibitors of BoNT/A, BoNT/B, and BoNT/F.
Table adapted from Kumar et al.⁴⁸

Peptide/Peptidomimetic BoNT/A	Inhibition (μM)		PDB ID	Reference
	IC ₅₀	K _i		
RRGC	1.5	0.157	3C88	31
RRGM		0.845	3C89	31
RRGL		0.660	3C8A	31
RRGI		0.786	3C8B	31
QRATKM	133.0		3DDA	49
RRATKM	95.0		3DDB	49
RRGF	0.9		3QW5	37
CRGF	1.5		3QW8	37
RRFC	1.8		3QW7	37
RRYC	5.4			37
CRGC	8.0			37
WRGC	10.0			37
QRGC	14.0			37
CRRGC	43.0			37
RRGCM	26.1			37
RRKRL	28.7			37
N-Ac-CRATKML		2.0	3BOO	34, 35
I1 (<i>DNP-DAB-RWT-DAB-ML</i>)		0.041	3DS9	50
I2		6.5		50
I6		8.3		50
I7		3.3		50
I8		0.98		50
I9		0.094		50
I10		0.05		50
I11		0.32		50
I12		0.1		50
I13		0.39		50
Compd: JTH-NB72-35		0.314		36
Compd: JTH-NB72-38		0.990		36
Compd: JTH-NB72-39		0.638	3NF3	36
Peptide/Peptidomimetic BoNT/B				
Buforin I: AGRGKQGGKVRAKAKTR SSRAGLQFPVGRVHLLR KGNK	1.0			43
Peptide/Peptidomimetic BoNT/F				
VAMP ₄₂₋₅₈ (Q58/D-C)		9.0		46
VAMP ₃₇₋₅₈ (Q58/D-C)		0.28		46
VAMP ₃₂₋₅₈ (Q58/D-C)		0.034		46
VAMP ₂₇₋₅₈ (Q58/D-C)		0.0019	3FII	46
VAMP ₂₂₋₅₈ (Q58/D-C)		0.0010	3FIE	46
VAMP ₁₇₋₅₈ (Q58/D-C)		0.0013		46

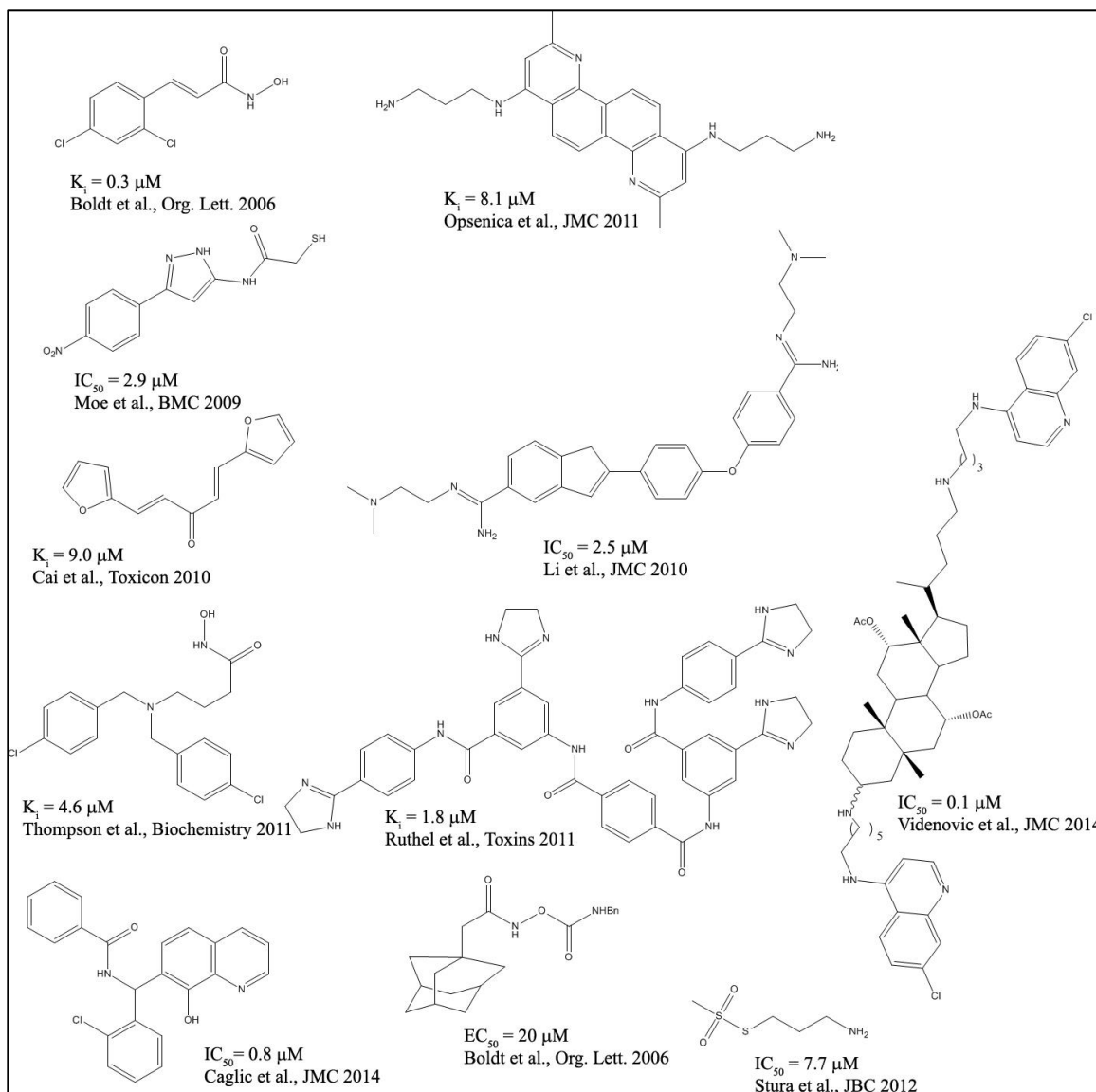


Figure 1.8: Chemical structures and inhibition data of all known small molecule BoNT/A-LC inhibitors.^{39, 51-59} Figure adapted from Kumar et al.⁴⁸

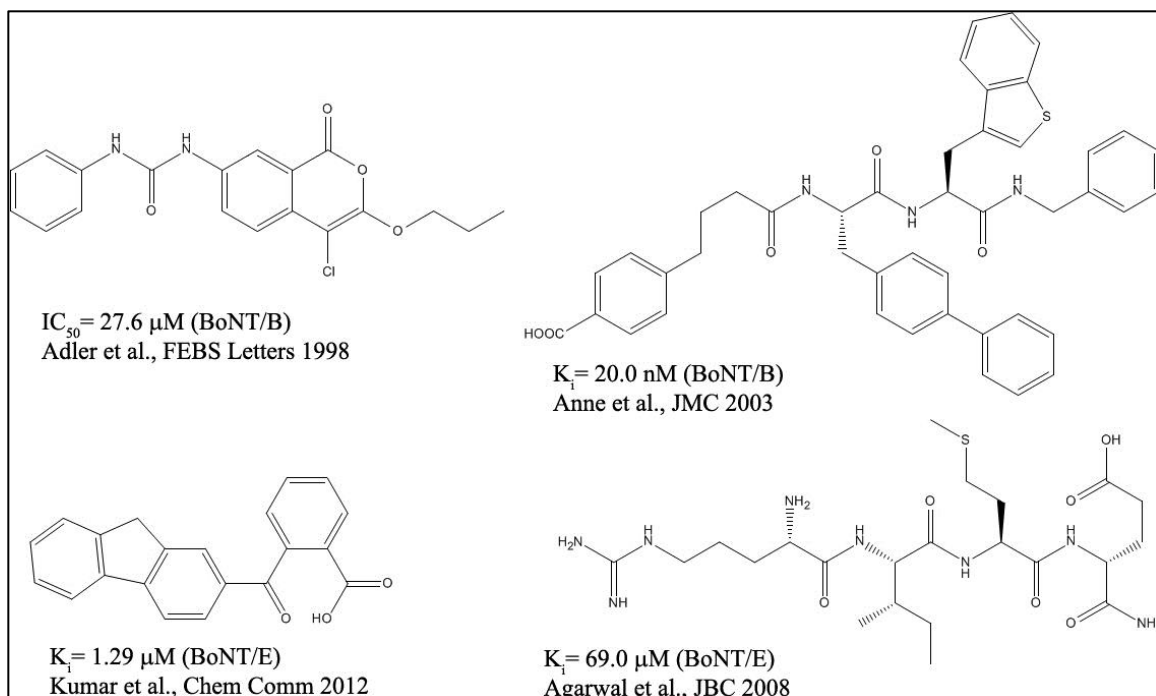


Figure 1.9: Structures and inhibition data for BoNT/B and BoNT/E inhibitors.
 Figure adapted from Kumar et al.⁴⁸

Current assays of BoNT light chain endopeptidase activity

The earliest assays developed for BoNT/A and BoNT/B involved immobilizing the C-terminus of segments of either VAMP (residues 60-94) or SNAP-25 (residues 137-206) on a microplate. The assay would proceed by adding buffer containing BoNT to the plate wells resulting in proteolysis, which would leave the C-terminal fragment of the substrate still bound to the plate well. An enzyme-linked antibody specific to the newly exposed N-terminus of the substrate would then be used to detect cleavage.^{60, 61} In effect, these assays are basically a form of enzyme-linked immunosorbent assay (ELISA) with many steps, and are therefore not very conducive to high throughput screening of inhibitors.

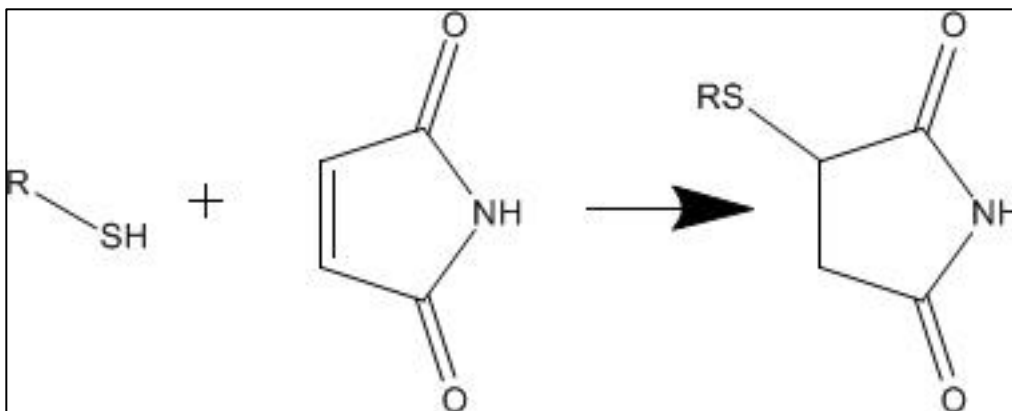
Another common way of assaying BoNT activity are HPLC based methods. In these assays, the substrate is a fragment of either VAMP or SNAP-25 spanning the

cleavage site of the BoNT serotype to be tested.⁶² For example, in the case of BoNT/A and BoNT/C, the 17-mer peptide ¹⁸⁷SNKTRIDEANQRATKML²⁰³ is used.⁶³ The substrate is incubated with BoNT in a suitable buffer and the reaction is allowed to proceed, taking care not to exceed the V_i of the reaction. Once the reaction is stopped by acidification of the reaction mixture, the cleavage fragments are analyzed via HPLC. Initial reaction velocities can then be obtained, typically by quantifying the amount of cleaved product formed by integration under the chromatogram peak for each species. However, like the ELISA based assay method, these HPLC assays are difficult to implement for high throughput screening.

Currently, the most heavily utilized BoNT assays are fluorescence-based, the most common of which are the commercial assays SNAPtideTM and BoTestTM.⁶⁴⁻⁶⁶ SNAPtide was the first widely available assay and is composed of the sequence (FITC)-TRIDEANQRATK(DABCYL)M. SNAPtide is a fluorescence resonance energy transfer (FRET) based substrate where the N-terminal fluorescein isothiocyanate (FITC) is quenched by 4-((4-(dimethylamino)phenyl)azo) benzoic acid (DABCYL), positioned on the C-terminal side of the peptide. After digestion by BoNT, the quench is released and a fluorescence signal can be measured. SNAPtide is specific to BoNT/A only, but similar substrates spanning the cleavage site of BoNT/B, BoNT/C, and BoNT/E are also available. These substrates can be used to perform continuous assays and are also easily scaled to run multiple reactions at once using a microplate reader. The primary downside of the SNAPtide substrates is their high cost and the requirement of a specific substrate for each BoNT serotype to be tested.

The BoTest assay represents several significant improvements over SNAPtide. Chiefly, BoTest does not require an individual substrate for each BoNT serotype and instead allows six of the seven serotypes to be assayed with only two separate substrates. These substrates are BoTest A/E (sequence: CFP-SNAP25₁₄₁₋₂₀₆-YFP) and BoTest B/D/F/G (sequence: CFP-Synaptobrevin2₃₃₋₉₄-YFP). The CFP and YFP domains of these substrates form a FRET pair. In the intact substrate, CFP excitation leads to YFP excitation resulting in quenched fluorescence of CFP and a corresponding increased fluorescence of YFP. Alternatively, when treated with BoNT, linkage between CFP and YFP is lost upon proteolysis and the opposite fluorescence pattern is seen, namely increased CFP fluorescence and decreased YFP fluorescence. Therefore, measuring the ratio of YFP to CFP emission allows for real time assaying of BoNT activity.

Lastly, Hines et al. have developed a recombinant substrate (called GFPSV) containing a SNAP-25/VAMP fusion protein linked to an N-Terminal GFP and a C-terminal cysteine residue. The cysteine was designed to allow for GFPSV to be immobilized on maleimide activated 96-well plates (Scheme 1.4), where it could subsequently be cleaved by the addition of a buffer containing BoNT-LC. Proteolysis results in movement of the N-terminal GFP linked segment of GFPSV into solution allowing for a fluorescence signal to be obtained. This substrate is both cost effective and allows for assaying all BoNT serotypes with a single substrate. However, GFPSV has proven to have extremely poor solubility and is prone to aggregation, making reliable purification difficult. Furthermore, the GFP signal obtained from this substrate is weak, resulting in poor run-to-run reproducibility and irregularly shaped progress curves.



Scheme 1.4: The maleimide and sulfhydryl groups react under pH conditions between 6.5 and 7.5, resulting in the formation of a stable thioether linkage.

Research Project Overview

The overall focus of this research project is the discovery of potent inhibitors of the bacterial protein targets FabZ and BoNT-LC. In chapter 2 we present the crystal structure of FabZ from two different organisms, *Yersinia pestis* (YpFabZ) and *Francisella tularensis* (FtFabZ). We next performed an enzymatic characterization of both FabZ isoforms, to include mutational studies of key residues involved in substrate recognition and catalysis. Additionally, we conducted an experimental screen of candidate FabZ inhibitors selected based on their complementarity to the FabZ active site. This screen identified two novel FabZ inhibitors, Mangostin and Stictic Acid, which demonstrate inhibitory activity against both YpFabZ and FtFabZ. Finally, we extended our inhibition studies to include compounds previously identified as inhibitors of HpFabZ, to determine which common scaffolds tend to be broadly cross-reactive against multiple FabZ isoforms.

In chapter 3, we shift our attention to the Botulinum neurotoxin light chains and present a novel method for performing high throughput assays of their activity. This assay utilizes a recombinant GFP based substrate called Repcon, containing the cleavage

recognition site for all seven BoNT serotypes (BoNT/A-G). We demonstrate that our substrate is cleaved at BoNT serotype specific locations, producing a strong GFP signal allowing for the calculation of initial velocity data for each serotype. We further characterize the reproducibility of our assay method, showing less than 2% run-to-run deviation. Lastly, we assayed known inhibitors of several BoNT serotypes using our Repcon substrate. The results showed serotype specific inhibition with IC₅₀ values generally in agreement with those obtained through established BoNT assay methods.

In chapter 4, we describe a virtual screen of 1.4 million compounds against BoNT/E light chain. From this initial screen, we selected 92 lead compounds based both on their DOCK energy score and visual analysis of their docked poses. These 92 compounds were initially assayed using the commercially available SNAP-Etide substrate resulting in nine compounds showing inhibition against BoNT/E at a concentration of 100 μ M. These nine compounds were further tested using our Repcon assay (described in chapter 3) to verify their activities against BoNT/E. After testing with Repcon, a single compound, C562-1101, maintained its activity and IC₅₀ calculations demonstrated that this compound is more potent than both currently published BoNT/E inhibitors.

Chapter 2: β -Hydroxyacyl-Acyl Carrier Protein Dehydratase (FabZ) from *Francisella tularensis* and *Yersinia pestis*: Structure Determination, Enzymatic Characterization, and Cross Inhibition Studies

Introduction

F. tularensis and *Y. pestis* are both pathogenic Gram-negative bacteria, classified as Tier 1 select agents by the U.S. government. This designation has been assigned due to their low infectious dose, high virulence, and relative ease by which these pathogens can be aerosolized. *F. tularensis* causes the disease Tularemia, which can have various clinical manifestations depending on the route of exposure to include ulceroglandular, oropharyngeal, respiratory, oculoglandular, and typhoidal. Of these potential manifestations, respiratory Tuleremia is the most severe, marked by chills, high fever, difficulty breathing, dry or productive cough, pharyngitis, chest pain, headache, drenching sweats, drowsiness and weakness. The possible development of pneumonia is extremely deadly with an average mortality rate of 30-60%, if left untreated.⁶⁷

Y. pestis is most commonly known as the causative agent of Bubonic Plague, but when encountered in aerosolized form it can cause a deadly pneumonic plague. This variant, in contrast to Bubonic plague, can be spread from person to person making the disease much more difficult to control.⁶⁸ Of further concern is that multidrug resistant forms of *Y. pestis* have been isolated in the last few decades while simultaneously the number of human cases of plague has been rising.⁶⁹ This highlights the general need for novel antimicrobials to combat the growing concern not only of *Y. pestis*, but also of multidrug resistant bacteria in general.

Several FabZ structures from a variety of different organisms have been solved in recent years and include *Helicobacter pylori* (HpFabZ)²⁰, *Pseudomonas aeruginosa*

(PaFabZ)⁷⁰, *Campylobacter jejuni* (CjFabZ)⁷¹, *Plasmodium falciparum* (PfFabZ)¹⁶, *Burkholderia thailandensis* (BtFabZ)⁷², and *Neisseria meningitides* (NmFabZ) (PDB ID: 4I83). Additionally, several co-crystal structures of inhibitors bound to either HpFabZ or PfFabZ have been reported.^{15, 18, 20-23, 73} Most inhibitors reported thus far for HpFabZ adopt either of two general binding poses referred to as “gate-blocking” or “channel blocking”. Channel blocking inhibitors have been found to bind at or near the FabZ enzymatic active site, making specific interactions near the strictly conserved catalytic histidine and glutamic acid residues found at this location. Alternatively, gate-blocking inhibitors bind far from the active site, interacting with conserved residues surrounding the entrance to the tunnel leading to the FabZ active site.²⁰ Inhibitors for PfFabZ generally adopt only a single pose close to the active site and make contacts with the catalytic residues.¹⁸

In this work, we report the crystal structures of FabZ from *F. tularensis* (FtFabZ) and *Y. pestis* (YpFabZ) as well as their enzymatic properties. In addition, we have identified two novel inhibitors of FabZ: the natural compounds Stictic Acid and Mangostin. Lastly, we tested our newly discovered inhibitors, as well as several inhibitors previously identified for HpFabZ, against both FtFabZ and YpFabZ. As will be described, these two FabZ homologs together represent much of the natural variation that occurs between all FabZ molecules, particularly in key areas within the active site tunnel that have been shown to interact with inhibitors. Therefore, an understanding of the relative strength of known inhibitors against these two homologs should be informative for the future design of broad-spectrum inhibitors of FabZ.

Methods

Cloning, Expression and Purification. Plasmids for both FtFabZ, and YpFabZ were obtained from the DNASU Plasmid repository (DNASU Plasmid ID's: FtCD00023067 and YpCD00016324). The FabZ gene for each of these constructs was sub-cloned into a pET28b vector, yielding the plasmids pET28b-FtFabZ and pET28b-YpFabZ, both of which contain an N-terminal 6-His tag. Several YpFabZ point mutations were also constructed using pET28b-YpFabZ as a template. These mutations were inserted within 5' phosphorylated primers that were used to prime extensions around the entire plasmid. After amplification the plasmid was reconstituted using T4 ligase. The primers used were as follows: H19A (Forward: 5'-GCACGTTTTCCGTTCTTGCTG-3' Reverse: 5'-CGGGAGTAGATCCAAAATCTC-3'), F79A (Forward: 5'-GCAAAGAGCCGCGGTAAACTG -3' Reverse: 5'-CGCCAAAATCCCTGTAGCCTG-3'), and Y91A (Forward: 5'- GCATACTTTGCTGGTATTGAT -3' Reverse: 5'-AAGCTCGCCCGGTTCCAGTTT -3'). Proper insertion of mutations was confirmed by sequencing.

All FabZ constructs were transformed into *E. Coli* BL21(DE3) and inoculated into ZYP-5052 medium containing 100 µg/mL of Kanamycin and grown at 37°C until A_{600} reached approximately 0.6. These cells were then allowed to autoinduce overnight at 20°C.⁷⁴ The cells were harvested by centrifugation at 5000g for 20 minutes and resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 5 mM imidazole), supplemented with 1X Bugbuster reagent (EMD Millipore). After 20 min of stirring at room temperature, benzonase was added and the solution was stirred for an additional 10 minutes. The mixture was centrifuged at 17,000g for 20 min at 4°C,

yielding a clear supernatant, which was applied to a column with Ni-NTA resin (Qiagen) pre-equilibrated in buffer A. The column was washed several times with buffer B (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 20 mM imidazole) and eluted with buffer C (20 mM, Tris-HCl, pH 8.0, 500 mM NaCl, and 200 mM imidazole). Fractions containing FabZ were then concentrated and loaded onto a size exclusion (S-200, GE Healthcare Life Sciences) column previously equilibrated with buffer D (2mM dithiothreitol, 200mM NaCl, and 20mM HEPES buffer at pH 7.4).

Crystallization and Structure Determination. FtFabZ crystals were grown at room temperature using the sitting drop vapor diffusion method in a 1:1 solution of protein (6 mg/ml in buffer D) and reservoir solution consisting of 0.2M ammonium sulfate, 0.1M BIS-TRIS pH5.5, 25% PEG 3350. Rectangular shaped crystals appeared after approximately 48 hours and were flash frozen in liquid nitrogen using 20% ethylene glycol or 20% glycerol as a cryoprotectant. Hexagonal shaped YpFabZ crystals were grown in a similar fashion using a reservoir solution of 8% tacsimate pH 8.0, 400mM sodium thiocyanate and 20% PEG 3350. X-ray diffraction data were collected at the x29 beamline at the National Synchrotron Light Source with an ADSC Quantum 315 detector. FtFabZ and YpFabZ crystallized in space groups C2 and P6₃, respectively. FtFabZ diffracted to at least 2.5Å, while for YpFabZ this value was 1.9 Å for the structure containing bound glycerol and 1.8Å for the structure containing bound ethylene glycol. All data was processed using the hkl2000 suite of programs.⁷⁵ Both structures were determined by molecular replacement via the CCP4 suite module, MOLREP.⁷⁶ The structure of PfFabZ (1ZHG) was used as a search model. Following the initial structure

determination, several iterative rounds of refinement via the CCP4 module⁷⁷, REFMAC⁷⁸ and model building with COOT⁷⁹, were performed. During the last stages of refinement, water molecules were selected based on peaks above 1σ in the $2F_{\text{obs}}-F_{\text{calc}}$ electron density maps. Further significant residual density, not corresponding to water molecules, was determined to be glycerol and ethylene glycol, which had been used as cryoprotectants for various crystals. Density corresponding to the planar structure of the additive sodium thiocyanate was also resolved in the final structure. The R and R-free converged to 0.16 and 0.18 for glycerol bound YpFabZ, 0.17 and 0.20 for ethylene glycol bound YpFabZ, and 0.21 and 0.28 for FtFabZ. The final refinement statistics are given in Table 2.1.

Table 2.1: Crystal data and refinement statistics of YpFabZ and FtFabZ

Name/code [*]		YP(Gly)	YP(EG)	FT
Cell dimensions	a (Å)	104.76	104.08	136.71
	b (Å)	104.76	104.08	56.03
	c (Å)	87.96	88.33	140.09
	α (°)	90.0	90.0	90.0
	β (°)	90.0	90.0	113.0
	γ (°)	120.0	120.0	90.0
Space group		P6 ₃	P6 ₃	C2
No. of molecules/ASU		2	2	6
Resolution range (Å)				
Overall		50–1.9	50–1.8	50 – 2.55
Last shell		1.96 -1.9	1.85-1.8	2.64 -2.55
Number of unique reflections		42366	49818	31876
Completeness (%)				
(Overall/Last shell)		98/85	99/89	99/91
R _{merge} ¹ overall/last shell		5.0/8.0	7.0/41	8.5/22
Refinement Statistics				
Resolution (Å)		50-1.9	50-1.8	50-2.55
R factor ² /Rfree (%)		15.7/18.3	17.3/19.7	21.0/28.0
R.M.S deviations from ideality				
Bond lengths (Å)		0.03	0.01	0.01
Bond angles (°)		2.4	2.3	1.70
Average B-factors (Å ²)				
Main chain		18.4	22.5	48.1
Side chain		26.3	29.5	54.0
Waters		30.1	31.1	39.6
Residues (%) in the core region of φ-ψ plot		87.0	88.0	85.3

¹R_{merge} = $\sum_j (|I_h - \langle I_h \rangle|) / \sum I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry equivalents

²R-factor = $\sum |F_{obs} - F_{calc}| / \sum |F_{obs}|$
^{*}YP(Gly): YpFabZ with glycerol bound; YP(EG): YpFabZ with glycerol bound; FT: FtFabZ

Enzymatic Characterization. The enzymatic activities of FtFabZ and YpFabZ were determined via the reported spectrophotometric method using the substrate analog crotonoyl-CoA.¹⁵ Initial reaction velocities were determined by monitoring the decreasing absorbance at 280 nm, resulting from the conversion of crotonoyl-CoA to β -hydroxybutyryl-CoA. An extinction coefficient of $3600 \text{ M}^{-1} \text{ cm}^{-1}$, corresponding to the hydration of the enoyl double bond, was used to convert absorbance readings to substrate concentrations.⁸⁰ Reactions were performed using a Beckman-Coulter DU730 UV/Vis Spectrophotometer. Reaction volumes of 100 μL were prepared in a buffer solution consisting of 2.0 μg FtFabZ, 2.0 μg YpFabZ (H19A) or 0.5 μg YpFabZ and 20mM TRIS pH 7.0. All reactions were performed at 25⁰C. K_m values for both FtFabZ and YpFabZ were determined by varying the substrate concentration. Inhibition reactions for IC_{50} calculations were performed under similar conditions with a fixed concentration of 100 μM crotonoyl-CoA. All tested inhibitors were dissolved in 100 % dimethyl sulfoxide (DMSO) and added to the reaction mixture to a final concentration of $\leq 1\%$ DMSO. Initial inhibitor screens were performed in identical reaction mixtures to those described above, except the assays were conducted using a Tecan Safire² Microplate Reader. Compounds for inhibitor screening were obtained from the National Cancer Institute's open repository.

Docking Studies. All docking calculations were performed with AutoDock Vina⁸¹ using AutoDock Tools³ for the preparation of ligand and receptor PDBQT files. Grid maps for

the docking calculations were made using 60 grid points with 0.375 Å spacing in the x, y, and z direction centered at the active site.

Results and Discussion

Structure of YpFabZ. Consistent with other reported FabZ structures, YpFabZ forms a hexamer with a trimer of dimers configuration (Fig. 2.1a). The FabZ monomer adopts a $\beta + \alpha$ hot-dog fold, where a 6-stranded β -sheet with the topology 1-2-4-5-6-3 wraps around a long central α -helix (α_3), with the β_3 strand containing a brief interrupted segment. A short α -helical region is formed near the N-terminus (α_1) as well as an additional short α -helix (α_2) between β_2 and α_3 . Short disordered regions separate individual β -strands, however the α -helices have longer loop regions in the intervening regions between themselves and adjacent secondary structure elements. These loops are defined as follows: Loop A resides between α_1 and β_1 , loop B lies between α_2 and α_3 and loop C is formed between α_3 and β_3 (Fig. 2.1b). These loops are present on the inside of the “hot-dog” making further contacts that stabilize either the dimer or the higher order hexamer.

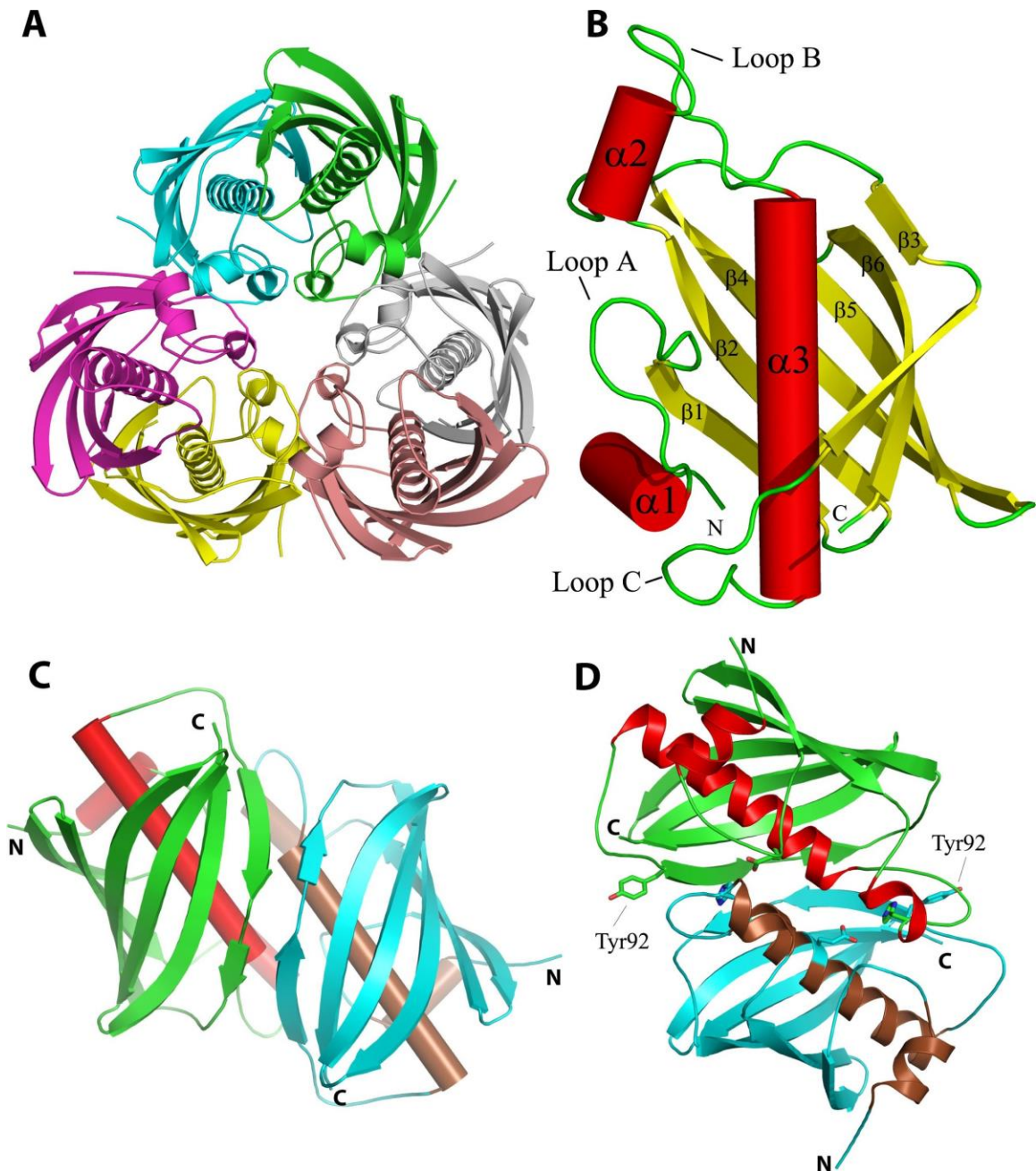


Figure 2.1: (A) Ribbon diagram of the YpFabZ hexamer. Dimer A/B is colored cyan/green, Dimer C/D is colored grey/salmon, and Dimer E/F is colored yellow/magenta. (B) Ribbon diagram of the YpFabZ monomer with α -helices in red, β sheets in yellow, and loop regions in green. (C) Ribbon diagram of YpFabZ dimer. (D) Ribbon diagram of YpFabZ dimer with locations of the two individual active sites marked by stick representation of the Tyr92 gatekeeper residue and the catalytic residues, His54 and Glu68. Tyr92 is labeled for clarity.

The dimer itself is formed by the association of each monomer along the groove of the β_3 strand creating a 12-stranded extended β sheet (Fig. 2.1c). The α_3 helices of each monomer are oriented antiparallel to one another making further contacts along the N-terminal two turns of each helix. The substrate-binding tunnel is formed in the dimeric interface and hence each dimer harbors two active sites located approximately 13 Å apart (Fig. 2.1d). The entrance to each active site tunnel is marked by Pro104 on one side of the entrance, opposed by Tyr92 on the opposite side (Fig. 2.2a). When structurally aligned, Tyr92 occupies the same position as Tyr100 in HpFabZ and Leu170 in PfFabZ, both of which are posited to act as entrance gatekeeper residues. Similar to the gatekeeper residues of HpFabZ and PfFabZ, Tyr92 is highly flexible and capable of completely blocking the tunnel entrance by adopting a “closed” conformation or alternatively can allow access to the tunnel by adopting an “open” conformation (Fig. 2.2d). Inside the tunnel, approximately in the region of the tunnel kink, the catalytic site formed by the conserved His54 and Glu68 residues is found. These residues are contributed by opposing monomers with His54* (*indicates the opposing monomer) coming off of loop B of one monomer and Glu68 coming off the N-terminus of α_3 on the other monomer. Deeper into the tunnel, past the catalytic site, the tunnel narrows and largely follows the path of the α_3 helix, passing Tyr91 and ending at Phe79. With the exception of the catalytic residues, the tunnel is almost entirely hydrophobic.

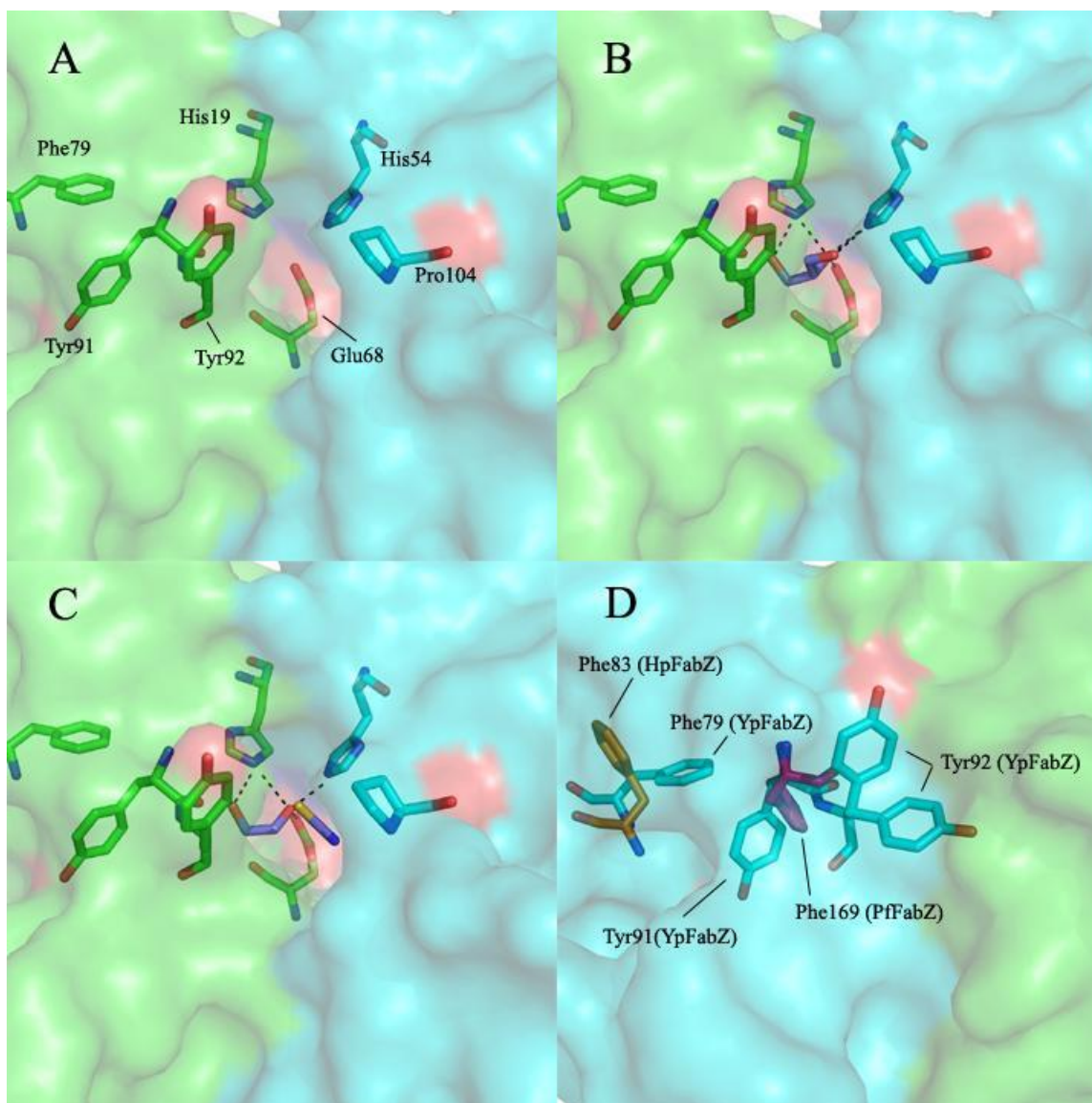


Figure 2.2: Surface representation of the entrance tunnel and active site region of YpFabZ with key residues and ligands shown in stick representation. Individual YpFabZ monomers are colored green or cyan in both surface and stick representation. Hydrogen bonds between ligands and YpFabZ residues are shown as black dashed lines. (A) YpFabZ entry and exit gate residues as well as active site residues are named. (B) Interactions between glycerol and active site residues. (C) Interactions between ethylene glycol and thiocyanate with active site residues. (D) Overlay of entrance and exit gate residues from YpFabZ, HpFabZ, and PfFabZ superimposed on YpFabZ surface representation. Residues from different FabZ isoforms are colored cyan (YpFabZ), magenta (PfFabZ), or orange (HpFabZ). Residues from HpFabZ and PfFabZ are shown in the “open” conformation. Tyr92 from YpFabZ is shown in both the “open” and “closed” conformation, while Phe79 and Tyr91 were found only in the “closed” conformation.

Analysis of the structures of PfFabZ and HpFabZ has led to two separate residues being proposed as potential exit gate residues, with both being found to adopt either an “open” or “closed” conformation, analogous to the tunnel entrance gate residue. For PfFabZ this residue is Phe169, which corresponds to Tyr91 in YpFabZ and in HpFabZ it is Phe83, which corresponds to Phe79 in YpFabZ. However, in the structure of YpFabZ, both Phe79 and Tyr91 were found only in the “closed” conformation (Fig. 2.2d).

In addition to the above structural features, a glycerol molecule, added as a cryo-protectant, was found bound deep in the active site tunnel making interactions with the catalytic dyad residues, His54* and Glu68. One of the hydroxyl groups of the glycerol molecule makes a bifurcated hydrogen bond with each of the catalytic dyad residues. Another hydroxyl group forms a hydrogen bond with the conserved His19 (Fig. 2.2b). Similarly, in a separate structure, ethylene glycol was found making comparable contacts (Fig. 2.2c). Additionally, in the ethylene glycol bound structure, a thiocyanate molecule occupies the space between the tunnel entrance and ethylene glycol.

Structure of FtFabZ. The overall architecture of FtFabZ is nearly identical to that of YpFabZ, except for an extended kinked portion of the FtFabZ $\alpha 3$ helix (Fig. 2.3a,c). The active sites of YpFabZ and FtFabZ are also very similar with both catalytic residues, as well as the second conserved histidine (His19 in YpFabZ and His22 in FtFabZ), having nearly identical positions in both structures (Fig. 2.3b). Also, Phe104 has a conserved position with respect to Tyr91 of YpFabZ. This residue corresponds to one of the two residues identified as potentially acting as an exit gate. This position is well conserved across all FabZ structures reported to date, normally occupied by a tyrosine or

phenylalanine, with the exception of HpFabZ and CjFabZ where it is a valine. One difference between YpFabZ and FtFabZ is at the entrance gate residue, where FtFabZ has a non-aromatic amino acid (Met105), in contrast to the tyrosine or phenylalanine that occupies this location in most FabZ isoforms. Met105 did not demonstrate the same flexibility as seen for the YpFabZ entrance gate residue and was found to occupy only a single “closed” conformation (Fig. 2.3b). The region surrounding the tunnel exit of FtFabZ adopts a conformation that is unique to the FabZ structures reported to date with the $\alpha 3$ helix containing an additional turn that kinks prior to the start of loop C (Fig. 2.3a,c). This feature is distinct, but somewhat reminiscent of the extra α -helix formed out of loop C in the HpFabZ structure.²⁰ This additional helical turn appears to derive from an approximately seven amino acid insertion in this region as compared to YpFabZ and HpFabZ. Based on sequence alignment,⁸² Phe90 would be expected to occupy the exit gate region of FtFabZ, however in the structure this residue is far removed with Glu83 taking its place (Fig. 2.3c). This feature is unusual in that the exit gate residue is hydrophobic in all other reported FabZ structures, with the predominant amino acid being a phenylalanine, the exceptions being PfFabZ and NmFabZ, which have a leucine and isoleucine, respectively.

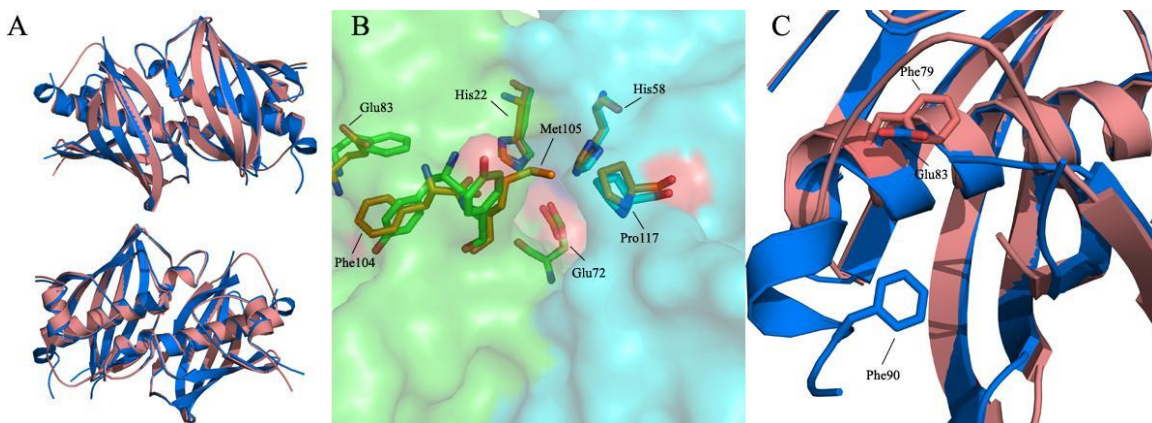


Figure 2.3: (A) Ribbon diagram of structural alignment of YpFabZ (salmon) and FtFabZ (blue) dimers. (B) Structural alignment of the gate and catalytic residues of YpFabZ and FtFabZ. FtFabZ residues are labeled and colored orange. Unlabeled YpFabZ residues are colored green and cyan. Background surface representation is of YpFabZ with individual monomers colored green and cyan. (C) Close up view of extended $\alpha 3$ helix of FtFabZ (blue) aligned to YpFabZ (salmon). Key exit gate residues are shown as sticks with colors corresponding to the ribbon representation.

Analysis of previous FabZ structures has identified three key residues that may act as gate residues. One of these residues, located at the entrance to the active site tunnel, is thought to control access to the catalytic residues. On the opposite end of the tunnel, two separate residues have been suggested as potential exit gate residues with one or both possibly functioning to control the length of the tunnel to accommodate various length carbon chains. Individually these three gate residues are most commonly either a tyrosine or phenylalanine, but often at least one will be a non-aromatic amino acid within a single FabZ isoform. The evidence to date that any of these three residue positions truly act as gatekeepers is largely based on structural information alone. A mutational study of the entrance gate of HpFabZ (Y100A) did show an approximate 50% reduction in activity over wild-type, however the mutant displayed an approximate 2.5 fold increase in K_m , which could indicate that the effects seen were not based solely on the side chain alone.²⁰ It is also difficult to discern what effect, if any, the entrance gate residue would

have on natural ACP bound substrates. In the case of the exit gate residues, our own mutational studies of YpFabZ (Y91A and F79A) did not show any change in either K_m or k_{cat} (data not shown), however the kinetic analysis was performed with crotonoyl-CoA as a substrate, which has a relatively short carbon chain that likely does not encounter these exit gate residues. Therefore, a proper understanding of the contribution of these residues to enzymatic activity requires a longer chain substrate that reaches deeper into the tunnel, thus contacting these residues.

Biochemical Properties of FtFabZ and YpFabZ. The activity of both FtFabZ and YpFabZ was assessed with a commonly used spectrophotometric assay that utilizes the substrate analog crotonoyl-CoA.¹⁵ The conversion of crotonoyl-CoA to β -hydroxybutyryl-CoA was measured by monitoring the decrease in absorbance at 280 nm as described in the experimental section. This reaction is analogous to the reverse reaction in vivo. The K_m for this reaction was determined to be $78.2 \pm 14.1 \mu\text{M}$ for YpFabZ and $61.8 \pm 8.5 \mu\text{M}$ for FtFabZ (Fig. 2.4 and Table 2.2). The k_{cat} and k_{cat}/K_m for YpFabZ is 1.45 s^{-1} and $14322 \text{ M}^{-1} \text{ s}^{-1}$ and for FtFabZ these values are 0.72 s^{-1} and $11650 \text{ M}^{-1} \text{ s}^{-1}$. Several FabZ homologs have been previously characterized using crotonoyl-CoA as a substrate with reported K_m values of $69.7 \mu\text{M}$ (CjFabZ),⁷¹ $82.6 \mu\text{M}$ (HpFabZ),⁸³ and $86.0 \mu\text{M}$ (PfFabZ).¹⁵ The K_m values obtained for YpFabZ and FtFabZ are of similar magnitude, however the reported k_{cat} value for the other homologs fall between 0.013 s^{-1} and 0.0515 s^{-1} , which is significantly below that of FtFabZ and YpFabZ. The reason for this discrepancy is not immediately apparent, however we found the high salt

concentration of 500 mM used in previous experiments to be strongly inhibitory to both YpFabZ and FtFabZ (Fig. 2.5).

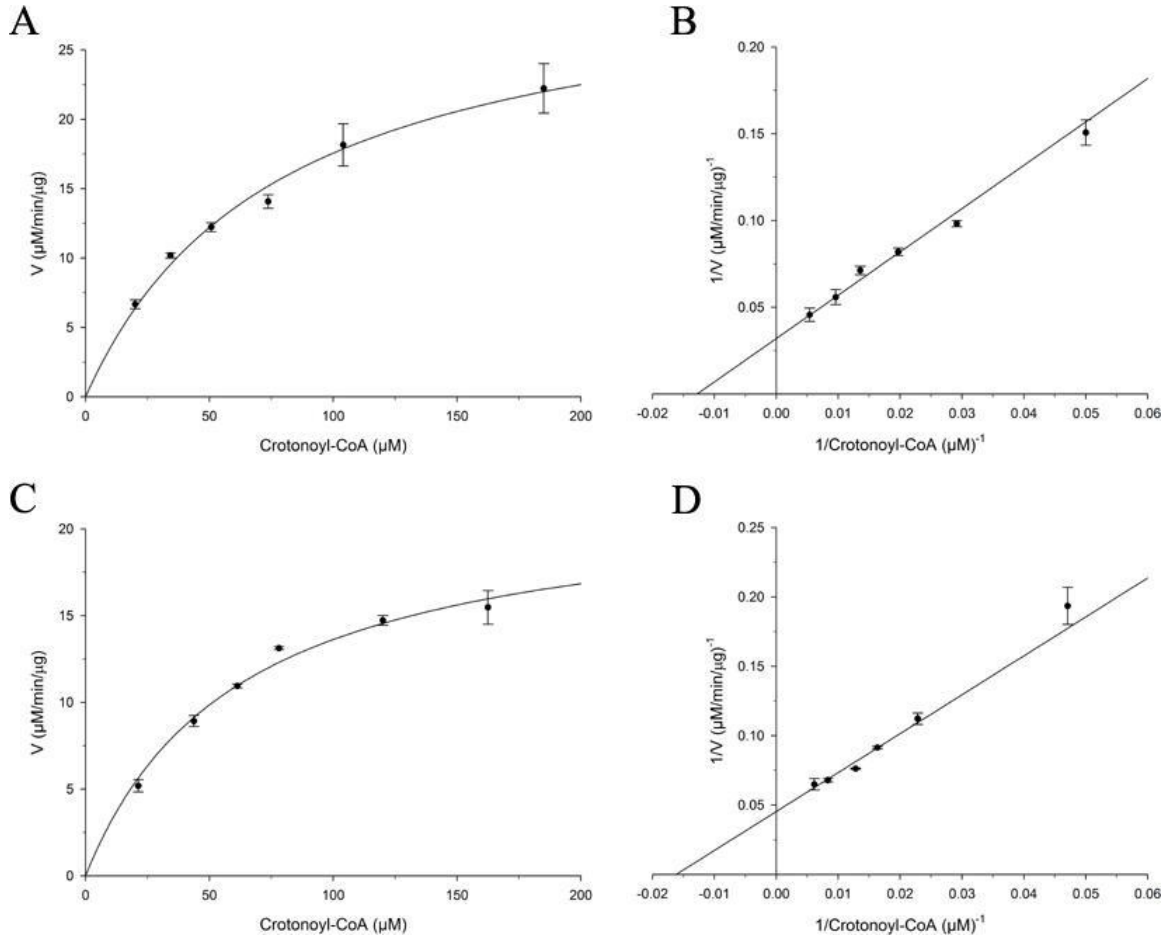


Figure 2.4: Kinetic analysis of YpFabZ (A,B) and FtFabZ (C,D). Initial velocities were determined using increasing concentrations of crotonoyl-CoA. The data were analyzed using non-linear regression analysis and K_m values of $78.2 \pm 14.0 \mu\text{M}$ and $61.8 \pm 8.5 \mu\text{M}$ were obtained for YpFabZ and FtFabZ, respectively.

Table 2.2: Kinetic values for YpFabZ and FtFabZ by spectrophotometric assay.

Enzyme	K_m (μM)	V_{max} ($\mu\text{M}/\text{min}/\mu\text{g}$)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{M}^{-1}\text{s}^{-1}$)
FtFabZ	61.8 ± 8.5	22.1 ± 1.2	0.72	11650
YpFabZ	78.2 ± 14.0	31.3 ± 1.3	1.45	14322

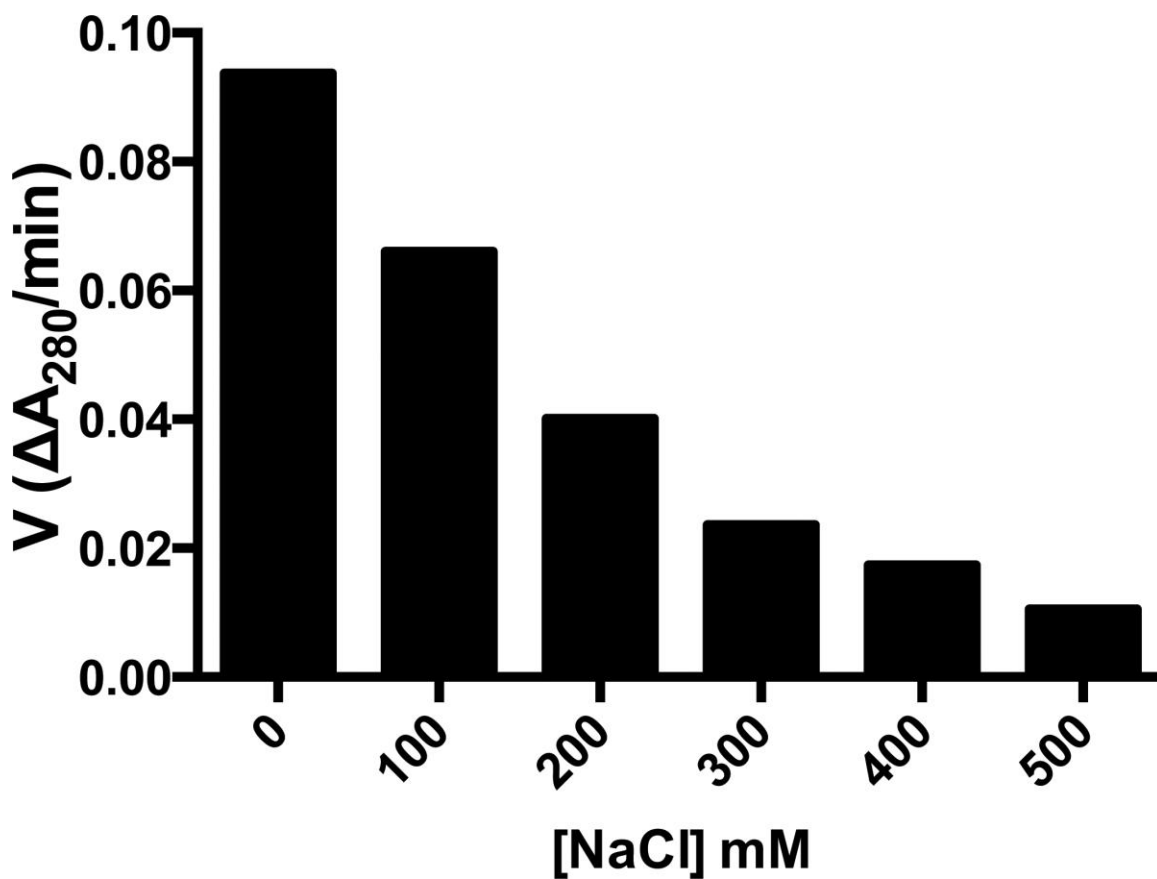


Figure 2.5: The effect of NaCl concentration on YpFabZ enzymatic activity. Each data point represents a single initial velocity calculation obtained at a constant enzyme and substrate concentration.

YpFabZ (H19A) Mutant Study. Our crystal structures of YpFabZ showed both a glycerol and ethylene glycol molecule making hydrogen bonds with both catalytic residues as well as a second histidine in the active site. The proximity of this second histidine to the catalytic residues as well as its complete conservation in all FabZ isoforms suggests it may have some role in enzymatic activity. To test this, we made a YpFabZ H19A mutation and performed a kinetic characterization (Fig. 2.6). The H19A mutation had no effect on K_m with a value of $73.5 \pm 7.6 \mu\text{M}$ as opposed to a K_m of $78.2 \pm 14.0 \mu\text{M}$ for wild-type YpFabZ. However, the k_{cat} of YpFabZ (H19A) was calculated to be 0.49 s^{-1} , which is approximately one third the value for YpFabZ. The agreement

between K_m values suggests that the H19A mutation did not affect the protein fold and this coupled with a modest reduction in k_{cat} points to a catalytic role for this second conserved histidine.

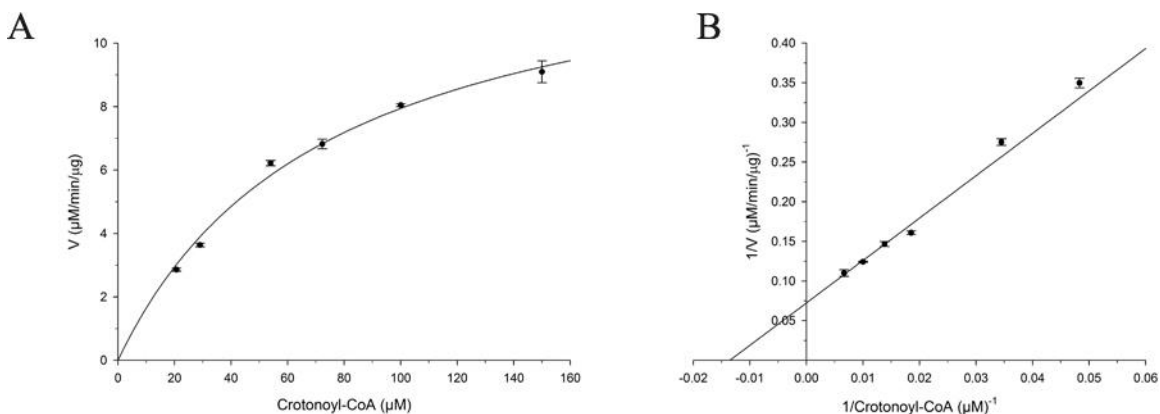


Figure 2.6 Kinetic analysis of YpFabZ H19A mutant. Initial velocities were determined using increasing concentrations of crotonoyl-CoA. The data were analyzed using non-linear regression analysis and a K_m value of $73.5 \pm 7.6 \mu\text{M}$ was obtained.

Inhibitor Screening and Evaluation. The NCI compound database natural product set III, containing 117 compounds, was screened using the spectrophotometric enzyme assay. This set was chosen since it contains numerous scaffolds similar to known FabZ inhibitors and most of the compounds in this set are of appropriate size for the FabZ active site tunnel. Initial screens were conducted at a concentration of 100 μM of each tested compound against YpFabZ. After screening, the two compounds Stictic Acid and Mangostin were identified as potential inhibitors of YpFabZ and were subjected to further analysis (Fig. 2.7b,d). The IC_{50} value of Mangostin against YpFabZ was determined to be $6.1 \pm 1.4 \mu\text{M}$. For Stictic Acid, the IC_{50} was $13.0 \pm 1.4 \mu\text{M}$ (Table 2.3). Inhibition mode characterization demonstrates that Mangostin and Stictic Acid are

competitive inhibitors of YpFabZ with K_i values of $1.4 \pm 0.13 \mu\text{M}$ and $22.6 \pm 2.0 \mu\text{M}$, respectively (Fig. 2.7a,c).

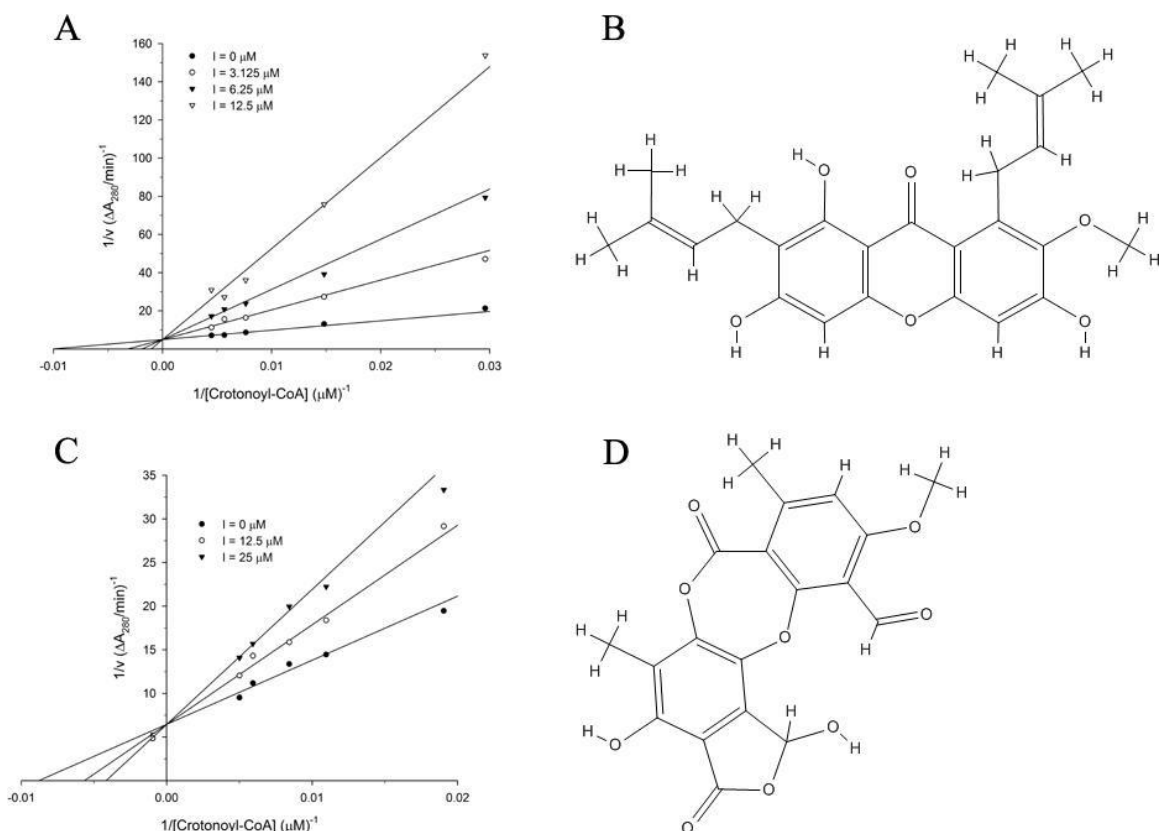


Figure 2.7: Double reciprocal plots of $1/V$ vs. $1/[\text{crotonoyl-CoA}]$ at different inhibitor concentrations are shown for Mangostin (A) and Stictic Acid (C) with YpFabZ. The lines for each inhibitor intercept at the $1/V$ axis, demonstrating that both Mangostin and Stictic Acid are competitive inhibitors of YpFabZ with K_i values of $1.4 \pm 0.13 \mu\text{M}$ and $22.6 \pm 2.0 \mu\text{M}$, respectively. The chemical structures of Mangostin (B) and Stictic Acid (D) are also shown.

Mangostin Docking Study. Docking analysis predicts that Mangostin adopts a binding pose spanning the tunnel entrance, making contacts with the catalytic residues and ending in the region just beyond the active site (Fig. 2.8a). Both of Mangostin's 2-methyl-2-butene groups make extensive hydrophobic interactions with YpFabZ. One of these

groups is predicted to sandwich between Tyr92 and Pro104*, analogous to the position occupied in the “gate blocking” pose of many of the HpFabZ inhibitors. The other 2-methyl-2-butene makes further interactions with the hydrophobic portion of the tunnel found beyond the catalytic residues towards the tunnel exit, making contacts with Tyr91, Ala71, and His19. The xanthone scaffold of Mangostin occupies the region between the tunnel entrance and the catalytic residues, stabilized by hydrogen bonds formed between the 1'-hydroxyl group with the backbone oxygen of Tyr91, as well as the 3'-hydroxyl group with the carboxyl group of Glu68. The 3'-hydroxyl group makes additional polar interactions with His54*. The predicted interaction energy of Mangostin is -10.4 kcal mol⁻¹.

Stictic Acid Docking Study. The predicted binding location of Stictic Acid is similar to that of Mangostin, found bound in the region spanning the tunnel entrance and making contacts near the catalytic residues (Fig. 2.8b). The 5'-methyl group of Stictic Acid is found between Tyr92 and Pro104*, analogous to contacts seen in the Mangostin docking pose. Deeper in the tunnel, near the catalytic residues, the benzodioxepin moiety of Stictic Acid is buried in the hydrophobic pocket formed primarily by Ala71, Leu90, Phe93, Tyr91, Phe55*, His54*, Ile 60*, and Pro62*. Additional hydrophobic interactions are seen between the 8'-methyl group and Leu90. Several polar interactions between Stictic Acid constituent groups and YpFabZ are also predicted, as follows: (1) the 1'-hydroxyl group with the backbone oxygen of Phe61*, (2) the 4'-hydroxyl group with the backbone nitrogen of Pro104*, and (3) the 11'-aldehyde group with the backbone nitrogen of Gly63*. The predicted interaction energy of Stictic Acid is -8.8 kcal mol⁻¹.

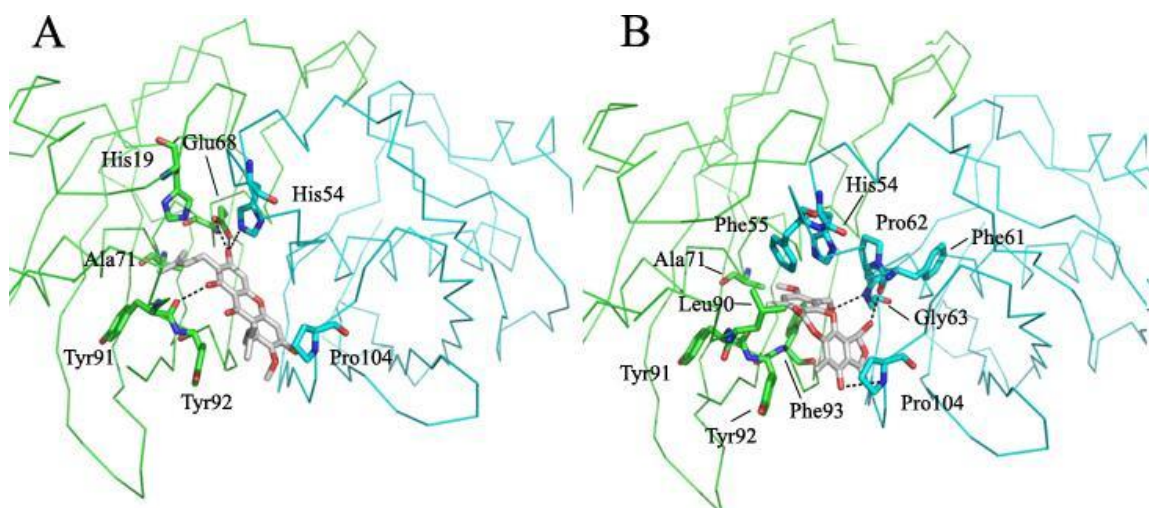


Figure 2.8: Images showing the interactions of Mangostin (A) and Stictic Acid (B) with YpFabZ. The two monomers of YpFabZ are shown as C α chain in green and cyan. Key interacting YpFabZ residues are shown as sticks. Hydrogen bonds are shown as dashed black lines. The predicted binding energies of Mangostin and Stictic Acid are $-10.4 \text{ kcal mol}^{-1}$ and $-8.8 \text{ kcal mol}^{-1}$, respectively. These values are in relative agreement with the calculated K_i values.

Inhibition Studies of Known FabZ Inhibitors. Inhibitor bound crystal structures exist for both HpFabZ and PfFabZ. A common finding among a majority of the inhibitors identified against HpFabZ is that they often have a binding pose at the entrance to the active site tunnel whereby an aromatic portion of the inhibitor is located between the phenol ring of the Tyr100 gate residue and the pyrrolidine ring of Pro112*, forming a sandwich structure.^{20-23, 73} Additionally, these compounds often have a second pose that is deeper into the tunnel and bound in the vicinity of the catalytic and/or exit gate residues. Unlike HpFabZ, the entrance gate residue of PfFabZ is non-aromatic (leucine) and inhibitors show only a single pose residing deep in the tunnel near the catalytic residues.¹⁸ In total, there are approximately 17 residue positions that make side chain interactions with one or more of the known inhibitors of either HpFabZ or PfFabZ. A

sequence alignment of the eight structurally characterized FabZ isoforms demonstrates that these 17 amino acid positions are all highly conserved, with seven positions being strictly conserved (Fig. 2.9). Of the remaining positions that are not strictly conserved, all are substitutions of chemically similar amino acids with a few exceptions. One exception, as has been described above, is at the exit gate residue of FtFabZ where there is a charged residue (Glu83), while the other 7 isoforms have a hydrophobic exit gate residue, which is typically a phenylalanine. Another exception is at the position corresponding to Leu90 on YpFabZ. This position is a hydrophobic residue in all isoforms except FtFabZ where it is a threonine. Lastly, the conserved proline at the tunnel entrance, corresponding to Pro104 on YpFabZ, is a valine in BtFabZ and a glutamine in NmFabZ.

With the exception of the conserved proline marking the tunnel entrance, YpFabZ and FtFabZ together represent much of the natural variation that exists among the binding sites of all the FabZ isoforms. For this reason, we tested our newly identified inhibitors, as well as several commercially available inhibitors of HpFabZ, against both YpFabZ and FtFabZ to determine whether or not these inhibitors are generally cross reactive or specific to a single isoform (Table 2.3).

YpFabZ	----MTTDTHTLHIEEILDLLPHRFPFLLVDRVLD FEE--GKFLRAVKNVSFNEPFFQGHF	55
NmFabZ	----MDVQLP-IEAKDIQKLLIPHRYPFLQLDRITAFEP-MKTLTAIKNVSINEPQFQGHF	54
PaFabZ	-----MMDINEIREYLPHRYPFLLVDRVVELDIEGKRIRAYKNVSINEPFFNGHF	50
BtFabZ	----MSTEKINFDIHKILTLLPHRYPILLVDRVLELEP-HKSIKALKNVTVNEPFFTGHF	55
PfFabZ	-----IDIEDIKKLLPHRYPFLLVDKVIYMQP-NKTIIGLKQVSTNEPFFNGHF	48
FtFabZ	-MSQFNQNNKQIDVMGIRKLLPHRYPFALLDKIVDWSVEDRTIVAQKNVTINEDFFNGHF	59
HpFabZ	MEQSHQNLOSQFFIEHILQILPHRYPMLLVDRIIELQA-NKKIVAYKNITFNEDEVFNGHF	59
CjFabZ	-----MIDVMQIQEILPHRYPFLLVDKITELKV-KEVVLGYKNISISDHVFMGHF	49
YpFabZ	PGKPIFPGVLILEAMAQATGILAFKSRG-----KLEPGELYFFAGIDEARFKRPVV	106
NmFabZ	PDLPVMPGVLIIEAMAQACGTLAILSEG-----GRKENEFFFFAGIDEARFKRQVI	105
PaFabZ	PEHPIMPGVLIIEAMAQAAGILGFKMLDV-----KPADGTLYYFVGSCLKLFRQPVL	102
BtFabZ	PKRPVMPGVLIIEALAQAAALLTFAEAEP-----KDPENTLYYFVGIDNARFKRVVE	107
PfFabZ	PQKQIMPQVLQIEALAQLAGILCLKSSD-----SQKNNLFLFAGVDGVRWKKPVL	98
FtFabZ	PDFPMPGVLIVEAMAQATAILGELMAETLFAHVVEKAGGGRRTFMLAGIDKVRVKRPVV	119
HpFabZ	PNKPIFPGVLIVEGMAQTGGFLAFTSLWGF-----DPEIAKTKLVYFMTIDKVKFRIPVT	114
CjFabZ	PGHPIYPGVLILEGMAQTGGVLAFFESME-----DKVDPKSKVVYFTGIDGAKFRNPVR	102
YpFabZ	PGDQMIMEVEFVKERR--GLTRFTGVAKVDGEIVCTATMMCARSKPAAPAESVVVKPDVV	164
NmFabZ	PGDQLVFEVELLTSRR--GIGKFNAVAKVDGQVAVEAIIIMCAKR-----VV-----	149
PaFabZ	PGDQLQLHAKFISVKR--SIWKFDCHATVDDKPVCSAEIICAERKL-----	146
BtFabZ	PGDQLILNVTFERYIR--GIWKFKAVAEDGKVAAEAELMCTVKTADAAP-----	155
PfFabZ	PGDTLTMQANLISFKSSLGIAKLSGVGYVNGKVVINISEMTFALSK-----	144
FtFabZ	PGDVLVIESRMVKQKN--ICTAESVAKVDGQIVCSAELMAAYKDY-----	163
HpFabZ	PGDRLEYHLEVLKHKG--MIWQVGGTAQVDGKVVAAEELKAMIAERD-----	159
CjFabZ	PGDRLDYEMSVVKNRG--NMWIFKGQAFVDGNLVAAEELKAMIVDK-----	146
YpFabZ	KPDVVKPDVVNPVVKES	181
NmFabZ	-----	
PaFabZ	-----	
BtFabZ	-----	
PfFabZ	-----	
FtFabZ	-----	
HpFabZ	-----	
CjFabZ	-----	

Figure 2.9: Sequence alignment of the eight FabZ isoforms with known structures. Residue positions that interact with known inhibitors are highlighted in red.

When tested, the compounds Apigenin and Quercetin showed no inhibition of YpFabZ or FtFabZ up to a concentration of 200 μ M (Table 2.3). The reasons for the inability of Quercetin and Apigenin to inhibit either YpFabZ or FtFabZ are not obvious. However, a study of the structure-activity relationship between a series of flavonoids against PfFabZ demonstrated that even slight alterations of the hydroxylation pattern of these compounds had major effects on their potency.⁸⁴ In fact, this study revealed that Quercetin, but not Apigenin, which differs by a single hydroxyl group, is an inhibitor of PfFabZ. Furthermore, examination of the co-crystal structures of HpFabZ, with both Quercetin

and Apigenin reveals a pronounced shift in the $\beta 3$ and $\beta 6$ strands, as compared to unliganded enzyme, resulting in the tunnel entrance opening up by approximately 3Å. A similar magnitude shift is seen in other HpFabZ inhibitor bound structures, but in the case of PfFabZ the shift appears less pronounced or even absent. This suggests that inhibitor binding to FabZ may be augmented by an induced fit mechanism. Therefore, the relative flexibility of these regions between isoforms could influence inhibitor binding. This is also supported by a published FabA-ACP complex structure with a crosslinked substrate (Fig. 2.10).⁸⁵ FabA and FabZ have similar function in the FAS pathway and their structures are homologous, with a similar fold and active site residues. In the FabA-ACP structure, the sulfonyl group (of the crosslinker) replaces the carbonyl group of the natural substrate of FabA. When the YpFabZ structure is superposed on FabA of the complex, the sulfonyl group of the crosslinker interacts with the N terminal region of $\beta 3$ (corresponding to Glu86-Phe93 of YpFabZ) of one monomer and the N-terminal region of the $\alpha 3$ (corresponding to Ile60-Ala71 of YpFabZ) of the other monomer. The superposition of FabA with YpFabZ shows the backbone nitrogen of gate residue Tyr92 and the backbone nitrogen of the conserved Gly63 making hydrogen bonds with the two oxygens of the sulfonyl group suggesting that these two regions guide the substrate through the tunnel (Fig. 2.10). This lends support to the hypothesis that during induced fit of substrate or inhibitor, the $\beta 3$ strand moves to allow entry.

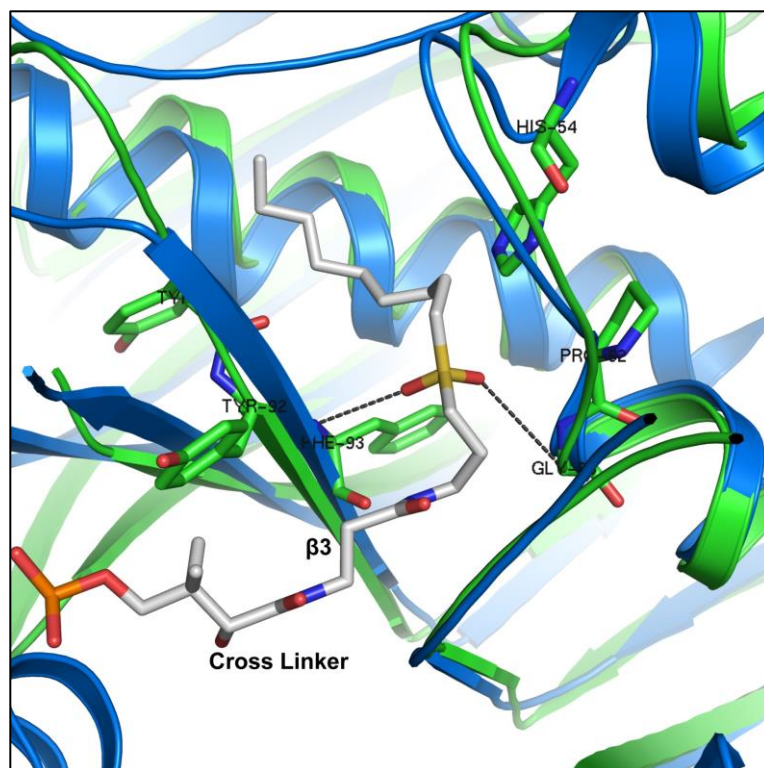


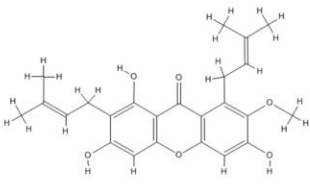
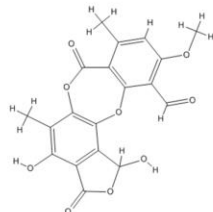
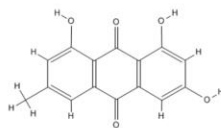
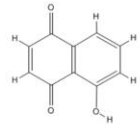
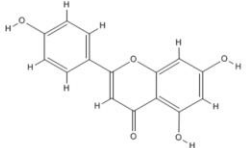
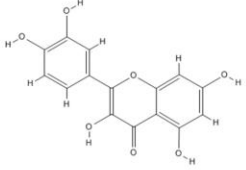
Figure 2.10: YpFabZ superposed on FabA of FabA-ACP complex. YpFabZ and FabA are shown in ribbon representation in green and blue, respectively. Tyr91-Phe93 on the β 3 strand and Pro62-Gly63 in the α 3 region are shown in stick model. The crosslinker is shown in silver grey. Dashed line represents hydrogen bonds.

The remaining compounds we tested inhibited both FabZ isoforms. A direct comparison of each compound against YpFabZ identified Juglone and Mangostin as the most potent inhibitors with IC_{50} values of $5.3 \pm 1.0 \mu\text{M}$ and $6.1 \pm 1.4 \mu\text{M}$, followed by Stictic Acid ($13.0 \pm 1.4 \mu\text{M}$) and Emodin ($29.7 \pm 6.0 \mu\text{M}$). When these compounds were tested with FtFabZ, they generally showed little to no difference in potency as compared to YpFabZ (Table 2.3). At a minimum, this suggests that these compounds are not dependent on an aromatic amino acid as the entrance gate residue since FtFabZ contains a methionine at this position rather than a tyrosine as in YpFabZ and HpFabZ. Supporting this finding is a previous mutational study of the HpFabZ entrance gate residue (Y100L), which showed no impact on the potency of the inhibitors Quercetin, Apigenin, or (S)-

Sakuranetin.²¹ More importantly, the entire range of residue substitutions and structural variations between YpFabZ and FtFabZ appear to be unimportant for the activity of these compounds. Taken together, this data demonstrates that the scaffolds of Mangostin, Stictic Acid, Emodin, and Juglone are all excellent starting points for further inhibitor design against all FabZ variants, while the flavonoids may be more isoform specific.

Our inhibition studies also allowed for a side-by-side comparison of Mangostin and Emodin, which share comparable 3-ringed scaffolds. The results show Mangostin to have an approximate 5-6 fold lower IC₅₀ than Emodin. Inspection of the dock pose for Mangostin with YpFabZ suggests that this effect could be due to Mangostin's ability to exploit hydrophobic regions of YpFabZ on both the entrance and exit sides of the active site tunnel through its dual 2-methyl-2-butene groups. Additionally, Mangostin makes polar contacts with both catalytic residues. This information provides insights into potential functional groups that could be added to existing inhibitor scaffolds for the design of more potent compounds.

Table 2.3: Inhibitory activities of Mangostin, Stictic Acid, Emodin, Juglone, Apigenin, and Quercetin against both YpFabZ and FtFabZ. N/O = None Observed.

Compound	Structure	IC ₅₀ for YpFabZ	IC ₅₀ for FtFabZ
Mangostin		6.1 ± 1.4 μM	7.7 ± 2.0 μM
Stictic Acid		13.0 ± 1.4 μM	27.8 ± 6.1 μM
Emodin		29.7 ± 6.0 μM	43.1 ± 9.2 μM
Juglone		5.3 ± 1.0 μM	5.4 ± 1.4 μM
Apigenin		N/O	N/O
Quercetin		N/O	N/O

Conclusion and Future Directions

In summary, we have reported the crystal structure of two new variants of FabZ (YpFabZ and FtFabZ) as well as their enzymatic characterization, including evidence of a catalytic role for the second histidine located in the active site. Additionally, we have discovered two new inhibitors of FabZ: Mangostin and Stictic Acid. Comparison of the dock pose of Mangostin with the crystallographic pose of Emodin has provided insights into rational modifications of existing inhibitors that may improve their potency. Lastly, we have compared the strengths of six inhibitors against two isoforms of FabZ to identify scaffolds that tend to be cross-reactive. In total, we expect the results of this study to help direct the synthesis of more potent FabZ inhibitors.

[The design of compounds presented in the remaining sections of this chapter was done in collaboration with the Ojima Lab, Stony Brook University]

Design and synthesis of novel FabZ inhibitors. On the basis of our docking studies as well as biological data, we hypothesize that combinations of a gate-blocking and active site binding compound would result in novel FabZ inhibitors with a substantial increase in potency due to dual binding, i.e., fragment-based inhibitor design. A hypothetical binding mode is shown in Figure 2.11, using Emodin as the gate-blocker fragment and NAS21-H, a modified NAS21, as the active site binder fragment, with the two fragments connected by a diester linker. The NAS series of compounds were previously summarized in Figure 1.5.

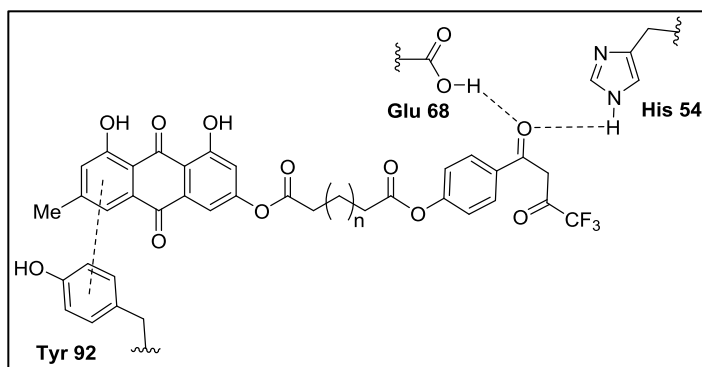


Figure 2.11: Hypothetical multi-binding mode of Emodin-NAS21-H conjugate in YpFabZ.

One of the benzene rings of the Emodin moiety should have a π - π aromatic interaction with Tyr 92 at the gate entrance based on the crystal structure of Emodin-*Hp*FabZ, while the benzoyl oxygen of NAS21-H should have hydrogen bonding interactions with Glu68* and His 54, based on the crystal structure of NAS21-PfFabZ. Figure 2.12 illustrates an initial library of conjugates, using Emodin, Mangostin and Juglone as gate-blocker fragments, and NAS21 derivatives, NAS91, NAS91-10 derivatives, and Juglone as active site binders. Additionally, Juglone can act as either a gate blocking or active site binding inhibitor and can be linked as a Bis-Juglone. As the linker for the two fragments, we will employ alkane dicarboxylic acids such as succinic anhydride, glutaric anhydride and adipic anhydride so that the distance between the two fragments can be varied systematically. We believe these three proposed linkers are sufficient for probing the optimum distance between the two fragments, judging from the available crystal structures. The compounds NAS21-H, NAS21-A, and NAS91-10Ser are unknown in the literature, but their synthesis is straightforward from commercially available starting materials.

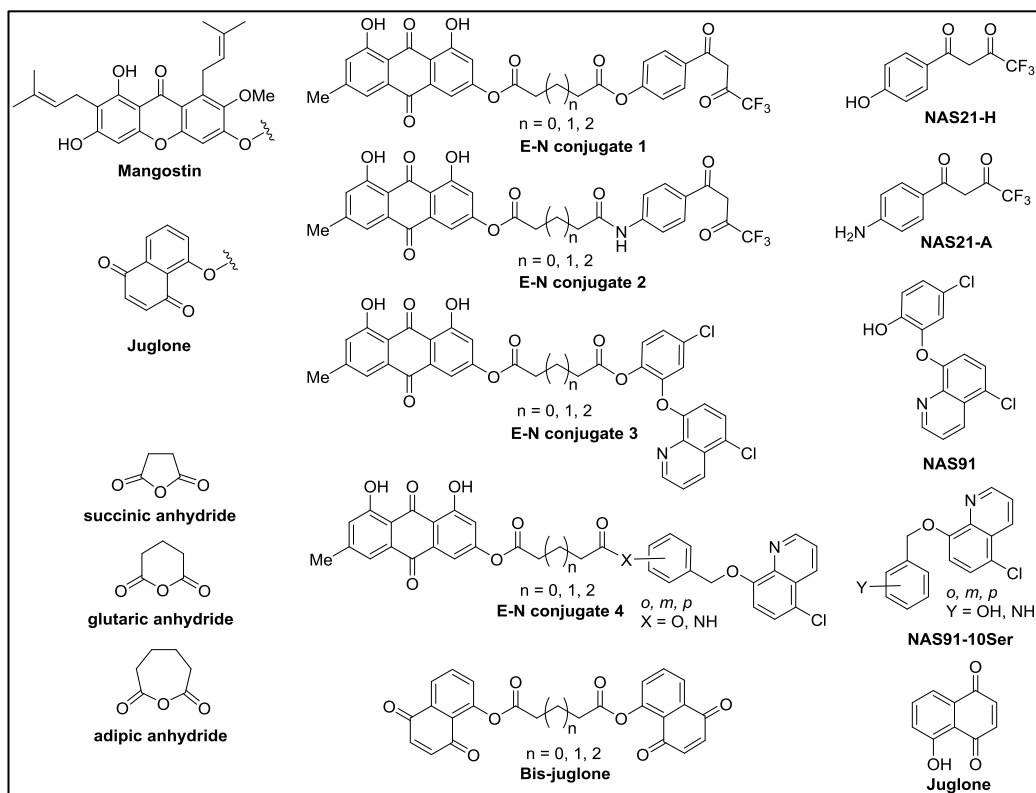


Figure 2.12: Library of gate-blocker/active site binder conjugates.

Although Stictic Acid has been identified as a potentially strong gate-blocking inhibitor, its structure includes rather fragile/reactive functional groups such as aldehyde, lactol, and lactone moieties and will not be included in the initial library design.

Inhibitors Targeting the ACP Binding Groove of FabZ. *In vivo*, the acyl chain substrates of FabZ are presented covalently bound to acyl carrier protein (ACP). For FabZ to act on the substrate, it must first recognize and bind to ACP. Once FabZ binding stabilizes ACP, the thioester region of the substrate inserts into the channel and interacts with the catalytic residues at the active site. Due to the necessity of this molecular recognition event between FabZ and ACP, the ACP binding site of FabZ is another potential target for inhibitor development. Additionally, due to the proximity of the ACP

binding site to the entrance to the active site tunnel, ACP recognition site binders could be further conjugated with gate or channel blocking compounds to increase potency.

The crystal structure of FabA in complex with AcpP from *E. coli* has been recently published and maps out the interactions between FabA and ACP. The interface has complementary charges, with FabA being positively charged and ACP being negatively charged along the binding interface. More specifically, arginines of FabA recognize ACP via glutamates and stabilize its conformation allowing FabA to bind to ACP. FabA is a homolog of FabZ, and FabA of *E. coli* and FabZ of *B. pseudomallei* share fairly good sequence homology with 31% identity in addition to 40% similarity. Based on this, we modeled the *Bp*FabZ:ACP complex (Fig. 2.13). At the interface two arginines (Arg120 and Arg123) and one lysine (Lys 149) of FabZ cluster together to provide a positive electrostatic potential and interact with a number of negatively charged glutamate residues (Glu48, Glu54, Glu61, etc.) of ACP providing both hydrogen bond and salt bridge interactions. Therefore, this site is an excellent target for a negatively charged small molecule or peptide-based inhibitor.

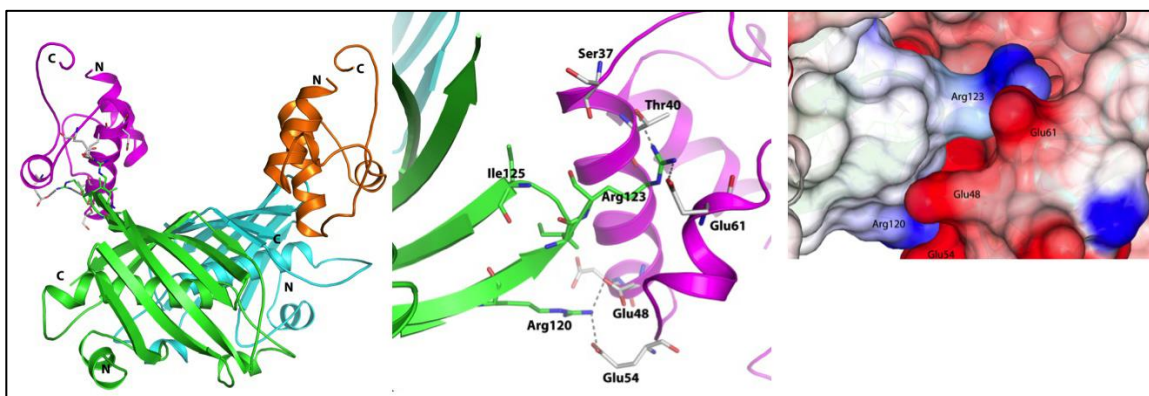


Figure 2.13: Left to Right. *Bp*FabZ:ACP complex based on FabA:ACP model; Interactions at the interface – Arg120 and Arg123 of FabZ interacting with Glu 48 and 61 of ACP; ES potential at the interface – the positive potential (in blue) of FabZ with interacting negative potential (in red) of ACP.

Chapter 3: A Recombinant Fluorescent Substrate With Cleavage Sites of All Seven BoNT Serotypes For High Throughput Screening of Inhibitor Compounds.

Introduction

One excellent target for drug discovery efforts is the Botulinum neurotoxin light chain (BoNT-LC), which contains the protease activity and is required for the neurotoxin's toxicity. Over the years, a number of assays of light chain activity have been developed of various designs.^{60, 61, 64, 86-89} One of the more commonly used assays makes use of commercially available fluorogenic peptide substrates from List Biological Laboratories.⁹⁰ These substrates consist of a short peptide flanking the cleavage site that contains a fluorescent dye donor moiety at one end, and a fluorescence resonance energy transfer (FRET) quencher at the other. In all, there are four distinct substrates for each of the serotypes A, B, C, and E, requiring the purchase of an expensive peptide for each individual serotype to be assayed. A further disadvantage of these fluorogenic peptides is their inability to make contacts with light chain exosite regions. A similar set of FRET based commercial substrates, BoTest A/E and BoTest B/D/F/G, contain several significant improvements.⁶⁴ These substrates cover a larger range of serotypes, are long enough to make exosite contacts, and cover six of the seven BoNT-LC serotypes using only two substrates. However, these substrates are still costly. Other, non-fluorogenic techniques are available such as HPLC based assays,^{62, 63, 91} but they too have the disadvantage of being expensive and also require individual substrates for each BoNT subtype assayed. These HPLC assays have the added drawback of being cumbersome to perform and are therefore not conducive to high throughput screening.

Recently, Hines et al. developed a recombinant fluorescent substrate containing the cleavage site for all seven BoNT subtypes, named GFPSV.⁹² This system relies on substrate immobilization on assay plates followed by liberation of Green Fluorescent Protein (GFP) into bulk solution upon cleavage by BoNT. This GFP based substrate allows for all BoNT subtypes to be assayed and has the added advantage of being able to be expressed and purified from *E. Coli* cells at very low cost. Although this is an excellent and cost effective assay, GFPSV has poor solubility and is prone to aggregation making purification difficult. In this work, we report an improved SNAP-25/VAMP2/GFP hybrid substrate called Repcon. Repcon contains an optimized codon sequence for *E. coli* expression, as well as several alterations in the linker regions separating SNAP-25, VAMP-2, and GFP. These optimizations have drastically improved Repcon's solubility, making it less likely to aggregate as compared to GFPSV. Additionally, Repcon contains an AviTag, which allows for easy and efficient in vivo biotinylation of Repcon, followed by a one-step streptavidin plate binding procedure. Also, mWasabi, which is among the brightest monomeric GFPs available, was selected as the fluorescent molecule resulting in a large boost in fluorescence signal strength, therefore allowing for more reproducible assay results. The assay was tested against known inhibitors of BoNT/E and BoNT/F to demonstrate its suitability in identifying inhibitors to both SNAP-25 and VAMP binding BoNT subtypes.

Methods

Cloning, expression, and purification of Repcon. Repcon, designed as a reporter construct for detection of specific proteolysis induced by all serotypes of botulinum

neurotoxin metalloproteases represent a fusion of an N-terminal poly His tag, followed by green fluorescent protein, a human SNAP-25 fragment (aa 128-206), human VAMP-2 (aa 2-94), and a 15 aa Avi tag (GLNDIFEAQKIEWHE, Avidity). The construct was optimized for expression/cloning in *E. coli* by using host-preferred codons and synthesized by GenScript. After digestion with restriction endonucleases NcoI and XhoI, the synthetic gene was cloned into a pET28a vector (Novagen). The plasmid pBirAcm (Avidity), encoding biotin ligase, was isolated from the *E. coli* strain EVB101 (Avidity), and grown on LB in the presence of chloramphenicol. *E. coli* strain BL21(DE3) was co-transformed with plasmids pBirAcm and pET28aRepcon, and doubly-positive clones were selected by growth on LB medium in the presence of chloramphenicol and kanamycin. A selected clone was grown overnight in the presence of both antibiotics, and 5 mL of the overnight culture was used to inoculate 500 mL of LB medium, and growth was continued until ~ 0.5 OD_{600nm} was reached (approximately 3 hours at 37°C with extensive aeration/shaking). The culture was cooled to 18°C, supplemented with D-biotin (100 μ M final concentration) and IPTG (1 mM final concentration) to induce expression and *in vivo* biotinylation of the protein. The culture was incubated at 18°C for 16 hours, and cells were harvested by centrifugation at 4,500 g in an HBB-6 rotor (Sorwall RC3C centrifuge). Cells were resuspended in approximately ten volumes (~ 100 mL) of buffer A (5 M urea, 100 mM NaCl, 25 mM sodium phosphate, 10 mM imidazole, pH 8.0) and homogenized with three passes through an EmulsiFlex-C3 homogenizer (Avestin). Cleared lysate was collected after ultracentrifugation of the homogenate at 185,000 g in a Ti 45 rotor (Beckman). The cleared lysate was loaded onto Ni²⁺-NTA affinity resin (30 mL of 50% slurry, IBA GmbH), and the column was sequentially washed with

approximately 20 volumes of buffer A, followed by 20 volumes of buffer B (5 M urea, 100 mM NaCl, 25 mM sodium phosphate, 50 mM imidazole, final pH 8.0). The protein was eluted with approximately 10 volumes (~ 150 mL) of buffer C (5 M urea, 100 mM NaCl, 25 mM sodium phosphate, 250 mM imidazole, final pH 8.0). The concentration of total protein in solution was determined to be approximately 0.35 mg/mL as measured by micro BCA protein assay kit (Thermo Fisher). For protein immobilization on streptavidin-coated 96-well plates (Thermo Fisher, Cat # 15503), protein solution was directly dispensed onto the plate (150 μ L per well) and incubated with low intensity shaking for 12 hours at 4°C. The unbound protein in solution was aspirated, wells were washed three times with the buffer D (50 mM HEPES, 10 mM Zn acetate, 2 mM DTT, final pH 7.0), and plates were kept at 4°C until the time of the assay. For evaluation of the specific digest of the Repcon construct on SDS PAGE, the protein was reconstituted with urea-free buffer as follows: Protein solution in buffer C was supplemented with Triton X-100 to a final concentration 0.5%, and saturated aqueous ammonium sulfate (~77g/100mL water) at 25°C was slowly added to the protein solution with constant stirring. The resulting green precipitate (paste) was washed three times with saturated aqueous ammonium sulfate, the liquid phase was carefully aspirated, and the pellet was dissolved in water and dialyzed against buffer D at least three times, 12 hours each at 4°C. Most of the protein precipitated during dialysis, and the concentration of the protein in solution did not exceed approximately 0.1 mg/mL according to measurements obtained with the micro BCA protein assay kit (Thermo Fisher).

Cloning, expression, and purification of BoNT/A-F light chains. Plasmids containing BoNT light chains from serotypes A-F were a gift from Thomas Binz, Institut für Biochemie/Physiologische, Hannover, Germany. The gene for each BoNT serotype was individually subcloned into a pET28b vector, yielding the six plasmids pET28b-BoNT-LC/A-F, all containing an N-terminal 6-His tag. All six BoNT constructs were transformed into *E. coli* BL21(DE3) and inoculated into ZYP-5052 medium containing 100 µg/mL of Kanamycin and grown at 37°C until A_{600} reached approximately 0.6. These cells were then allowed to autoinduce overnight at 20°C.⁷⁴ The cells were harvested at 5000g for 20 minutes and resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 5 mM imidazole), supplemented with 1X Bugbuster reagent. After 20 min of stirring at room temperature, benzonase was added and the solution was stirred for an additional 10 minutes. The mixture was centrifuged at 17,000g for 20 min at 4°C, yielding a clear supernatant, which was applied to a column with Ni-NTA resin (Qiagen) pre-equilibrated in buffer A. The column was washed several times with buffer B (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 20 mM imidazole) and eluted with buffer C (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 200 mM imidazole). Fractions containing BoNT-LC were then concentrated and loaded onto a size exclusion (S-200) column previously equilibrated with buffer D (2mM dithiothreitol, 200mM NaCl, and 20mM HEPES buffer at pH 7.4).

Recon solution studies. An in-solution digest of Recon with BoNTs/A-F was performed in a 30 µL reaction mixture containing 20 mM HEPES, pH 7.4, 2 mM DTT, 10 µM zinc acetate, and 2 µg of Recon. Additionally, ~5ng of each BoNT-LC was added to

the mixture. Each reaction was incubated at room temperature for 1 hour before being analyzed by SDS-PAGE.

Fluorescence Based Continuous Assay. Assays were performed in Repron bound 96-well black streptavidin coated plates (PierceTM catalog # 15503) in a buffer solution containing 20 mM Hepes pH 7.4 and 2mM DTT. All reactions were conducted at 37°C and were monitored using a Tecan Sapphire² Microplate reader at excitation and emission wavelengths of 460 nm and 506 nm, respectively. An initial background signal was present, due to GFP located at the periphery of the plate reader light path. Subsequent initiation of the reaction by addition of BoNT-LC resulted in Repron proteolysis and liberation of GFP into the bulk solution, allowing for the visualization of a normal enzymatic progress curve with a maximum signal far in excess of background. IC₅₀ values were obtained by fitting the initial velocity data to a sigmoid dose-response equation using Prism 6 software.

Results

In-solution Repron digest. The Repron substrate contains the cleavage recognition sites of all seven BoNT serotypes. The relative positions of each cleavage site are shown in figure 3.1. The cleavage sites for BoNTs/A,C, and E are located on the SNAP-25 segment, while that of BoNTs/B,D,F, and G are on the VAMP segment. BoNT/A and BoNT/C cleave adjacent bonds with BoNT/A cleaving between Q197-R198 and BoNT/C cleaving between R198-A199. BoNT/E cleaves between R180 and I181. On the VAMP segment, BoNTs/F and D cleave the adjacent bonds Q58-K59 and K59-L60.

Additionally, BoNT/B cleaves the Q76-F77 bond and BoNT/G cleaves between A81-A82.

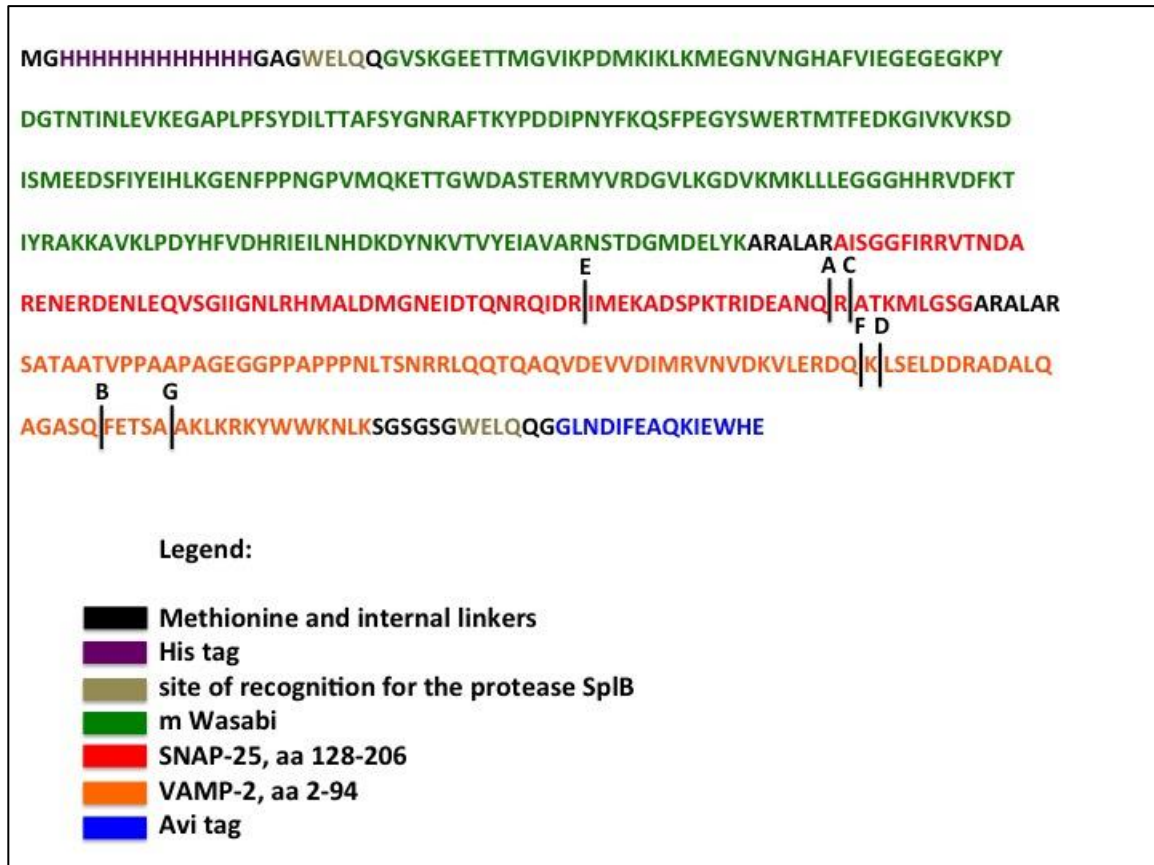


Figure 3.1: Sequence and domain organization of the Repon substrate. Each Repon domain is individually color-coded as indicated in the figure. The cleavage site locations of BoNT/A-G are also shown.

We performed an in solution digest of Repon with BoNTs/A-F to verify that proteolysis occurs at the expected location for each serotype. BoNT/G was not available in our lab and was not tested. As shown in figure 3.2, incubation of BoNTs/A-F with Repon produces a characteristic set of bands for each serotype. The cleavage position of BoNT/A and BoNT/C differs by a one-residue shift on SNAP-25 and expectedly the cleavage fragments are of approximately the same size. The BoNT/E cleavage position is

17 amino acids in the N-terminal direction from the BoNT/A position, resulting in both BoNT/E cleavage fragments falling between those of BoNT/A. The BoNT/B, BoNT/D, and BoNT/F cleavage sites are located near the C-terminus of Repcon, resulting in cleavage products of approximately 43 Kda and 7 Kda.



Figure 3.2: SDS-PAGE of Repcon digest with each BoNT-LC serotype. Undigested Repcon is shown for comparison. Each digestion results in two cleavage fragments of the expected size for each BoNT serotype. A light band of approximately 45 Kda, corresponding to the BoNT light chain is also visible in the BoNT/C and BoNT/D lanes.

Repcon Fluorescence Assay. Having determined that Repcon undergoes site-specific cleavage by BoNTs/A-F, we next tested each serotype in a fluorescence-based assay. After immobilizing biotinylated Repcon on streptavidin-coated plates, buffer solution containing 50 nM of a single BoNT subtype was added to individual wells. Figure 3.3

displays the resulting progress curves and initial velocities for BoNT/A-F. As has been previously discussed, this progress curve results from the increase in fluorescence signal as immobilized GFP at the periphery of the plate reader light path is liberated into bulk solution by BoNT proteolysis of Repcon. The six subtypes tested all produced a robust fluorescence signal as compared to the blank control. The relative initial velocities for each serotype are shown in figure 3.3b. The V_i of BoNT/C was significantly below that of the other serotypes; however, this is not unexpected based on the low level of activity of BoNT/C observed in previous studies.^{91, 93, 94} The buffer-only control showed no signal increase, demonstrating that the progress curves are due to BoNT proteolysis alone rather than nonspecific release of Repcon into solution.

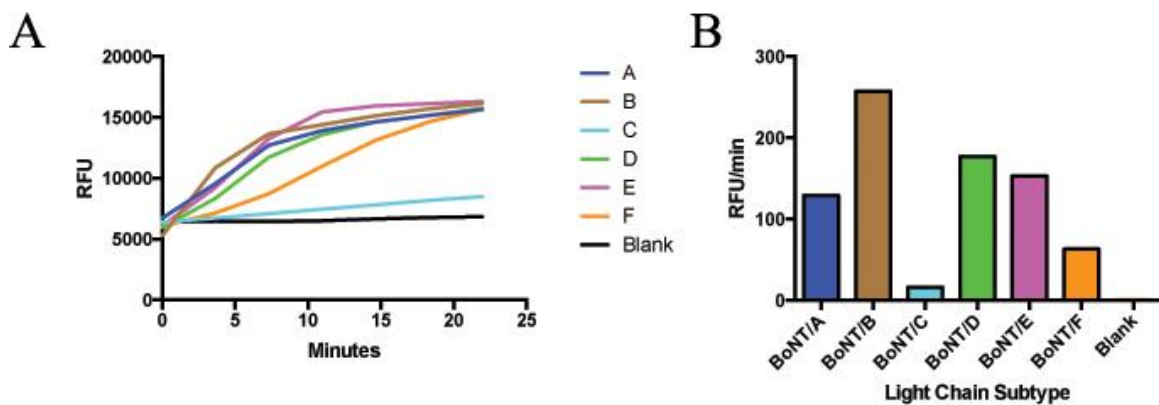


Figure 3.3: Fluorescence signal produced by cleavage of Repcon by BoNT/A-F. (A) Raw progress curves for each BoNT serotype and blank control. (B) Bar graph showing relative initial velocities calculated from progress curves shown in panel A.

Fluorescence assay reproducibility. To be useful for high throughput screening of candidate BoNT-LC inhibitor compounds, the fluorescence assay must exhibit reliable and reproducible kinetics. Figure 3.4a depicts raw progress curves after addition of 50 nM of BoNT/E to five randomly selected wells from a Repcon bound 96-well plate.

Visual inspection of the curves reveals a similar time course for each of the five runs. The median initial and maximum signals were 4496 ± 293.5 RFUs and 18155 ± 258.9 RFUs, respectively. More importantly, the median V_i was 275.0 ± 4.9 RFU/min, which represents less than a 2% variation between each data set (Fig. 3.4b). We next tested the reaction velocities as a function of BoNT/E enzyme concentration at a fixed concentration of Repton. The results show a linear relationship, again demonstrating that the observed reaction progress is due to BoNT proteolysis alone (Fig. 3.5). Furthermore, this analysis shows that enzyme concentrations as low as 2.0 nM are sufficient to produce adequate signal strength for assay purposes.

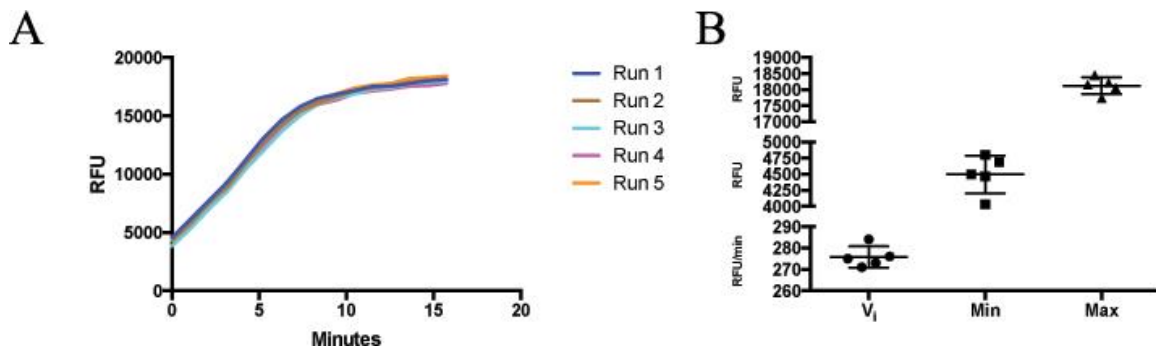


Figure 3.4: Demonstration of the reproducibility of the Repton fluorescence-based assay. (A) Raw progress curves for five independent assays using a fixed concentration of BoNT/E. (B) Scatter plot showing relative V_i , minimum, and maximum signal from five independent runs from panel A.

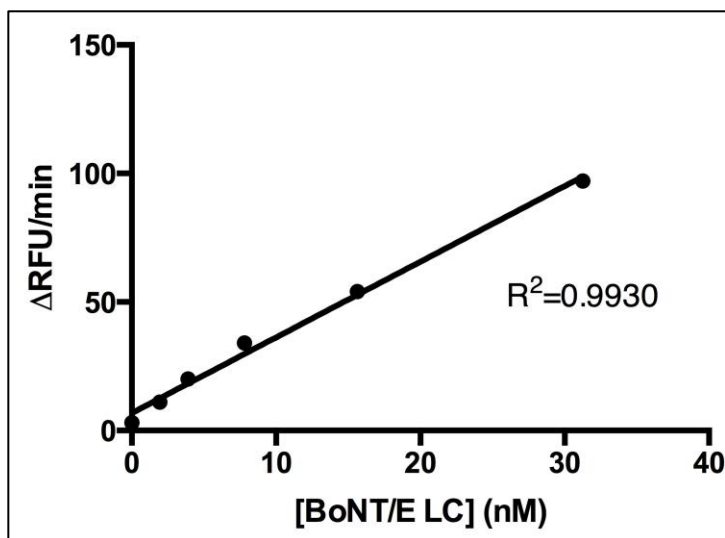


Figure 3.5: Initial velocities plotted against BoNT/E concentrations of 32.0 nM, 16.0 nM, 8.0 nM, 4.0, nM, and 2.0 nM. Plot shows a linear dependence of initial velocity on BoNT/E enzyme concentration.

Testing of known inhibitors using Repcon fluorescence assay. As an additional verification of the usefulness of our Repcon assay, we next tested whether the assay could correctly identify known BoNT inhibitors. The first inhibitor we tested was 2-(9H-fluorene-2-carbonyl) benzoic acid, designated NSC-77053 in the National Cancer Institute repository. This compound has previously been identified as an inhibitor of BoNT/E with a K_i value of 1.29 μM .⁴⁵ When assayed using Repcon, NSC-77053 also showed inhibition against BoNT/E with a calculated IC_{50} value of 48.6 μM (Fig 3.6a). Next we tested the known BoNT/F inhibitor, VAMP₂₂₋₅₈(Q58/D-C).^{46, 47} This inhibitor, which contains VAMP residues 22-58 with a Glu58 substitution to D-cysteine, is one of the most potent BoNT inhibitors identified to date, with a K_i value of 1.0 nM. When tested with Repcon as the substrate, the IC_{50} was determined to be 128.0 nM (Fig. 3.6b). Although K_i and IC_{50} values are not directly comparable, our Repcon results are

consistent in identifying the relative potency of these two inhibitors. Also, the IC_{50} values we calculated are generally of similar magnitude to the K_i values.

As an additional confirmation, we tested NCI-77053 with BoNT/F and $VAMP_{22-58}(Q58/D-C)$ with BoNT/E to ensure that our inhibition results were serotype specific and not due to an artifact of our Repcon assay. When tested, NCI-77053 failed to inhibit BoNT/F and similarly $VAMP_{22-58}(Q58/D-C)$ did not inhibit BoNT/E, even at inhibitor concentrations far in excess of that needed to show 100% inhibition in the correctly paired assay (Fig 3.6c,d). Again, this demonstrates that our Repcon assay can faithfully identify inhibitors of specific BoNT serotypes.

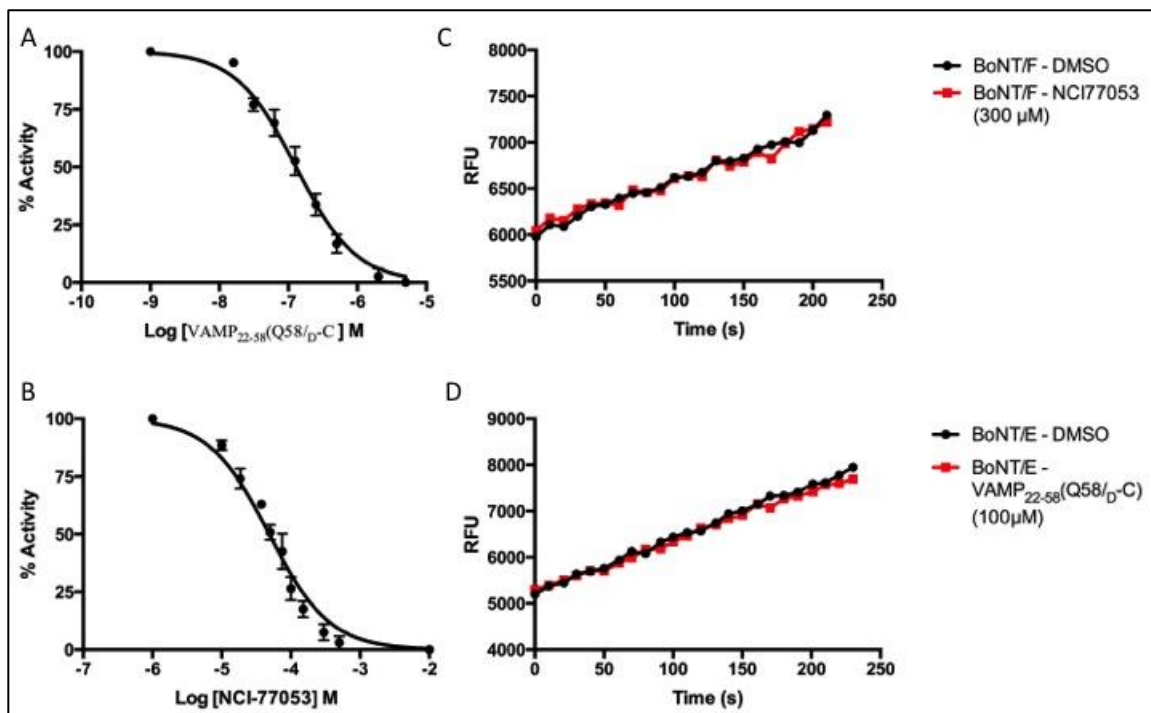


Figure 3.6: Inhibition data for known inhibitors of BoNT/E and BoNT/F when tested using the Repcon assay. (a) IC_{50} curve for compound NCI-77053 with BoNT/E. (b) IC_{50} curve for compound $VAMP_{22-58}(Q58/D-C)$ with BoNT/F. (c) Initial velocity curve for NCI-77053 at a concentration of 300 μ M tested with BoNT/F. (d) Initial velocity curve for $VAMP_{22-58}(Q58/D-C)$ at a concentration of 100 μ M tested with BoNT/E.

Discussion

The first hybrid BoNT substrate (GFPSV) developed contained SNAP-25 (residues 127-206) and VAMP-2 (residues 1-94) separated by linker. In addition, this substrate contained a GFP reporter at the N-terminus and a free cysteine at the C-terminus for chemical conjugation to maleimide bound plates. Although an excellent assay in principle, GFPSV suffers from several severe drawbacks that make its use somewhat impractical. One issue with GFPSV is the choice of maleimide-sulphydryl conjugation for substrate immobilization. The single C-terminal cysteine could lead to the formation of intermolecular disulfide bond formation, which would prevent efficient maleimide conjugation. Additionally, maleimide as a chemical compound is not entirely specific for S-H groups only, which brings into question the precision of immobilization.⁹⁵ A second major issue with GFPSV is that it is heavily prone to aggregation and thus poor solubility. The likely cause of this is internal folding or self-association of protein segments due to the presence of both SNAP-25 and VAMP-2 on the GFPSV substrate. As a result, the purification process for GFPSV is extremely cumbersome and results in poor protein yield. Lastly, the signal obtained from GFPSV is low, resulting in irregularly shaped enzymatic progress curves that make an accurate calculation of initial velocities difficult. The low signal could be the result of several factors, either individually or in combination, including poor or non-specific substrate immobilization or in the choice of GFP molecule.

In this study, we present our recombinant BoNT substrate, Repcon, which represents several major improvements over GFPSV. Like GFPSV, our Repcon substrate is cost effective and allows for all BoNT serotypes to be assayed using a single multi-

domain substrate. However, our optimizations have the added advantage of (1) improving protein solubility and reducing aggregation, thus simplifying the purification process, (2) utilizing an *in vivo* method of substrate biotinylation for easy streptavidin plate binding, and (3) increasing the assay signal strength by utilizing the bright and monomeric mWasabi GFP variant. These changes have resulted in a highly reliable and reproducible method for screening candidate inhibitor compounds of all BoNT serotypes using either a continuous or endpoint assay. Analysis of the variability within the assay setup has demonstrated minimal run-to-run deviation, assuring that any observed reductions in initial velocities during high throughput screening can confidently be attributed to enzyme inhibition alone. Lastly, our examination of known inhibitors of BoNT/E and BoNT/F demonstrates that our assay can faithfully identify inhibitors of both VAMP and SNAP-25 cleaving BoNT serotypes.

Conclusion and Future Directions

Our Repcon based assay is a useful tool for performing high throughput screens of candidate inhibitor compounds for all seven BoNT serotypes. However, the assay has some limitations. Although Repcon is excellent for obtaining IC_{50} values it cannot be used to perform more extensive kinetic analysis, such as the calculation of K_i values, due to the limited amount of substrate that can be bound to streptavidin plates. However, this shortcoming is not specific to Repcon. SNAPtide, which is among the more commonly used BoNT assays, has a reported K_m of approximately 1.0 mM.⁹⁶ As a result, this assay will be conducted under substrate limiting conditions. The high K_m value of SNAPtide is likely due to its short sequence and thus inability to make more extensive contacts at the

BoNT exosite. Therefore, for both Repcon and SNAPtide, any lead compounds identified require an alternative method, such as HPLC, for finer kinetic analysis.

It is plausible that certain modifications can be made to Repcon that will allow for the use of higher substrate concentrations, while maintaining Repcon's high binding affinity, cost effectiveness, and ease of production. One possibility is that Repcon could be divided into two separate substrates: one for VAMP-2 and a second for SNAP-25. Both of these constructs would retain both the N-terminal GFP and C-terminal Avi Tag. However, rather than utilizing the Avi Tag for streptavidin plate binding, it could instead be used for conjugation to dye-labeled streptavidin, such as the AlexaFluor555, which is an excellent FRET pair for GFP. As a result, the substrate would no longer require immobilization and instead could be used for solution studies. Of course, this design is dependent on the Förster distance of the FRET pair, which may limit the acceptable length of the substrate, resulting in similar problems to those seen with SNAPtide. However, the FRET based BoTest substrates, which utilizes SNAP-25 residues 141-206 (BoTest A/E) and VAMP residues 33-94 (BoTest B/D/F/G), maintain K_m values of 0.67 μM and 15.9 μM , respectively.⁶⁴ Therefore, a FRET based redesign of Repcon is conceptually plausible.

It could be argued that these proposed changes are simply a recreation of the BoTest assay. However, the BoTest substrates are extremely difficult to express and purify, and if done incorrectly, the result is poor yield and an unacceptably low fluorescence signal. These difficulties are likely related to poor folding and maturation of the dual fluorophores present in the BoTest constructs and the sensitivity of the BoTest preparation is why these substrates are typically purchased rather than made in-house.⁶⁴

Repcon, however, only has a single GFP fluorophore, and the entire Repcon construct has demonstrated excellent behavior during the protein expression and purification process. The final post-translational coupling of a streptavidin-conjugated fluorophore is a simple process analogous to streptavidin plate binding and is unlikely to adversely affect Repcon's stability in solution.

Chapter 4: Lead Identification Targeting Botulinum Neurotoxin Serotype E Using Computational and Biochemical Methods

Introduction

Botulinum neurotoxin type E recognizes and cleaves SNAP-25 at the Arg180-Ile181 bond. There are currently only two published inhibitors of BoNT/E; the small molecule NSC-77053 and the substrate based peptide inhibitor, RIME, which represents the four residues spanning the scissile bond (P1:Arg180 P1': Ile181 P2': Met182 P3': Glu183). A crystal structure of BoNT/E in complex with RIME has been solved, containing well-defined density for the entire tetrapeptide. This has provided a wealth of information about the specific interactions of each substrate residue with its corresponding S1, S1', S2', and S3' region on BoNT/E, as summarized below and depicted in Fig 4.1.⁴⁴

P1-Arg180: The carbonyl oxygen of Arg180 was found to interact with both the catalytic zinc as well as the strictly conserved Tyr350 residue. Also, Arg180 was found to displace the nucleophilic water present in the apo BoNT/E structures.

P1'-Ile181: The backbone N of Ile181 makes contacts with Oe1 and Oe2 of the catalytic base residue, Gly212, as well as with Thr159. The backbone oxygen of Ile181 interacts with the NH1 and NH2 groups of the Arg347 side chain. These P1' interactions are likely to be common to all BoNT subtypes since they occur between conserved residues of the enzyme. However, the hydrophobic side chain of Ile181 is stabilized within the S1' pocket formed by the residues Thr159, Phe191, and Thr208, which are not conserved. It

has been suggested that at least part of the cleavage site specificity between the BoNTs is derived from differences in the S1' pocket. This theory is bolstered by the observation that the changes to the P1' residue for several BoNTs results in a drastic reduction in K_{cat} . These findings all suggest that the P1'-S1' interaction is critical to the molecular recognition event between BoNTs and their substrate.

P2'-Met182: The backbone oxygen of Met182 interacts with the OH group of Tyr350. The Met182 side chain resides in a hydrophobic pocket formed by Phe191, TYR354, and Tyr356. Phe191 also forms hydrophobic interactions with the P1'-Ile181, as described earlier and the presumed importance of this residue, based on the crystal structure interactions, is supported by an F191A mutational study that found a 75 fold decrease in K_{cat} .⁹⁷

P3'-Glu183: BoNT/E residues 234-244, correspond by sequence alignment to a region on BoNT/A termed the 250-loop. In the apo structure of BoNT/E, the 250-loop was disordered and unable to be resolved. However, interactions with Glu183, particularly at residues Ile240, Thr241, Asn242, and Thr246, stabilize the 250-loop allowing it to be resolved in the crystal structure.

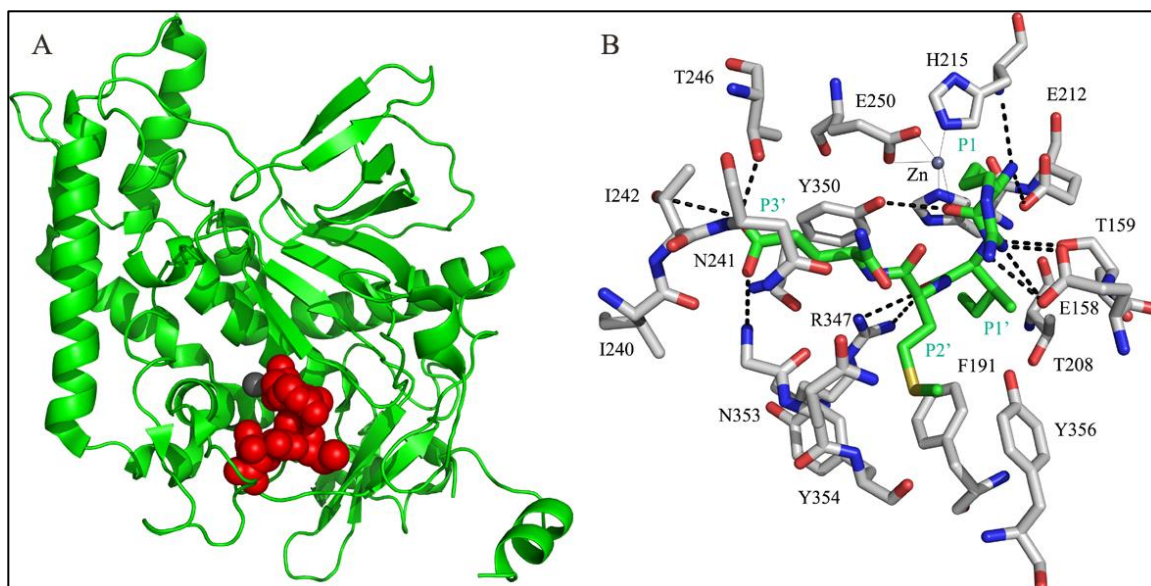


Figure 4.1: The structure of the tetrapeptide inhibitor, RIME, in complex with BoNT/E. (A) Schematic diagram of the BoNT/E-RIME complex. BoNT/E is shown in green cartoon representation. RIME is shown as red spheres. The active site zinc is shown as a grey sphere. (B) Close up view of the BoNT/E active site with RIME bound. BoNT/E residues are shown as grey sticks. RIME is shown as green sticks. The active site zinc is shown as a grey sphere. Black dashed lines show key interactions between RIME and various residues of BoNT/E. Grey lines depict zinc ion coordinating interactions.

Methods

SNAP-Etide assay. Initial testing of inhibitor compounds identified through virtual screening was performed using the well-established SNAP-Etide assay. All reactions were carried out in a 100 μ L reaction mixture containing an optimized buffer solution containing 20 mM Hepes pH 7.4, 2mM DTT, 2.5 nM of BoNT/E light chain, and 100 μ M of each test inhibitor. Each inhibitor was incubated with BoNT/E light chain for 15 minutes and the reaction was initiated by addition of 10 μ M of SNAP-Etide substrate. All reactions were conducted at 37°C in 96-well, black, flat-bottomed, non-binding plates (Corning® Costar®) and monitored using a Tecan Sapphire2 Microplate reader. The excitation and emission wavelengths were set at 320 nm and 420 nm, respectively.

REPCON Fluorescence Assay. Assays were performed in Repcon bound 96-well black streptavidin coated plates (PierceTM catalog # 15503) in a 125 μ L buffer solution containing 20 mM Hepes pH 7.4, 2mM DTT, and 20 nM of BoNT/E. All tested inhibitors were dissolved in 100 % DMSO with a final reaction concentration of < 2% DMSO. Reactions were conducted at 37°C and were monitored using a Tecan Sapphire² microplate reader at excitation and emission wavelengths of 460 nm and 506 nm, respectively. IC₅₀ values were obtained by fitting the initial velocity data to a sigmoid dose-response equation using Prism 6 software with a hillslope = 1.

Results and Discussion

Virtual Screen. Our collaborators in the Rizzo lab performed a virtual screen of 1.4 million compounds selected from the ZINC database. This screen was performed using the program DOCK 6.6.⁹⁸ The receptor coordinates selected for virtual screening studies were from the 2.25 Å resolution crystal structure of BoNT/E light chain in complex with the SNAP-25 substrate derived tetrapeptide, RIME (PDB: 3D3X). The DOCK program, used for virtual screening, does not allow for sampling of the flexibility of the receptor molecule during docking calculations. For this reason, having a crystal structure of the receptor in complex with an inhibitor is preferable since it would presumably represent a conformation that can accommodate small molecule inhibitors. This may not be the case for apo receptor structures, which may have smaller or more constrained binding sites than is the case dynamically. For virtual screening purposes, chain B of the BoNT/E homodimer was selected rather than chain A since the RIME tetrapeptide was not

completely resolved in chain A and four BoNT/E residues (235 - 238), in close proximity to the binding site, were missing. Chain B also had missing residues (59,60), but they are relatively far from the binding site ($> 15 \text{ \AA}$). Also, the N-terminus of RIME was modeled as a deprotonated neutral NH₂ group, based on the crystal structure showing the N-terminus coordinating with the catalytic zinc ion. After the initial virtual screen, the top 100,000 scored compounds were rank ordered and, after visual inspection, 92 compounds were selected for biochemical testing.

SNAP-Etide Screen. The 92 compounds selected from the virtual screen were tested for inhibitory activity against BoNT/E-LC using the established SNAP-Etide assay. The compounds were tested in sets of 10, at a concentration of 100 μM , and were run simultaneously with two DMSO controls. The initial velocities of the control runs were averaged and compared to the initial velocity of each test compound. The percent inhibition of each compound was then calculated as $1 - (\text{TEST}_{(vi)} \div \text{DMSO}_{(vi)})$. During optimization of the SNAP-Etide assay, run-to-run deviations of initial velocity data between control runs was up to $\pm 10\%$, therefore a cutoff of $>20\%$ inhibition was used as the criterion for a “positive hit”. In total, nine compounds met this threshold with activities ranging from 25-71% inhibition (Fig. 4.2, 4.3a).

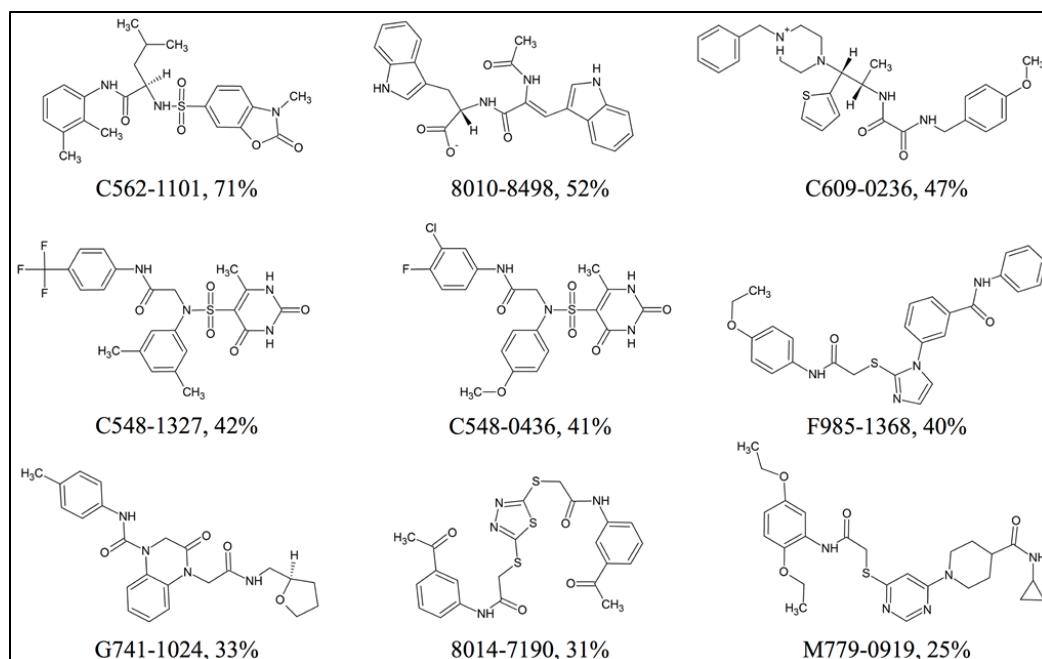


Figure 4.2: 2D structures, compound identifiers, and % inhibition for 9 hits out of 92 experimentally tested compounds using the SNAP-Etide assay

Repon Screen. In general, a significant number of false positives are common to fluorescence-based assays, likely due to fluorescence quenching by the inhibitors themselves. For this reason, we desired a further verification of the SNAP-Etide results. Although Repon is also a fluorescence-based assay, the excitation and emission wavelengths of the Repon GFP fluorophore do not overlap with those of SNAP-Etide (SNAP-Etide Ex: 320 nm and Em: 420 nm, Repon: Ex: 460 nm and Em: 506 nm). Therefore, any quenching effect is unlikely to be common to both assay methods. A further advantage of Repon is that it contains SNAP-25 residues 128-206, which encompasses the SNAP region believed to make exosite contacts with BoNT/E. Therefore this substrate is more physiologically reasonable and is a better indicator of whether test inhibitors can truly compete with the natural substrate for binding.

Using the Repon assay, the nine compounds identified through the SNAP-Etide assay were retested in triplicate at a concentration of 100 μ M of each inhibitor. After

testing, the compound C562-1101 demonstrated the highest activity against BoNT/E, with approximately 90% inhibition (Fig 4.3b). The structure of C562-1101 [N-(2,3-dimethylphenyl)-4-methyl-2-(3-methyl-2-oxo-2,3-dihydro-1,3-benzoxazole-6-sulfonamido)pentanamide] is shown in Figure 4.2. The known BoNT/E inhibitor, NSC-77053 was used as an additional positive control and displayed approximately 70% inhibition. The other eight compounds tested did not show any significant activity. IC₅₀ values were computed for compound C562-1101 as well as for NSC-77053 in order to obtain a side-by-side comparison of their activities. The IC₅₀ value of C562-1101 was 14.2 ± 1.7 μM, and for NSC77053 this value was 48.5 ± 6.5 μM, demonstrating that C562-1101 has superior potency (Fig. 4.4).

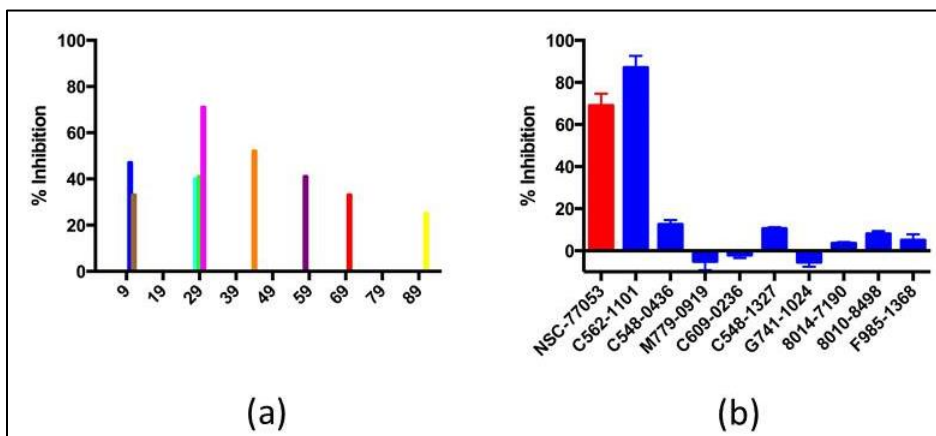


Figure 4.3: Raw inhibitor screening results from both the SNAP-Etide and Repcon assays. (A) Percent inhibition relative to blank control of each of the 92 compounds identified through the virtual screen. Only those compounds showing greater than 20% inhibition are shown. (B) Percent inhibition relative to blank control for the nine compounds identified from SNAP-Etide assay. The known BoNT/E inhibitor, NSC-77053, is shown in red for comparison.

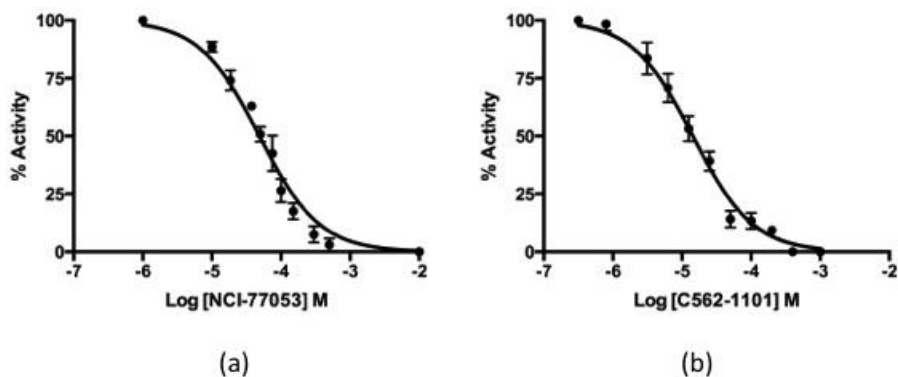


Figure 4.4: IC_{50} curves for (a) NSC-77053 (b) and C562-1101 against BoNT/E. The IC_{50} values for NSC-77053 and C562-1101 were $48.5 \pm 6.5 \mu\text{M}$ and $14.2 \pm 1.7 \mu\text{M}$, respectively.

Structural analysis of compound C562-1101. After virtual screening, the docked compounds were further rescored by our collaborators using the footprint similarity (FPS) scoring function.^{99, 100} FPS attempts to quantify the similarity of interaction patterns of a candidate compound with those made by a known ligand, using a per residue energy decomposition of the ligand-receptor interactions. Theoretically, further pruning of the highest scoring compounds from docking studies, based on how similar their interactions are to those made by known ligands, should result in more chemically reasonable compounds that are likely to be true inhibitors. For our studies, the RIME tetrapeptide was used as the FPS reference. The most promising lead compound identified in experimental testing, C562-1101, had originally been selected from the top ranked FPS_{sum} list, indicating that it is among the compounds making the most similar interaction patterns to RIME.

Figure 4.5 compares the predicted binding geometry of C562-1101 with that of the crystallographically determined binding pose of RIME in the active site of BoNT/E. From this alignment, we see that the benzoxazole group mimics the P3':Glu183 of RIME, locating in the vicinity of the 250-loop and making contacts with Ile240, Thr241,

Asn242, and Thr246. Also, in the RIME structure, the carboxyl group of P3':Glu183 makes polar contacts with the backbone oxygen of Ile240 and the backbone nitrogen of Gly352. In the docked pose of C562-1101, these same contacts are reconstituted by the carbonyl oxygen and ring oxygen of the benzoxazole group. Additionally, the benzoxazole carbonyl oxygen occupies a position near a water molecule that was shown to hydrogen bond with the carboxyl group of P3':Glu183 in the RIME structure.

In the RIME structure, at the P1':Ile181 position, the backbone oxygen makes polar contacts with Arg347 of BoNT/E. The sulfonamide oxygen of C562-1101 overlays with the backbone oxygen of P1':Ile181 and makes an identical contact with Arg347. In addition, the propyl group of C562-1101 occupies the same position as the P1' Ile181 side chain and is buried in the S1' hydrophobic pocket formed by Thr159, Phe191, and Thr208.

The backbone oxygen of P1:Arg180 of RIME interacts with the zinc ion, as well as Tyr350 of BoNT/E. This oxygen also displaces the nucleophilic water present in the BoNT/E apo structure. The carbonyl oxygen of the pentanamide moiety of C562-1101 occupies this same position and makes identical contacts with the zinc ion and Tyr350.

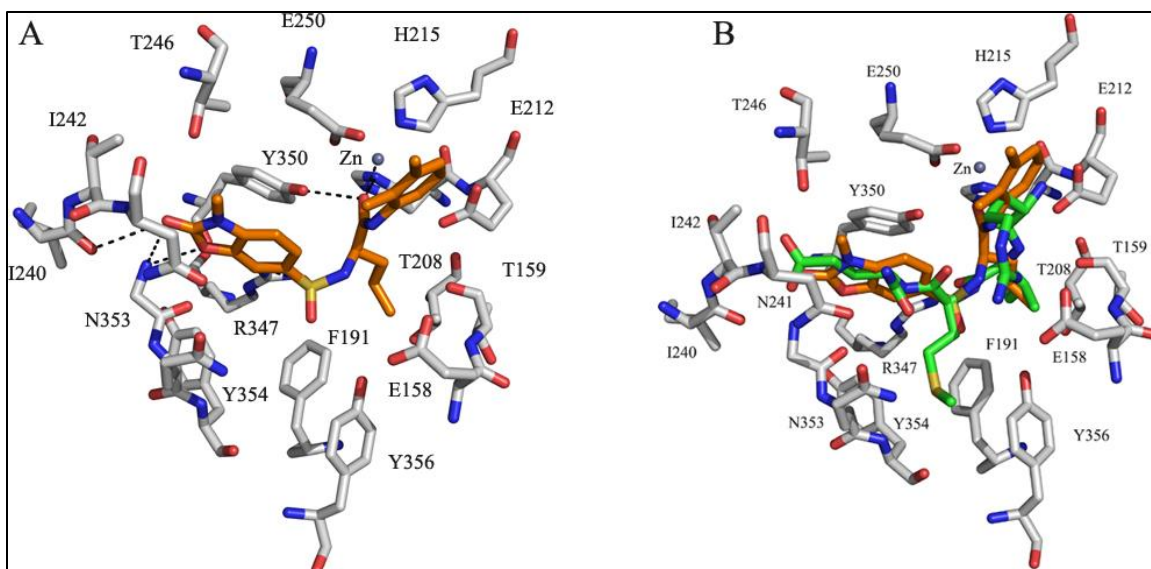


Figure 4.5: Comparison of the Dock pose of compound C562-1101 with the crystal structure pose of the RIME tetrapeptide. (a) The dock pose of compound C562-1101 with BoNT/E light chain. Polar contacts are shown as black dashed lines. (b) Compound C562-1101 (orange) overlaid with tetrapeptide RIME (green) in the BoNT/E light chain active site binding pocket.

This structural analysis was also quantitatively supported by the footprint analysis conducted by our collaborators in the Rizzo Lab. Figure 4.6 shows the footprints of both C562-1101 and the reference tetra-peptide RIME. The footprint analysis shows the per residue energy contribution to the total binding energy between the ligand and receptor. In this analysis, the 50 residues contributing most to the binding energy are shown individually, while the energies of the remaining residues are summed and reported as a single value. The van der Waals footprints show overall good overlap between the C562-1101 and the reference RIME, as expected based on the structural analysis. Additionally, the electrostatic interaction footprint indicates that C562-1101 binds strongly to the zinc ion via a carbonyl oxygen, contributing greatly to the overall binding energy. This is in agreement with previous studies that have found the ability of BoNT inhibitors to coordinate the active site zinc ion to be a strong predictor of potency. Overall, C562-1101

does not make as many favorable electrostatic interactions as the reference since it lacks the two charged residues, arginine and glutamic acid, present in RIME. These residues are capable of forming a salt bridge with Glu158 and a favorable hydrogen bond with Thr246, respectively, which is not possible in C562-1101. However, C562-1101 still has several favorable electrostatic interactions in addition to those with the zinc ion, including interactions with Glu158, Glu212 and Arg347.

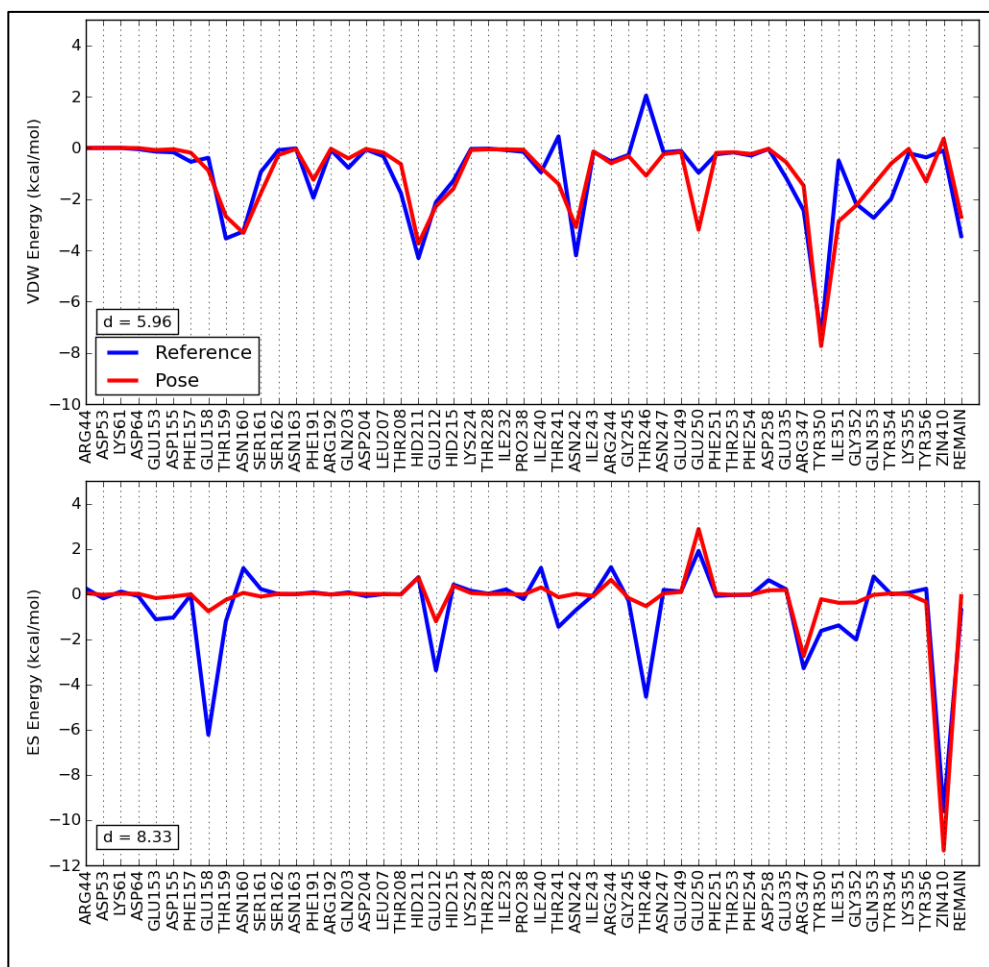


Figure 4.6: Footprint plot of C562-1101 (red) and the reference tetra-peptide RIME (blue) interacting with BoNT/E light chain active domain. Fifty significant residues are shown explicitly with the remaining residues combined into a single aggregate value.

Conclusion and Future Directions

This work represents the initial phases of the structure based drug design process. Our virtual screen identified compound C562-1101 as a low micromolar inhibitor of BoNT/E, the most potent BoNT/E inhibitor discovered to date. The next step in the process is to co-crystallize C562-1101 in complex with BoNT/E, and visualize the specific binding interactions between these two molecules. However, in the absence of a crystal structure, we can use the docking pose to guide modifications to C562-1101 that may improve potency. One interesting feature of C562-1101 is that it contains a substituted pentanamide group, a portion of which is identical to the side chain of leucine. In the Dock pose of C562-1101, this pentanamide group overlays the isoleucine of RIME. In the BoNT/E-RIME structure, this isoleucine resides in the hydrophobic pocket formed by Thr159, Phe191, and Thr208 of BoNT/E. Therefore, a possible modification of C562-1101 is to change the leucine moiety to an isoleucine to mimic RIME. This substitution may make more extensive hydrophobic interactions with the BoNT/E S1' pocket. Also, further inspection of the C562-1101 DOCK pose shows the absence of any significant interactions with the S2' portion of BoNT/E. In the RIME structure, P2':Met182 is found buried in the S2' pocket, formed by Phe191, Tyr354, and Tyr356 and Phe191. Therefore, it is likely that the potency of C562-1101 could be improved by the addition of a properly oriented hydrophobic functional group in the position of the C562-1101 sulfanamide group. An optimal addition at this location will be pursued using a fragment based drug design approach.

Appendix A: Botulinum Neurotoxin Serotype C (BoNT/C): Substrate Recognition and Cleavage Site Specificity

Introduction

There are several interesting characteristics of BoNT/C, that makes it unique among the BoNTs. First, BoNT/C is able to recognize and cleave two separate substrates, namely Syntaxin and SNAP-25.²⁸ At this time, it is unknown how BoNT/C achieves this dual substrate specificity. The current theory assumes that BoNT/C interacts with its substrate in a manner similar to BoNT/A (for which much more structural information is known),⁹⁴ however, this theory fails to account for all observations on BoNT/C to date. We hypothesize that there is an unidentified BoNT/C exosite that is able to recognize a specific sequence that is highly conserved between SNAP-25 and Syntaxin, which accounts for BoNT/C's ability to recognize both substrates. A second interesting feature of BoNT/C is that its SNAP-25 cleavage site is shifted a single amino acid from that of the BoNT/A cleavage site (Figure A.1).^{93,101} Furthermore, BoNT/C is able to distinguish between closely neighboring and identical lysine-alanine peptide bonds (Figure A.2).¹⁰¹ It is unlikely that such precision is due to active site effects alone. Therefore we hypothesize that this remarkable specificity is due not simply to active site architecture, but is largely driven by interactions at more distant exosites.

Our long-term goal is the discovery of broad-spectrum inhibitors to the BoNT's as well as to determine structural features that are amenable for toxin reengineering. The overall objective of this study is to (1) Identify the key BoNT/C-substrate interactions that are responsible for its cleavage site precision (2) Determine the structural basis for BoNT/C's dual substrate specificity and (3) Develop novel inhibitors that target BoNT/C.

Our central hypothesis is that exosite interactions are the driving force for BoNT/C's dual substrate selectivity and also serve as the anchor point for spatially positioning the substrate for proper cleavage. To accomplish our first two objectives, we will use x-ray crystallographic studies to solve both the BoNT/C-SNAP-25 and BoNT/C-Syntaxin structures. We will also use a combination of computational and biochemical techniques to probe for exosites on the surface of BoNT/C. Aim 3 will be accomplished using the rational drug design method incorporating multiple rounds of virtual screening, analog searches, binding assays, and finally crystallographic studies.

As shown in Figure A.1, both BoNT/A and BoNT/C cleave SNAP-25 close to its C-terminal end. Remarkably, both BoNT subtypes are highly specific for their particular cleavage site, with BoNT/A exclusively hydrolyzing the Gln197-Arg198 bond, while BoNT/C hydrolyzes only the Arg198-Ala199 bond. Understanding the reasons for this one amino acid shift should reveal much about the precise details of the BoNT cleavage mechanism. Currently, the structure of an inactive double mutant of BoNT/A in complex with a truncated form of SNAP-25 (residues 145-202) is known. Additionally, wild type BoNT/A in complex with two non-cleavable SNAP-25 analogs (sequences QRATKM and RRATKM) has also been successfully co-crystallized. Although we have much structural information about BoNT/A-substrate complexes, only the apo form of BoNT/C is known. This structure is inadequate for full comparison to the BoNT/A structures necessitating a co-crystal structure of BoNT/C with SNAP-25. Although no substrate bound structures of BoNT/C exist, analysis of the active site does provide some insights into why the BoNT/C cleavage site differs from that of BoNT/A. Comparisons of the catalytic regions of BoNT/A and BoNT/C reveal that the putative S1' position of

BoNT/C is much smaller than the S1' region of BoNT/A. It is hypothesized that size constraints at the S1' pocket simply will not accommodate the large bulky side chain of arginine. Instead, the substrate shifts by one residue allowing the much smaller alanine side chain to occupy the S1' site.⁹⁴ Of further importance, mutation studies have demonstrated that a positive residue (either Arginine or lysine) must occupy the P1 site on SNAP-25 or Syntaxin in order for BoNT/C to show full activity.⁹³ This contrasts with the other BoNT subtypes, which do not appear to have any dependence on the P1 residue. Clearly there are multiple active site interactions required for the precision of BoNT/C.

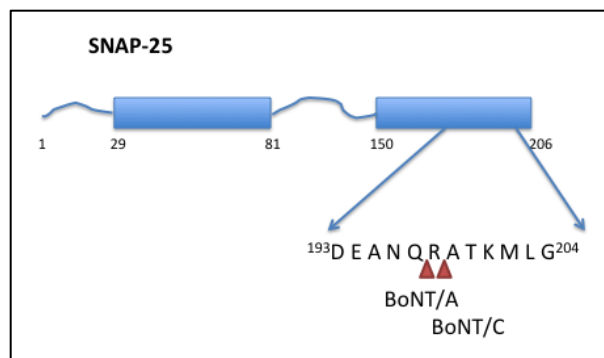


Figure A.1: Schematic drawing of SNAP-25. α -helical regions are depicted as rectangles. The red triangles indicate the cleavage sites for BoNT/A and BoNT/C. Figure adapted from Binz et al.¹⁰²

However, the complexity of the BoNT/C-substrate interaction does not reside solely in the active site, and is additionally affected by more distant exosite interactions. As shown in Figure 1.8, BoNT/C cleaves Syntaxin at the Lys253-Ala254 bond, but never at the identical and closely neighboring Lys260-Ala261 bond (Fig. A.2, red circle).¹⁰¹ Due to the close proximity of these bonds, it would be expected that either one could be

cleaved by BoNT/C. The fact that this does not occur suggests that Syntaxin is physically restrained, possibly by exosite binding, and this ultimately determines which residues are accessible for cleavage at the catalytic site.

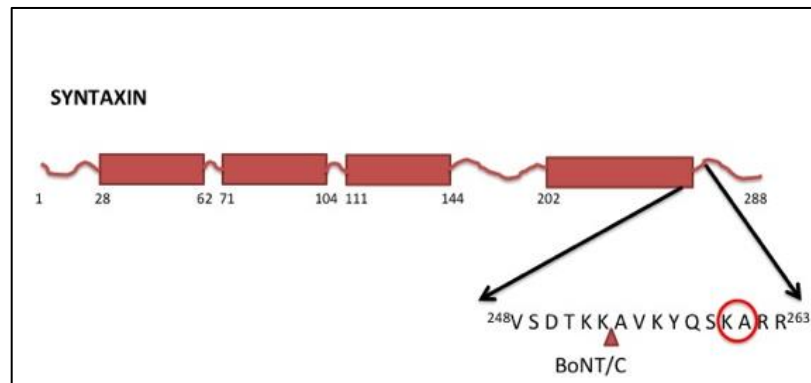


Figure A.2 Schematic drawing of Syntaxin with α -helical regions depicted as rectangles. The cleavage site is depicted by a red triangle. The red circle highlights the closely neighboring Lys-Ala peptide bond. Figure adapted from Binz et al.¹⁰²

Another interesting difference between BoNT/A and BoNT/C is that they each require different length SNAP-25 constructs for maximum enzyme activity. Specifically, BoNT/A requires a construct consisting of residues 141-202, while BoNT/C requires a significantly longer construct of residues 92-202.⁹³

As demonstrated by the BoNT/A-SNAP-25 complex structure, the α -exosite is occupied by an α -helical region of SNAP-25. This region is composed of residues 146 to 167 and contains a domain known as a SNARE motif.^{32, 103} The SNARE motifs are conserved regions found on all three members of the SNARE Complex (VAMP, Syntaxin, and SNAP-25). These motifs have an α -helical secondary structure with 1/3 of the helix being occupied by negatively charged residues and 1/3 being occupied by

hydrophobic residues. The final 1/3 is of variable amino acid constituency.¹⁰³ These regions are responsible for assembly of the ternary SNARE complex from its constituent parts and are the primary areas of association between these three individual proteins.¹⁰⁴ It was previously hypothesized that the BoNTs may also exploit these conserved regions during the process of substrate recognition. To test this hypothesis, short peptides composed of individual SNARE motifs were synthesized and tested for their ability to inhibit BoNT activity. It was found that several of these SNARE motif peptides had strong cross reactivity, as demonstrated by their ability to inhibit the activity of several BoNT subtypes.

Figure A.3a shows the locations of the SNARE motifs on syntaxin and SNAP-25. For both SNAP-25 and Syntaxin, the relative distance in amino acids between the cleavage site and the nearest SNARE motif is also depicted. Finally, a bracket marks the distance from the cleavage site to the N-terminal end of the optimum length BoNT/C SNAP-25 construct (residues 92-202). As previously mentioned, the BoNT/A-SNAP-25 structure shows that the α -exosite interacts strongly with a region of SNAP-25 containing a SNARE domain. Inspection of figure A.3a reveals that this is the S4 SNARE domain. The current theory as to how BoNT/C achieves dual substrate specificity assumes that a region on BoNT/C similar to the BoNT/A α -exosite recognizes the SNAP-25 S4 SNARE domain. Additionally, it has been postulated that the syntaxin region, ¹⁹⁷TRHSEIHKLEN²⁰⁷, binds to this same exosite, thus explaining BoNT/C's dual substrate specificity.⁹⁴ There are three major problems with this theory. First, BoNT/C cleavage studies of different length SNAP-25 constructs demonstrated that cleavage activity was unchanged when comparing a SNAP-25 construct containing residues 146-

202 with one that removed the S4 domain (residues 156-202). This suggests that the S4 domain may not be an important binding region for BoNT/C. A second problem with this theory is that it does not address the reason why BoNT/C requires a longer SNAP-25 construct than BoNT/A for maximum activity. Lastly, although the Syntaxin region, ¹⁹⁷TRHSEIIKLEN²⁰⁷, contains many similar residues to the S4 domain (¹⁴⁵EMDENLEQVSG¹⁵⁵), there is not a single conserved residue between these two sequences.

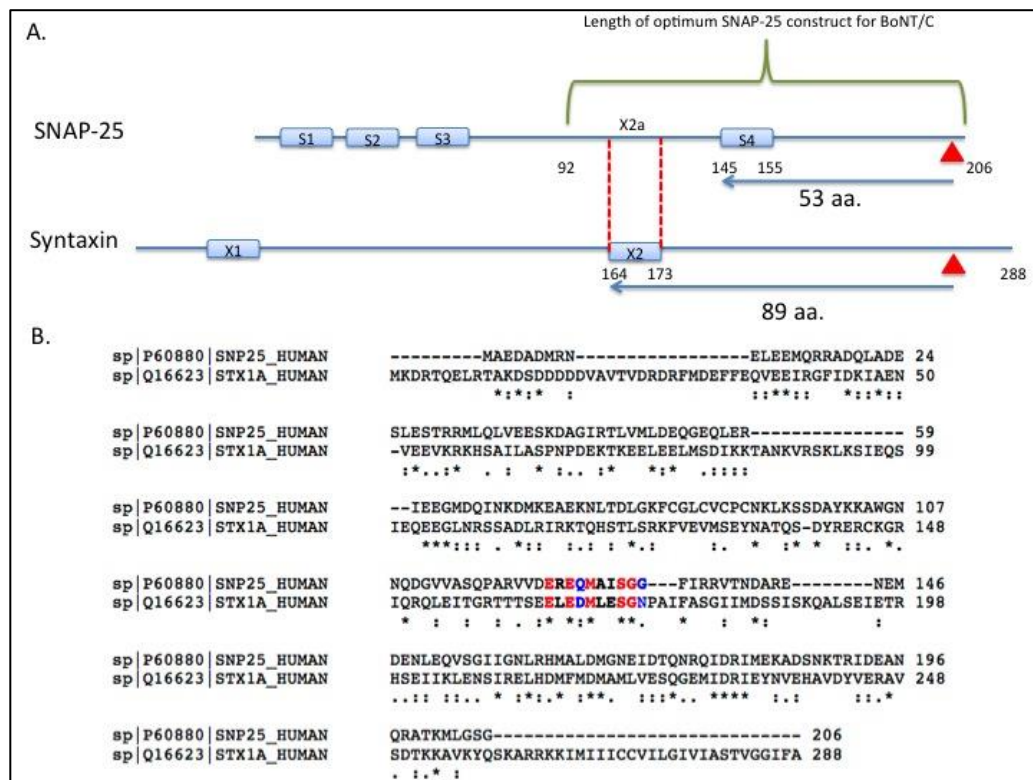


Figure A.3: (a) SNAP-25 and Syntaxin with SNARE regions depicted as blue rectangles and cleavage sites as red triangles. Shown with arrows is the distance in amino acids from the point of cleavage to the start of the SNARE domain. Green brackets show the length of the optimum SNAP-25 construct for BoNT/C (see text). Red dashed lines map the X2 domain to the complementary region on SNAP-25. (b) Alignment of SNAP-25 and syntaxin with X2 domain and complementary region bolded. Red colored amino acids are conserved, blue colored amino acids are similar.

We propose an alternative theory of BoNT/C dual substrate specificity whereby both SNAP-25 and Syntaxin bind a BoNT/C exosite in a region removed from the homologous BoNT/A α -exosite. Superimposing SNAP-25 and Syntaxin demonstrates that the additional amino acids required for optimal BoNT/C activity extend approximately the number of amino acids necessary to line up spatially with the Syntaxin X2 SNARE motif. Figure A.3b depicts a sequence alignment between Syntaxin and SNAP-25. Importantly, the region of SNAP-25 that is expected to spatially overlap with the Syntaxin X2 domain also contains a region of high sequence identity to the X2 domain (which we refer to as the X2a domain) composed of the sequence ¹²¹EREQMAISGG¹³². Sequence comparison of this domain and the X2 domain shows 50% identity, with two additional similar amino acids. We hypothesize that the X2 and X2a domains exploit a yet unknown exosite on BoNT/C and this may explain BoNT/C's dual substrate selectivity. Furthermore, this theory explains why BoNT/C and BoNT/A require different length SNAP-25 constructs for maximum activity.

Results

Cloning, expression, and purification of BoNT/C and SNAP-25 constructs. Three separate constructs of BoNT/C were cloned for crystallographic studies. The first construct consists of residues 1-430 of wild type BoNT/C. The second construct is of an inactive mutant BoNT/C containing an E to Q point mutation at residue position 230. Lastly, we created a BoNT/C triple Mutant (E230Q, M221A, and Y375F), in the event that the single mutant maintained residual activity. Using established protocols from the Swaminathan lab⁹¹, high-level expression in *E. Coli* of all three BoNT/C forms was

achieved via the auto-induction method and purified using size exclusion chromatography.

In addition, we successfully cloned, overexpressed, and purified six SNAP-25 constructs: (1) Full length SNAP-25. (2) SNAP-25 residues 145-202, which coincides with the start of the S4 SNARE domain and is the identical construct to that which has been successfully co-crystallized with BoNT/A (3) SNAP-25 residue, 121-202, which begins with our identified X2a domain. (4) SNAP-25 residues 92-202, which is the previously identified optimal length for maximum BoNT/C cleavage efficiency. (5) SNAP-25 residues 145-202 (A199C) mutant. (6) SNAP-25 residues 145-202 (A199R) mutant. The mutant SNAP-25 constructs are non-cleavable and were used for co-crystallization trials with wild-type BoNT/C. Table A.1 summarizes the various primers used for making BoNT/C and SNAP-25 constructs. BoNT/C wild-type, BoNT/C E230Q, full length SNAP-25, and SNAP-25 residues 145-202, were previously available in the lab and are not listed.

Table A.1: Primers for design of BoNT/C and SNAP-25 constructs.

Name	Primer
Syntaxin Forward	5'-GGTGGTGGTGCTCGAGTTACATGATCTTCTTCCTGC-3'
Syntaxin Reverse	5'-CGCGCGGCAGCCATATGACGACCAGTGAGGAGTTG-3'
Snap25 Reverse	5'-GGTGGTGGTGCTCGAGTTAACCACTCCAGCATCT-3'
Snap25.aa92_Forward	5'-CGCGCGGCAGCCATATGAACAAGCTTAAATCCAGTGA-3'
Snap25.aa121 Forward	5'-CGCGCGGCAGCCATATGCGTGTGGTGGATGAACGGGA-3'
BotC.M221A Forward	5'-CGGATCCAATACTAATTTTAA-3'
BotC.M221 Reverse	5'-CGCAAAATTCAGACTTAGAAA-3'
BotC.Y375F Forward	5'-TTCTTTCAAATGTATATACTC-3'
BotC.Y375F Reverse	5'-ATATTTTCTATTTTGTACA-3'
Snap25.A199C Forward	5'-GCACAAAGATGCTGGGAAGTGG-3'
Snap25.A199C Reverse	5'-AACGTTGGTTGGCTTCATCAA-3'
Snap25.A199R Forward	5'-GTACAAAGATGCTGGGAAGTGG-3'
Snap25.A199R Reverse	5'-GACGTTGGTTGGCTTCATCAA-3'

Crystallographic Studies. The currently published crystal structure of BoNT/C light chain was grown in a high salt crystallization solution containing 1.6M Sodium Formate and 0.1M Sodium Citrate (pH 4.6-5.0).⁹⁴ Our first goal was to discover a crystallization condition with a low salt content, which is more suitable for co-crystallization studies. To this end, we ran initial crystallization screens using screen

kits from Hampton Research. After further optimization of initial screen hits, large crystals were grown after 24 hours at 4°C by vapor diffusion in a 1:1 (v/v) ratio of protein and reservoir solution containing 0.5 mM Cadmium Chloride, 0.04 M citric acid/0.06 M BIS-TRIS Propane pH 6.4, and 20% PEG 3350. The crystals were flash frozen in liquid nitrogen using 20% glycerol as a cryoprotectant. Diffraction data were collected at the X29 Beamline of the National Synchrotron Light Source. The crystals diffracted to 1.9 Å and belonged to space group P3₂21 with unit cell dimension a=b=96.5, c=152.8, $\gamma=120$ with one molecule per asymmetric unit. This structure, shown in Figure A.4, is similar to the currently reported BoNT/C structure and contains a clear active site amenable for co-crystallization studies. A cadmium ion, in place of the typical zinc ion, was also present in the active site.

We were also able to crystallize the BoNT/C E230Q mutant. Using the crystallization solution discovered for the wild type form of BoNT/C as a starting point, we were able to grow crystals after 48 hours at 4°C by vapor diffusion in a 1:1 (v/v) ratio of protein and reservoir solution containing 0.5 mM Cadmium Chloride, 20mM ATP, 0.04M citric acid/ 0.06 M BIS-TRIS Propane pH 6.4, 5% glycerol, and 20% PEG 3350. These crystals diffracted to 1.8 Å with unit cell dimensions a=55.93, b=54.06, c=74.34, $\alpha=\gamma=90$ $\beta=104.2$ and belonged to space group P2₁ with one molecule per asymmetric unit.

Building on the initial crystals we obtained, multiple co-crystallization trials were attempted with a variety of different substrates. None of these trials resulted in a successful complex and consisted only of the BoNT/C molecule. However, BoNT/C and BoNT/C E230Q were crystallized in a variety of different space groups as summarized in

Table A.2. Comparison of these various structures showed changes in solvent accessible loops; however, the overall architectures were highly similar.

Table A.2: Summary of BoNT/C light chain crystallographic data.

Experiment	Crystallization	Data Collection	Resolution	Structural Solution	Refinement
BoNT/C – RRGC	0.5 mM Cadmium Chloride, 0.04 M citric acid/0.06 M BIS-TRIS Propane pH 6.4, 20% PEG 3350	P2 ₁ a=55.7,b=54.3,c=76.2 (Å) α=γ=90 β=104 (°)	2.0Å	MolRep (3DEB)	R factor=0.26 R free = 0.29
BoNT/C-M - 17MER	0.5 mM Cadmium Chloride, 20mM ATP, 0.04M citric acid/ 0.06 M BIS-TRIS Propane pH 6.4, 5% glycerol, 20% PEG 3350	P3 ₂ 21 a=b=96.4,c=153 α=β=90 γ=120	2.5Å	MolRep (3DEB)	
BoNT/C-M-SNAP25 (Full Length)	0.1M Sodium Malonate pH 5.0, PEG 3350, 30% v/v Methanol	P3 ₂ 21 a=b=96.6,c=153.5 α=β=90 γ=120	2.4Å	MolRep (3DEB)	R Factor= 0.23 R Free = 0.27
BoNT/C-M-SNAP25 (93-202)	0.2M Sodium phosphate dibasic dihydrate, 20% PEG 3350, 5% n-Octyl-β-D-glucoside	P6 ₅ 22 a=b=97.7,c=261.6 α=β=90 γ=120	2.8Å	MolRep (3DEB)	R Factor= 0.23 R Free = 0.29
BoNT/C-M-SNAP25 (146-202)	4% v/v Tacsimate pH 8.0, 12% PEG 3350, 0.1M Nickel Chloride hexahydrate	P2 ₁ a=55.9,b=54.0,c=74.3 α=γ=90 β=104.2	1.8Å	MolRep (3DEB)	
BoNT/C-SNAP25 (117-202)	1% w/v tryptone, 0.05M HEPES Sodium 7.0, 12% PEG 3350, 5% w/v n-Dodecyl-β-D-maltoside	C222 ₁	2.4Å	MolRep (3DEB)	
BoNT/E-M-SNAP25 (163-202)	0.1 M TRIS hydrochloride pH 8.5 2.0 M Ammonium phosphate monobasic 30% v/v ethano	P222 ₁ a=82.0,b=92.0,c=146.0	3.5Å	MolRep (3DEB)	

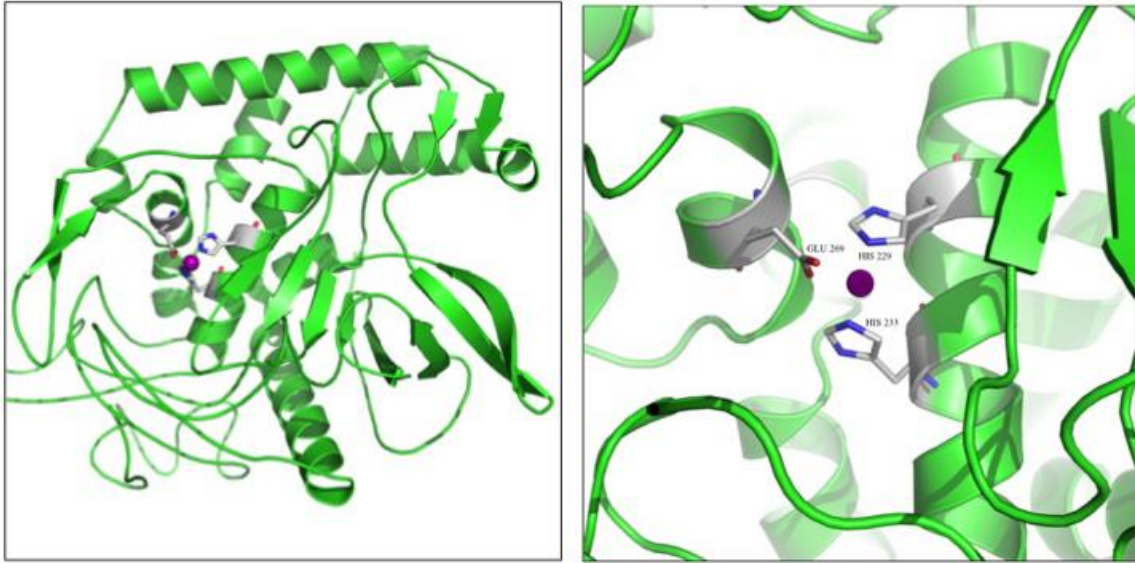


Figure A.4: Structure of BoNT/C light chain. (A) Ribbon diagram of BoNT/C. (B) Close up of BoNT/C active site with catalytic residues shown as sticks.

Appendix B: Cloning, Purification and Inhibition Studies of β -hydroxyacyl-ACP dehydrase from *Burkholderia pseudomallei* (BpFabZ) and Additional Inhibitor Studies.

Results

Cloning of β -hydroxyacyl-ACP dehydrase from *Burkholderia pseudomallei*. The BpFabZ gene was directly amplified from genomic DNA from strain Bp82, an avirulent *B. Pseudomallei* (Δ purM mutant), deficient in adenine biosynthesis, provided as a gift from the Tonge Lab. The primers used are summarized in table B.1.

Table B.1: Primers used for cloning of BpFabZ.

Name	Primer
BpFabZ Forward	5'-CGCGCGGCAGCCATATGCAGTCATCACCGCGCAGCCA-3'
BpFabZ Reverse	5'-GGTGGTGGTGCTCGAGTCAGGGCGCCGCGTCGGCCGTC-3'

Purification of β -hydroxyacyl-ACP dehydrase from *Burkholderia pseudomallei*.

During initial attempts at purification, BpFabZ was found to form insoluble inclusion bodies. Efforts to solubilize BpFabZ through altered expression conditions such as lower temperatures, were unsuccessful. Additionally, BpFabZ remained insoluble in up to Octyl β -D-glucopyranoside or 5% *n*-Dodecyl β -D-maltoside and the use of 2% sodium lauroyl sarcosinate resulted in inactive protein. As a result, a protein denaturing-refolding strategy was conducted, summarized as follows:

Inclusion bodies containing *BpFabZ* protein were solubilized in buffer containing 6M Urea, 500mM NaCl, and TRIS buffer pH 8.0 and applied to a 5mL HisTrap HP column (GE Healthcare) previously equilibrated with the solubilizing buffer. The HisTrap column was placed in an AKTA protein purification system and the solubilizing buffer

was replaced with a refolding buffer containing 20% glycerol, 2 mM DTT, 200 mM NaCl, and 20mM HEPES pH 7.5 using a linear gradient over the course of 2 hours. After refolding, *BpFabZ* protein was eluted from the column using a step gradient of imidazole (Fig. B.1). Proper refolding of *BpFabZ* protein was confirmed via testing of enzymatic activity.

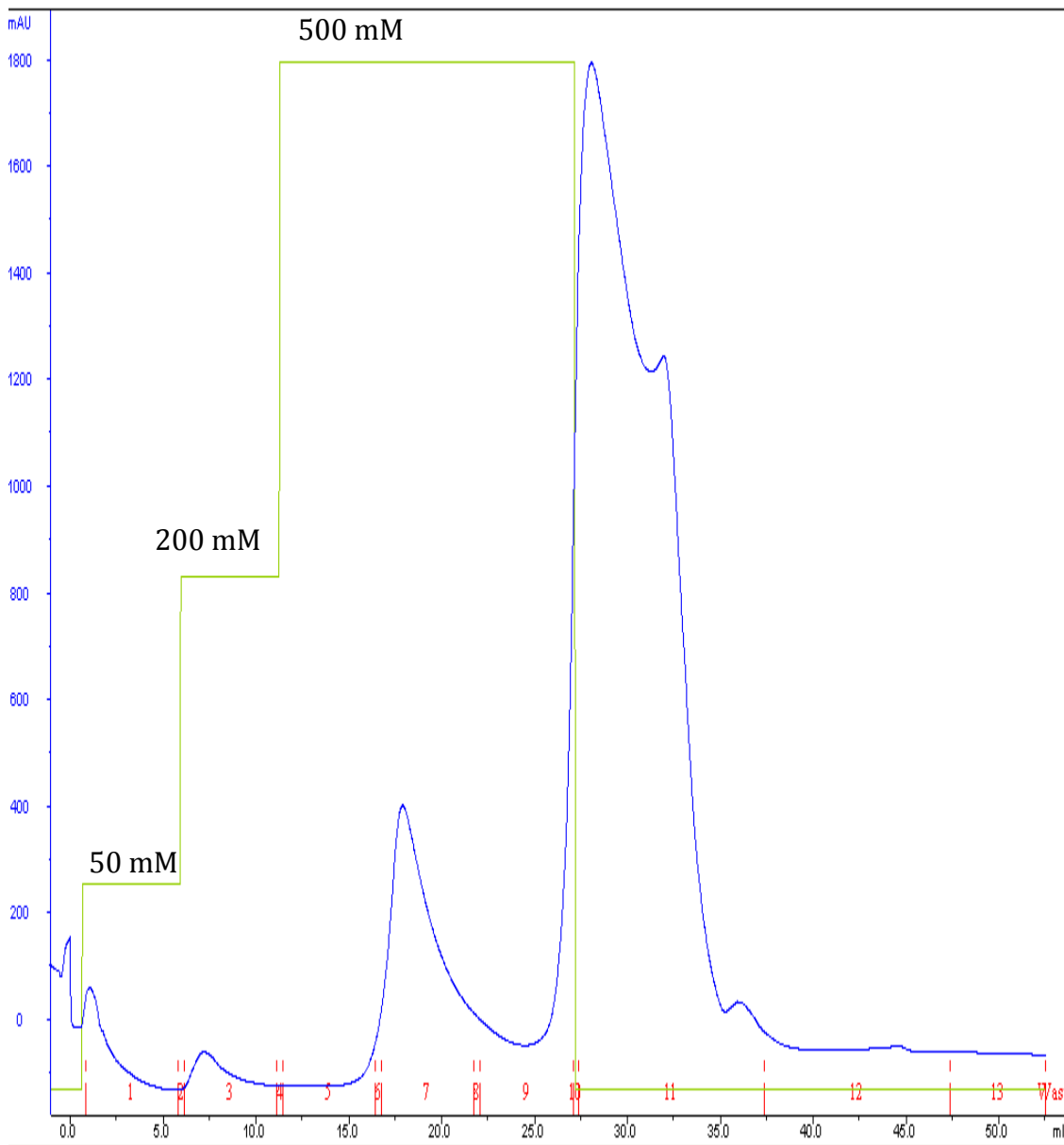


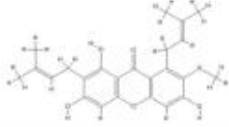

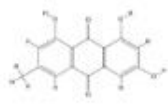
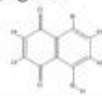
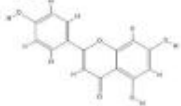
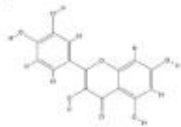
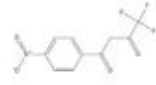
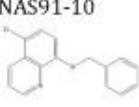
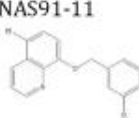
Figure B.1: Chromatogram depicting imidazole elution of BpFabZ. Absorbance is shown as a blue line, with absorbance values labeled on the Y-axis; Imidazole concentrations are shown as a yellow line and are labeled at each step increase. The x-axis shows total volume in black, with each 1.5 mL fraction labeled in red. Two major BpFabZ fractions were recovered with fraction 7 demonstrating activity in enzymatic assays. Fraction 11 contained protein aggregates and was inactive.

Preliminary inhibition data for select compounds against YpFabZ, FtFabZ, and BpFabZ. In addition to the inhibition studies detailed in chapter 2, the natural

compounds Emodin, Juglone, Stictic Acid, Mangostin, Apigenin, and Quercetin as well as the NAS series of synthetic compounds (NAS21, NAS91-10, and NAS91-11) were also tested against BpFabZ. Additionally, the NAS series of compounds were tested against YpFabZ and FtFabZ.

Although soluble and active BpFabZ was obtained through our refolding strategy, attempts to fully remove imidazole after elution from the HisTrap column have resulted in protein precipitation. The presence of imidazole and/or contamination with partially folded or unfolded BpFabZ has led to unacceptable levels of variability in enzymatic assays. For this reason, IC_{50} data for BpFabZ is considered preliminary. Additionally, only single concentration runs of all three NAS compounds have been conducted and careful evaluation of their potencies is not yet complete. These preliminary results are summarized in Table B.2.

Table B.2: Inhibitory activities of select compounds against YpFabZ, FtFabZ, and BpFabZ. N/O = none observed.

	YpFabZ	FtFabZ	BpFabZ
<p>Mangostin</p> 	6.1 ± 1.4 μM	7.7 ± 2.0 μM	> 50 μM
<p>Stictic Acid</p> 	13.0 ± 1.4 μM	27.8 ± 6.1 μM	38.0 ± 13.0 μM
<p>Emodin</p> 	29.7 ± 6.0 μM	43.1 ± 9.2 μM	177.0 ± 67.0 μM
<p>Juglone</p> 	5.3 ± 1.0 μM	5.4 ± 1.4 μM	1.5 ± 0.42 μM
<p>Apigenin</p> 	N/O	N/O	N/O
<p>Quercetin</p> 	N/O	N/O	N/O
<p>NAS-21</p> 	N/O	N/O	N/O
<p>NAS91-10</p> 	N/O	N/O	N/O
<p>NAS91-11</p> 	N/O	N/O	N/O

Bibliography

1. Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; et al. Rational design of peptide-based HIV proteinase inhibitors. *Science* **1990**, 248, 358-61.
2. Anderson, A. C. The process of structure-based drug design. *Chem Biol* **2003**, 10, 787-97.
3. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **2009**, 30, 2785-91.
4. Chirala, S. S.; Huang, W. Y.; Jayakumar, A.; Sakai, K.; Wakil, S. J. Animal fatty acid synthase: functional mapping and cloning and expression of the domain I constituent activities. *Proc Natl Acad Sci U S A* **1997**, 94, 5588-93.
5. Waller, R. F.; Keeling, P. J.; Donald, R. G.; Striepen, B.; Handman, E.; Lang-Unnasch, N.; Cowman, A. F.; Besra, G. S.; Roos, D. S.; McFadden, G. I. Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **1998**, 95, 12352-7.
6. Campbell, J. W.; Cronan, J. E., Jr. Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. *Annu Rev Microbiol* **2001**, 55, 305-32.
7. Payne, D. J.; Warren, P. V.; Holmes, D. J.; Ji, Y.; Lonsdale, J. T. Bacterial fatty acid biosynthesis: a genomics-driven target for antibacterial drug discovery. *Drug Discov Today* **2001**, 6, 537-544.
8. Heath, R. J.; Rock, C. O. Fatty acid biosynthesis as a target for novel antibacterials. *Curr Opin Investig Drugs* **2004**, 5, 146-53.
9. Heath, R. J.; White, S. W.; Rock, C. O. Lipid biosynthesis as a target for antibacterial agents. *Prog Lipid Res* **2001**, 40, 467-97.
10. Cronan, J. E., Jr.; Waldrop, G. L. Multi-subunit acetyl-CoA carboxylases. *Prog Lipid Res* **2002**, 41, 407-35.
11. Marrakchi, H.; Zhang, Y. M.; Rock, C. O. Mechanistic diversity and regulation of Type II fatty acid synthesis. *Biochem Soc Trans* **2002**, 30, 1050-5.
12. White, S. W.; Zheng, J.; Zhang, Y. M.; Rock, C. O. The structural biology of type II fatty acid biosynthesis. *Annu Rev Biochem* **2005**, 74, 791-831.
13. Price, A. C.; Zhang, Y. M.; Rock, C. O.; White, S. W. Structure of beta-ketoacyl-[acyl carrier protein] reductase from *Escherichia coli*: negative cooperativity and its structural basis. *Biochemistry* **2001**, 40, 12772-81.
14. Leesong, M.; Henderson, B. S.; Gillig, J. R.; Schwab, J. M.; Smith, J. L. Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* **1996**, 4, 253-64.
15. Sharma, S. K.; Kapoor, M.; Ramya, T. N.; Kumar, S.; Kumar, G.; Modak, R.; Sharma, S.; Surolia, N.; Surolia, A. Identification, characterization, and inhibition of *Plasmodium falciparum* beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ). *J Biol Chem* **2003**, 278, 45661-71.
16. Swarnamukhi, P. L.; Sharma, S. K.; Padala, P.; Surolia, N.; Surolia, A.; Suguna, K. Packing and loop-structure variations in non-isomorphous crystals of FabZ from *Plasmodium falciparum*. *Acta Crystallogr D Biol Crystallogr* **2007**, 63, 458-64.

17. Kostrewa, D.; Winkler, F. K.; Folkers, G.; Scapozza, L.; Perozzo, R. The crystal structure of PfFabZ, the unique beta-hydroxyacyl-ACP dehydratase involved in fatty acid biosynthesis of *Plasmodium falciparum*. *Protein Sci* **2005**, *14*, 1570-80.
18. Maity, K.; Venkata, B. S.; Kapoor, N.; Surolia, N.; Surolia, A.; Suguna, K. Structural basis for the functional and inhibitory mechanisms of beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) of *Plasmodium falciparum*. *J Struct Biol* **2011**, *176*, 238-49.
19. Sivaraman, S.; Sullivan, T. J.; Johnson, F.; Novichenok, P.; Cui, G.; Simmerling, C.; Tonge, P. J. Inhibition of the bacterial enoyl reductase FabI by triclosan: a structure-reactivity analysis of FabI inhibition by triclosan analogues. *J Med Chem* **2004**, *47*, 509-18.
20. Zhang, L.; Liu, W.; Hu, T.; Du, L.; Luo, C.; Chen, K.; Shen, X.; Jiang, H. Structural basis for catalytic and inhibitory mechanisms of beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ). *J Biol Chem* **2008**, *283*, 5370-9.
21. Zhang, L.; Kong, Y.; Wu, D.; Zhang, H.; Wu, J.; Chen, J.; Ding, J.; Hu, L.; Jiang, H.; Shen, X. Three flavonoids targeting the beta-hydroxyacyl-acyl carrier protein dehydratase from *Helicobacter pylori*: crystal structure characterization with enzymatic inhibition assay. *Protein Sci* **2008**, *17*, 1971-8.
22. Chen, J.; Zhang, L.; Zhang, Y.; Zhang, H.; Du, J.; Ding, J.; Guo, Y.; Jiang, H.; Shen, X. Emodin targets the beta-hydroxyacyl-acyl carrier protein dehydratase from *Helicobacter pylori*: enzymatic inhibition assay with crystal structural and thermodynamic characterization. *BMC Microbiol* **2009**, *9*, 91.
23. Kong, Y. H.; Zhang, L.; Yang, Z. Y.; Han, C.; Hu, L. H.; Jiang, H. L.; Shen, X. Natural product juglone targets three key enzymes from *Helicobacter pylori*: inhibition assay with crystal structure characterization. *Acta Pharmacol Sin* **2008**, *29*, 870-6.
24. Montecucco, C.; Molgo, J. Botulinum neurotoxins: revival of an old killer. *Curr Opin Pharmacol* **2005**, *5*, 274-9.
25. Cordivari, C.; Misra, V. P.; Catania, S.; Lees, A. J. New therapeutic indications for botulinum toxins. *Mov Disord* **2004**, *19* Suppl 8, S157-61.
26. Huttner, W. B. Cell biology. Snappy exocytosis. *Nature* **1993**, *365*, 104-5.
27. Turton, K.; Chaddock, J. A.; Acharya, K. R. Botulinum and tetanus neurotoxins: structure, function and therapeutic utility. *Trends Biochem Sci* **2002**, *27*, 552-8.
28. Schiavo, G.; Matteoli, M.; Montecucco, C. Neurotoxins affecting neuroexocytosis. *Physiol Rev* **2000**, *80*, 717-66.
29. Aa, A.; E, T.; Tv, I. Clinical Management of Potential Bioterrorism-Related Conditions. *Pediatr Infect Dis J* **2015**.
30. Lacy, D. B.; Tepp, W.; Cohen, A. C.; DasGupta, B. R.; Stevens, R. C. Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nat Struct Biol* **1998**, *5*, 898-902.
31. Kumaran, D.; Rawat, R.; Ludivico, M. L.; Ahmed, S. A.; Swaminathan, S. Structure- and substrate-based inhibitor design for *Clostridium botulinum* neurotoxin serotype A. *J Biol Chem* **2008**, *283*, 18883-91.
32. Breidenbach, M. A.; Brunger, A. T. Substrate recognition strategy for botulinum neurotoxin serotype A. *Nature* **2004**, *432*, 925-9.

33. Schmidt, J. J.; Stafford, R. G.; Bostian, K. A. Type A botulinum neurotoxin proteolytic activity: development of competitive inhibitors and implications for substrate specificity at the S1' binding subsite. *FEBS Lett* **1998**, 435, 61-4.
34. Schmidt, J. J.; Stafford, R. G. A high-affinity competitive inhibitor of type A botulinum neurotoxin protease activity. *FEBS Lett* **2002**, 532, 423-6.
35. Silvaggi, N. R.; Wilson, D.; Tzipori, S.; Allen, K. N. Catalytic features of the botulinum neurotoxin A light chain revealed by high resolution structure of an inhibitory peptide complex. *Biochemistry* **2008**, 47, 5736-45.
36. Zuniga, J. E.; Hammill, J. T.; Drory, O.; Nuss, J. E.; Burnett, J. C.; Gussio, R.; Wipf, P.; Bavari, S.; Brunger, A. T. Iterative structure-based peptide-like inhibitor design against the botulinum neurotoxin serotype A. *PLoS One* **2010**, 5, e11378.
37. Kumar, G.; Kumaran, D.; Ahmed, S. A.; Swaminathan, S. Peptide inhibitors of botulinum neurotoxin serotype A: design, inhibition, cocrystal structures, structure-activity relationship and pharmacophore modeling. *Acta Crystallogr D Biol Crystallogr* **2012**, 68, 511-20.
38. Hale, M.; Oyler, G.; Swaminathan, S.; Ahmed, S. A. Basic tetrapeptides as potent intracellular inhibitors of type A botulinum neurotoxin protease activity. *J Biol Chem* **2011**, 286, 1802-11.
39. Videnovic, M.; Opsenica, D. M.; Burnett, J. C.; Gomba, L.; Nuss, J. E.; Selakovic, Z.; Konstantinovic, J.; Krstic, M.; Segan, S.; Zlatovic, M.; Sciotti, R. J.; Bavari, S.; Solaja, B. A. Second generation steroidal 4-aminoquinolines are potent, dual-target inhibitors of the botulinum neurotoxin serotype A metalloprotease and *P. falciparum* malaria. *J Med Chem* **2014**, 57, 4134-53.
40. Adler, M.; Nicholson, J. D.; Cornille, F.; Hackley, B. E., Jr. Efficacy of a novel metalloprotease inhibitor on botulinum neurotoxin B activity. *FEBS Lett* **1998**, 429, 234-8.
41. Anne, C.; Turcaud, S.; Quancard, J.; Teffo, F.; Meudal, H.; Fournie-Zaluski, M. C.; Roques, B. P. Development of potent inhibitors of botulinum neurotoxin type B. *J Med Chem* **2003**, 46, 4648-56.
42. Anne, C.; Blommaert, A.; Turcaud, S.; Martin, A. S.; Meudal, H.; Roques, B. P. Thio-derived disulfides as potent inhibitors of botulinum neurotoxin type B: implications for zinc interaction. *Bioorg Med Chem* **2003**, 11, 4655-60.
43. Garcia, G. E.; Moorad, D. R.; Gordon, R. K. Buforin I, a natural peptide, inhibits botulinum neurotoxin B activity in vitro. *J Appl Toxicol* **1999**, 19 Suppl 1, S19-22.
44. Agarwal, R.; Swaminathan, S. SNAP-25 substrate peptide (residues 180-183) binds to but bypasses cleavage by catalytically active *Clostridium botulinum* neurotoxin E. *J Biol Chem* **2008**, 283, 25944-51.
45. Kumar, G.; Agarwal, R.; Swaminathan, S. Discovery of a fluorene class of compounds as inhibitors of botulinum neurotoxin serotype E by virtual screening. *Chem Commun (Camb)* **2012**, 48, 2412-4.
46. Schmidt, J. J.; Stafford, R. G. Botulinum neurotoxin serotype F: identification of substrate recognition requirements and development of inhibitors with low nanomolar affinity. *Biochemistry* **2005**, 44, 4067-73.
47. Agarwal, R.; Schmidt, J. J.; Stafford, R. G.; Swaminathan, S. Mode of VAMP substrate recognition and inhibition of *Clostridium botulinum* neurotoxin F. *Nat Struct Mol Biol* **2009**, 16, 789-94.

48. Kumar, G.; Swaminathan, S. Recent developments with metalloprotease inhibitor class of drug candidates for botulinum neurotoxins. *Curr Top Med Chem* **2015**, *15*, 685-95.
49. Kumaran, D.; Rawat, R.; Ahmed, S. A.; Swaminathan, S. Substrate binding mode and its implication on drug design for botulinum neurotoxin A. *PLoS Pathog* **2008**, *4*, e1000165.
50. Zuniga, J. E.; Schmidt, J. J.; Fenn, T.; Burnett, J. C.; Arac, D.; Gussio, R.; Stafford, R. G.; Badie, S. S.; Bavari, S.; Brunger, A. T. A potent peptidomimetic inhibitor of botulinum neurotoxin serotype A has a very different conformation than SNAP-25 substrate. *Structure* **2008**, *16*, 1588-97.
51. Boldt, G. E.; Kennedy, J. P.; Janda, K. D. Identification of a potent botulinum neurotoxin a protease inhibitor using in situ lead identification chemistry. *Org Lett* **2006**, *8*, 1729-32.
52. Cai, S.; Lindo, P.; Park, J. B.; Vasa, K.; Singh, B. R. The identification and biochemical characterization of drug-like compounds that inhibit botulinum neurotoxin serotype A endopeptidase activity. *Toxicon* **2010**, *55*, 818-26.
53. Opsenica, I.; Burnett, J. C.; Gussio, R.; Opsenica, D.; Todorovic, N.; Lanteri, C. A.; Sciotti, R. J.; Gettayacamin, M.; Basilico, N.; Taramelli, D.; Nuss, J. E.; Wanner, L.; Panchal, R. G.; Solaja, B. A.; Bavari, S. A chemotype that inhibits three unrelated pathogenic targets: the botulinum neurotoxin serotype A light chain, P. falciparum malaria, and the Ebola filovirus. *J Med Chem* **2011**, *54*, 1157-69.
54. Moe, S. T.; Thompson, A. B.; Smith, G. M.; Fredenburg, R. A.; Stein, R. L.; Jacobson, A. R. Botulinum neurotoxin serotype A inhibitors: small-molecule mercaptoacetamide analogs. *Bioorg Med Chem* **2009**, *17*, 3072-9.
55. Li, B.; Pai, R.; Cardinale, S. C.; Butler, M. M.; Peet, N. P.; Moir, D. T.; Bavari, S.; Bowlin, T. L. Synthesis and biological evaluation of botulinum neurotoxin a protease inhibitors. *J Med Chem* **2010**, *53*, 2264-76.
56. Thompson, A. A.; Jiao, G. S.; Kim, S.; Thai, A.; Cregar-Hernandez, L.; Margosiak, S. A.; Johnson, A. T.; Han, G. W.; O'Malley, S.; Stevens, R. C. Structural characterization of three novel hydroxamate-based zinc chelating inhibitors of the Clostridium botulinum serotype A neurotoxin light chain metalloprotease reveals a compact binding site resulting from 60/70 loop flexibility. *Biochemistry* **2011**, *50*, 4019-28.
57. Ruthel, G.; Burnett, J. C.; Nuss, J. E.; Wanner, L. M.; Tressler, L. E.; Torres-Melendez, E.; Sandwick, S. J.; Retterer, C. J.; Bavari, S. Post-intoxication inhibition of botulinum neurotoxin serotype A within neurons by small-molecule, non-peptidic inhibitors. *Toxins (Basel)* **2011**, *3*, 207-17.
58. Caglic, D.; Krutein, M. C.; Bompiani, K. M.; Barlow, D. J.; Benoni, G.; Pelletier, J. C.; Reitz, A. B.; Lairson, L. L.; Houseknecht, K. L.; Smith, G. R.; Dickerson, T. J. Identification of clinically viable quinolinol inhibitors of botulinum neurotoxin A light chain. *J Med Chem* **2014**, *57*, 669-76.
59. Stura, E. A.; Le Roux, L.; Guitot, K.; Garcia, S.; Bregant, S.; Beau, F.; Vera, L.; Collet, G.; Ptchelkine, D.; Bakirci, H.; Dive, V. Structural framework for covalent inhibition of Clostridium botulinum neurotoxin A by targeting Cys165. *J Biol Chem* **2012**, *287*, 33607-14.

60. Hallis, B.; James, B. A.; Shone, C. C. Development of novel assays for botulinum type A and B neurotoxins based on their endopeptidase activities. *J Clin Microbiol* **1996**, *34*, 1934-8.
61. Keller, J. E.; Nowakowski, J. L.; Filbert, M. G.; Adler, M. Rapid microplate assay for monitoring botulinum neurotoxin B catalytic activity. *J Appl Toxicol* **1999**, *19* Suppl 1, S13-7.
62. Schmidt, J. J.; Bostian, K. A. Proteolysis of synthetic peptides by type A botulinum neurotoxin. *J Protein Chem* **1995**, *14*, 703-8.
63. Schmidt, J. J.; Bostian, K. A. Endoproteinase activity of type A botulinum neurotoxin: substrate requirements and activation by serum albumin. *J Protein Chem* **1997**, *16*, 19-26.
64. Ruge, D. R.; Dunning, F. M.; Piazza, T. M.; Molles, B. E.; Adler, M.; Zeytin, F. N.; Tucker, W. C. Detection of six serotypes of botulinum neurotoxin using fluorogenic reporters. *Anal Biochem* **2011**, *411*, 200-9.
65. Ouimet, T.; Duquesnoy, S.; Poras, H.; Fournie-Zaluski, M. C.; Roques, B. P. Comparison of fluorogenic peptide substrates PL50, SNAPtide, and BoTest A/E for BoNT/A detection and quantification: exosite binding confers high-assay sensitivity. *J Biomol Screen* **2013**, *18*, 726-35.
66. Boldt, G. E.; Kennedy, J. P.; Hixon, M. S.; McAllister, L. A.; Barbieri, J. T.; Tzipori, S.; Janda, K. D. Synthesis, characterization and development of a high-throughput methodology for the discovery of botulinum neurotoxin a inhibitors. *J Comb Chem* **2006**, *8*, 513-21.
67. *WHO Guidelines on Tularaemia*. World Health Organization: 2007.
68. Perry, R. D.; Fetherston, J. D. *Yersinia pestis*--etiologic agent of plague. *Clin Microbiol Rev* **1997**, *10*, 35-66.
69. Galimand, M.; Guiyoule, A.; Gerbaud, G.; Rasoamanana, B.; Chanteau, S.; Carniel, E.; Courvalin, P. Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid. *N Engl J Med* **1997**, *337*, 677-80.
70. Kimber, M. S.; Martin, F.; Lu, Y.; Houston, S.; Vedadi, M.; Dharamsi, A.; Fiebig, K. M.; Schmid, M.; Rock, C. O. The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from *Pseudomonas aeruginosa*. *J Biol Chem* **2004**, *279*, 52593-602.
71. Kirkpatrick, A. S.; Yokoyama, T.; Choi, K. J.; Yeo, H. J. *Campylobacter jejuni* fatty acid synthase II: structural and functional analysis of beta-hydroxyacyl-ACP dehydratase (FabZ). *Biochem Biophys Res Commun* **2009**, *380*, 407-12.
72. Bryan, C. M.; Bhandari, J.; Napuli, A. J.; Leibly, D. J.; Choi, R.; Kelley, A.; Van Voorhis, W. C.; Edwards, T. E.; Stewart, L. J. High-throughput protein production and purification at the Seattle Structural Genomics Center for Infectious Disease. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2011**, *67*, 1010-4.
73. He, L.; Zhang, L.; Liu, X.; Li, X.; Zheng, M.; Li, H.; Yu, K.; Chen, K.; Shen, X.; Jiang, H.; Liu, H. Discovering potent inhibitors against the beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) of *Helicobacter pylori*: structure-based design, synthesis, bioassay, and crystal structure determination. *J Med Chem* **2009**, *52*, 2465-81.
74. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **2005**, *41*, 207-34.

75. Otwinowski, Z.; Minor, W.; W Jr, C. C. Processing of X-ray diffraction data collected in oscillation mode. **1997**.
76. Vagin, A.; Teplyakov, A. MOLREP: an automated program for molecular replacement. *Journal of applied crystallography* **1997**, 30, 1022-1025.
77. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011**, 67, 235-42.
78. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **1997**, 53, 240-55.
79. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, 60, 2126-32.
80. Hofstein, H. A.; Feng, Y.; Anderson, V. E.; Tonge, P. J. Role of glutamate 144 and glutamate 164 in the catalytic mechanism of enoyl-CoA hydratase. *Biochemistry* **1999**, 38, 9508-16.
81. Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010**, 31, 455-61.
82. McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **2013**, 41, W597-600.
83. Liu, W.; Luo, C.; Han, C.; Peng, S.; Yang, Y.; Yue, J.; Shen, X.; Jiang, H. A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from *Helicobacter pylori*: Molecular cloning, enzymatic characterization, and structural modeling. *Biochem Biophys Res Commun* **2005**, 333, 1078-86.
84. Tasdemir, D.; Lack, G.; Brun, R.; Rüedi, P.; Scapozza, L.; Perozzo, R. Inhibition of *Plasmodium falciparum* Fatty Acid Biosynthesis: Evaluation of FabG, FabZ, and FabI as Drug Targets for Flavonoids. *Journal of medicinal chemistry* **2006**, 49, 3345-3353.
85. Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R.; Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, F.; O'Dowd, B.; McCammon, J. A.; Opella, S. J.; Tsai, S. C.; Burkart, M. D. Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* **2014**, 505, 427-31.
86. Anne, C.; Cornille, F.; Lenoir, C.; Roques, B. P. High-throughput fluorogenic assay for determination of botulinum type B neurotoxin protease activity. *Anal Biochem* **2001**, 291, 253-61.
87. Mangru, S.; Bentz, B. L.; Davis, T. J.; Desai, N.; Stabile, P. J.; Schmidt, J. J.; Millard, C. B.; Bavari, S.; Kodukula, K. Integrated bioassays in microfluidic devices: botulinum toxin assays. *J Biomol Screen* **2005**, 10, 788-94.
88. Schmidt, J. J.; Stafford, R. G.; Millard, C. B. High-throughput assays for botulinum neurotoxin proteolytic activity: serotypes A, B, D, and F. *Anal Biochem* **2001**, 296, 130-7.

89. Schmidt, J. J.; Stafford, R. G. Fluorogenic substrates for the protease activities of botulinum neurotoxins, serotypes A, B, and F. *Appl Environ Microbiol* **2003**, *69*, 297-303.
90. N. R. Shine, U. P. A., 504,006 B1, 2003.
91. Rawat, R.; Ashraf Ahmed, S.; Swaminathan, S. High level expression of the light chain of botulinum neurotoxin serotype C1 and an efficient HPLC assay to monitor its proteolytic activity. *Protein Expr Purif* **2008**, *60*, 165-9.
92. Hines, H. B.; Kim, A. D.; Stafford, R. G.; Badie, S. S.; Brueggeman, E. E.; Newman, D. J.; Schmidt, J. J. Use of a recombinant fluorescent substrate with cleavage sites for all botulinum neurotoxins in high-throughput screening of natural product extracts for inhibitors of serotypes A, B, and E. *Appl Environ Microbiol* **2008**, *74*, 653-9.
93. Vaidyanathan, V. V.; Yoshino, K.; Jahnz, M.; Dorries, C.; Bade, S.; Nauenburg, S.; Niemann, H.; Binz, T. Proteolysis of SNAP-25 isoforms by botulinum neurotoxin types A, C, and E: domains and amino acid residues controlling the formation of enzyme-substrate complexes and cleavage. *J Neurochem* **1999**, *72*, 327-37.
94. Jin, R.; Sikorra, S.; Stegmann, C. M.; Pich, A.; Binz, T.; Brunger, A. T. Structural and biochemical studies of botulinum neurotoxin serotype C1 light chain protease: implications for dual substrate specificity. *Biochemistry* **2007**, *46*, 10685-93.
95. Brewer, C. F.; Riehm, J. P. Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins. *Analytical Biochemistry* **1967**, *18*, 248-255.
96. Capkova, K.; Hixon, M. S.; McAllister, L. A.; Janda, K. D. Toward the discovery of potent inhibitors of botulinum neurotoxin A: development of a robust LC MS based assay operational from low to subnanomolar enzyme concentrations. *Chem Commun (Camb)* **2008**, 3525-7.
97. Chen, S.; Barbieri, J. T. Unique substrate recognition by botulinum neurotoxins serotypes A and E. *J Biol Chem* **2006**, *281*, 10906-11.
98. Brozell, S. R.; Mukherjee, S.; Balias, T. E.; Roe, D. R.; Case, D. A.; Rizzo, R. C. Evaluation of DOCK 6 as a pose generation and database enrichment tool. *J Comput Aided Mol Des* **2012**, *26*, 749-73.
99. Balias, T. E.; Mukherjee, S.; Rizzo, R. C. Implementation and evaluation of a docking-rescoring method using molecular footprint comparisons. *J Comput Chem* **2011**, *32*, 2273-89.
100. Balias, T. E.; Allen, W. J.; Mukherjee, S.; Rizzo, R. C. Grid-based molecular footprint comparison method for docking and de novo design: application to HIVgp41. *J Comput Chem* **2013**, *34*, 1226-40.
101. Schiavo, G.; Shone, C. C.; Bennett, M. K.; Scheller, R. H.; Montecucco, C. Botulinum neurotoxin type C cleaves a single Lys-Ala bond within the carboxyl-terminal region of syntaxins. *J Biol Chem* **1995**, *270*, 10566-70.
102. Binz, T.; Sikorra, S.; Mahrhold, S. Clostridial neurotoxins: mechanism of SNARE cleavage and outlook on potential substrate specificity reengineering. *Toxins (Basel)* **2010**, *2*, 665-82.
103. Rossetto, O.; Schiavo, G.; Montecucco, C.; Poulain, B.; Deloye, F.; Lozzi, L.; Shone, C. C. SNARE motif and neurotoxins. *Nature* **1994**, *372*, 415-6.

104. Weimbs, T.; Low, S. H.; Chapin, S. J.; Mostov, K. E.; Bucher, P.; Hofmann, K. A conserved domain is present in different families of vesicular fusion proteins: a new superfamily. *Proc Natl Acad Sci U S A* **1997**, 94, 3046-51.