Synthetic, Biochemical, X-ray Crystallographic, Computational and High-Throughput
Screening Approaches Toward Anthrax Toxin Lethal Factor Inhibition

DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Elbek Kadamovich Kurbanov

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Elizabeth A. Amin

October 2015

Acknowledgements

I would like to express my deepest gratitude to my advisor, Prof. Elizabeth A. Amin, for her support, guidance, and optimism. I am extremely grateful for her encouragement, patience, and understanding over the five years I spent in her lab. Without Prof. Amin, this dissertation would not be possible. I would also like to sincerely thank Prof. Rodney L. Johnson for his instruction and guidance, and for being so instrumental in helping me to acquire essential synthetic skills and knowledge. I am also very grateful to Prof. Barry C. Finzel for his productive collaboration, and his insightful comments and suggestions over the course of my PhD training. My interactions with him have nurtured my academic growth and his lab's research accomplishments have significantly improved this dissertation. I am also indebted to Prof. Donald G. Truhlar for his perceptive feedback, critiques, and instruction that greatly contributed to my early work and to this dissertation.

I would also like to thank the entire lethal factor project team. Especially, I want to thank Dr. Jon E. Hawkinson and Jonathan Solberg for developing the LF assays presented in this dissertation, and for testing my compounds and analyzing the data. I would like to thank Kim M. Maize, Dr. Teresa De La Mora-Rey and Dr. Todd W. Geders for X-ray crystallographic studies. For analytical support, I would like to thank Dr. Subhashree Francis as she performed the quality control analysis on all of my synthesized compounds. I would like to thank Dr. Michael A. Walters for useful discussions, including but not limited to, synthesis and PAINS molecules. I would also like to acknowledge the Amin group, Dr. Xia Zhang and Dr. Ting-Lan Chiu, for providing an enjoyable and fun environment to learn over the last five years.

Dedication

This thesis is dedicated to my grandmother, Yakutjan Masharipova, my parents, Kadam

Kurbanov and Lolakhon Masharipova, and Dr. Margaret E. Olson, my future wife.

Abstract

The lethal factor (LF) enzyme secreted by *Bacillus anthracis* is chiefly responsible for anthrax-related cytotoxicity. In this dissertation, I present the computational design, synthesis, biochemical testing, structural biology, and virtual and high-throughput screening approaches to identify binding requirements for LF inhibition. To this end, we designed ~50 novel compounds to probe design principles and structural requirements for LF. Specifically, in Chapters 2 and 3, computational, synthetic, biochemical and structural biology methods to explore the underinvestigated LF S2′ binding subsite are described. We discovered that LF domain 3 is very flexible and results in a relatively unconstrained S2′ binding site region. Additionally, we found that the S1′ subsite can undergo a novel conformational change resulting in a previously unreported tunnel region, which we term S1′*, that we expect can further be explored to design potent and selective LF inhibitors. Using this novel LF configuration, we virtually screened ~11 million drug-like compounds for activity against LF and have identified a novel compound that inhibits LF with an IC_{50} of 126 μM.

In the course of this work, we found that reliable representation of zinc and other transition metal centers in macromolecules is nontrivial, due to the complexity of the coordination environment and charge distribution at the catalytic center. In Chapter 7, I will present work on applying and optimizing quantum mechanical methods developed by the Truhlar group to accurately calculate bond dissociation energies at low computational cost for various representative Zn^{2+} and Cd^{2+} model systems. By analyzing errors, we developed a prescription for an optimal system fragmentation strategy for our models. With this scheme, we find that the EE-3B-CE method is able to reproduce 53 conventionally calculated bond energies with an average absolute error of

only 0.59 kcal/mol. Therefore, one could use the EE-3B-CE approximation to obtain accurate results for large systems and/or identify better parameters for Zn centers for use in virtual screening.

Finally, we present the results of a large-scale *in vitro* HTS campaign of ~250,000 small-molecules against LF. After extensive validation, involving secondary assays and hit synthesis we were able to prioritize a key lead for further prosecution.

Table of Contents

List of Tables

List of Figures

xv

List of Schemes

Chapter 1

INTRODUCTION

1.1 Anthrax

Despite the successful dissemination of *Bacillus anthracis* (*B. anthracis)* in 2001, which resulted in five American deaths, anthrax remains a significant threat to society because effective treatments remain elusive. *B. anthracis* is considered a serious bioterrorism threat because anthrax spores can be easily produced en masse, aerosolized, and these spores can remain in the environment for extended periods of time (up to 60 years).[1–4] The Centers for Disease Control (CDC) continuously categorizes *B. anthracis* as a Tier 1 Select Agent, a pathogen that poses great risk to national security, economy, and critical infrastructure, because it can be easily disseminated to cause high mortality rates.[5] In recent history, the former Soviet Union is rumored to have built aerosolization facilities for the large-scale production of *B. anthracis*.[3,6] As evidence, a 1979 outbreak of anthrax in Sverdlovsk, former Soviet Union, has been attributed to inhalation of anthrax spores accidentally released at a military microbiological facility, which killed 64 people.[7] Also, during WWII, the Japanese army used anthrax as a biological weapon against the Chinese in Manchuria in 1940.[8] It is also a serious concern that antibiotic-resistant strains of *B. anthracis* can be engineered and used in a bioterrorist attack. Strains resistant to penicillin, doxycycline, and ciprofloxacin, the frontline therapies for anthrax treatment, have been engineered *in vitro* by several groups.[9–12]

Anthrax is caused by *B. anthracis*, a Gram-positive, spore-forming, rod-shaped, facultatively anaerobic bacterium which affects both animals and humans. There are three main ways in which humans can contract anthrax: gastrointestinal (through the GI tract), cutaneous (through the skin), and inhalational (through the lungs). It has recently been reported that a fourth type of anthrax transmission, injection anthrax, is possible.

Injection anthrax appeared in heroin-drug users with mortality rate of 30%.[13] Gastrointestinal anthrax develops upon consumption of animal products infected with anthrax spores. If not treated in a timely manner, this type of anthrax results in a high mortality rate of >50%.[14–16] The most common type of anthrax transmission is cutaneous, which comprises 95% of all anthrax infections.[14,15] A person is infected with cutaneous anthrax when spores enter the body through broken skin. Even though it is rarely fatal, without treatment mortality rates can approach 20%.[3,17] Inhalation anthrax is the most fatal of all three modes of infection. If left untreated, the mortality rate can reach 90%.[18] Infection with inhalational anthrax occurs when the spores are inhaled. Symptoms vary depending on type of anthrax infection.

There are two virulence factors that are responsible for anthrax related cell death. These virulence factors are encoded in pXO1 and pXO2 plasmids and secreted by *B. anthracis*.[19] Both plasmids are required for full virulence.[20] pXO1 encodes anthrax toxin, which consists of edema factor (EF), lethal factor (LF) and a protective antigen (PA).[21,22] Edema factor and protective antigen form edema toxin, whereas lethal factor and protective antigen form lethal toxin. pXO2 encodes for antiphagocytic poly-γ-D-glutamic acid capsule.[23] The capsule prevents the bacterium from being digested by macrophages.[24,25] A strain that lacks pXO2 plasmid, called the Sterne strain, has been effectively used as a vaccine for animals in the US.[26–28]

1.2 Current Treatments and Their Limitations

1.2.1 Vaccines

The first vaccine against anthrax was developed in 1880 and used in livestock.[29] The first vaccine for human use was developed in 1954 as a cell-free filtrate precipitated

with aluminum potassium sulfate (alum).[30,31] Later, in the 1960s, the manufacturing procedure for the anthrax vaccine was modified resulting in BioThrax® (anthrax vaccine adsorbed). It is made by aluminum hydroxide precipitation of PA from filtrates of cultures of non-encapsulated Sterne anthrax strain V770-NP1-R.[32] Currently, it is the only prophylactic, pre-exposure anthrax vaccine approved in the US. It stimulates the host immune system to produce antibodies against PA, the protein responsible for the transportation of anthrax toxin components such as lethal factor and edema factor into cells.[33] Once PA is neutralized by antibodies, the toxic effects of lethal factor and edema factor are eliminated. There are no efficacy studies on BioThrax®; however, a 1954 BioThrax® predecessor vaccine prevented inhalational and cutaneous anthrax with an efficacy of 92.5% according to a 1962 study.[34] Primarily, BioThrax® is used in vaccinations of military and research personnel who are at high risk of anthrax exposure. Biothrax® usage is limited due to the following disadvantages of the vaccine: 1) Intensive vaccination schedule of 18 months duration followed by yearly boosters, 2) Possible severe allergic reactions (anaphylactic shock), 3) Vaccine is not licensed for post-exposure use, 4) Efficacy and safety have not been established in children and elderly people, 5) Vaccine may not protect all individuals who get the shots, especially people with immunodeficiency[35], 6) Must be stored at 2-8 °C otherwise unstable[36]. These disadvantages and perception by the public that anthrax is rare disease prevent widespread civilian immunizations.

1.2.2 Monoclonal and Polyclonal Antibodies

After the 2001 anthrax attacks, the Project Bioshield Act was signed into law and provided funding for the development of new therapies aimed at anthrax, especially monoclonal antibodies. It was recognized that new treatments were needed since

existing therapies were severely limited in treating inhalational anthrax.[3,37] Other treatment options include antibody-based therapeutics, which exhibit their own set of limitations. Antibody-based treatments belong to a well-established drug class that has a fairly high success rate for clinical approval, and are usually well tolerated by humans, but key caveats exist: they are very expensive due to the high cost of manufacturing and the often large doses required; rare but serious adverse effects have been reported; and antibodies can also display significant pharmacokinetic liabilities, limited tissue accessibility, and impaired interactions with the immune system.[38,39] Consistent with antibody-based target restriction to those on the surface or exterior of host cells, all anti-anthrax antibodies developed to date target the protective antigen in order to interfere with LF translocation into host cells. One of the first fully human monoclonal antibodies developed was Valortim. However, post-exposure studies performed on non-human primates (NHPs) showed that Valortim could only achieve 70% efficacy. In 2012, the Food and Drug Administration (FDA) approved raxibacumab (Abthrax) in combination with antibiotics for the treatment of inhalational anthrax. Raxibacumab is a recombinant human immunoglobulin G1λ monoclonal antibody that prevents binding of PA to cell surfaces. Despite achieving FDA approval, efficacy studies performed on *Cynomolgus* macaques and New Zealand white (NZW) rabbits showed that raxibacumab is inferior to the currently approved antimicrobials.[40,41]

More recently, the FDA approved Anthrasil, a purified human immunoglobulin G, in combination with antibiotics to treat patients with inhalational anthrax. Anthrasil's mechanism of action is similar to that of raxibacumab. Efficacy studies performed with *Cynomolgus* macaques and NZW rabbits demonstrated that even in combination with antimicrobials, complete protection of animals from *B. anthracis* by Anthrasil remains challenging.[42] Thus, for post-exposure anthrax treatment, there is still a key unmet need

for novel therapeutics that reliably and effectively protect against the effects of anthrax toxin. Given the critical role of LF in anthrax pathogenesis and lethality, the discovery of strategies to inhibit this enzyme is still the most promising approach to combat post-exposure anthrax.[43,44]

1.2.3. Antibiotics

The major limitation of relying on antibiotic treatment is that antibiotics have no effect on toxin itself and must be administered early in the disease cycle due to rapid exotoxin secretion and consequent host death. Unfortunately, early stage infection is also when diagnosis is the most difficult.[14] At the later stages of infection, antibiotics fail to fight anthrax infection as high levels of anthrax toxin have already been secreted systemically, causing fatal septicemia. This has prompted various groups to research new therapies that neutralize anthrax toxin. For example, out of the eleven people who were diagnosed with inhalational anthrax and treated after the 2001 anthrax attacks, only six survived. This demonstrates some of the limitations of antibiotic treatments of inhalational anthrax.[3] Antibiotics such as ciprofloxacin, penicillin, and doxycycline are effective against the causative agent *B. anthracis*, and approved by the FDA for the treatment of inhalational anthrax.

Ciprofloxacin inhibits bacterial proliferation by binding to the active site of DNA gyrase, an essential enzyme in DNA replication. Therefore, ciprofloxacin only stops bacterial cell division (bacteriostatic effect) and has no effect on exotoxin secretion. Penicillin, on the other hand, kills bacteria by inhibiting cell wall synthesis (bactericidal effect). Doxycycline is also bacteriostatic and stalls bacterial growth and division by inhibiting protein synthesis. There are some antibiotics such as sulfamethoxazole:trimethoprim (Bactrim®), cefuroxime, cefotaxime sodium, aztreonam,

and ceftazidime that cannot be used to treat anthrax due to complete resistance.[45–47] Another concern is patient compliance as a 2-month long antibiotic regimen for exposure prophylaxis is required. After the 2001 anthrax attacks, only 40% of the sampled postal workers in Washington reported full adherence and 18% had completely stopped taking antibiotics.[48] As anthrax toxin plays a crucial role in the pathogenesis and lethality of anthrax, the discovery of strategies to inhibit the exotoxin is a promising approach to treat post-exposure anthrax.[43,44]

1.3 Anthrax Toxin Lethal Factor

LF Structure. The lethality of anthrax results from the anthrax toxin, which is composed of three proteins: lethal factor (LF), a calmodulin-activated adenylate cyclase (edema factor; EF), and protective antigen (PA).[49] LF, an 89-kDa zinc metalloprotease is primarily responsible for anthrax pathogenesis. LF consists of four domains: the N-terminal domain (I), the large central domain (II), a small helical domain (III), and the C-terminal catalytic domain (IV) (**Figure 1.1**).[50] The N-terminal domain I (residues 1-263) binds to PA and is responsible for LF translocation into host cells. The functions of domains II (residues 264-297 and 385-550) and III (residues 303-382) are not fully understood; however, it is known that domain III plays an important role in LF substrate selectivity. The catalytic C-terminal domain (IV) (residues 552-776) contains the LF active site, which is defined by a catalytic Zn^{2+} coordinated to three active site residues: His686, His690, and Glu735. Residues His686, His690, and Glu687 form part of the signature Zn metalloproteinase HEXXH consensus motif that is characteristic of most matrix metalloproteinases (MMPs).

Figure 1.1. Anthrax toxin lethal factor. The N-terminal domain I is in dark blue, domain II in yellow, domain III in red, and domain IV in green with catalytic zinc in grey sphere (PDB ID 1JKY).[50] (Schrödinger Maestro Discovery Suite 9.4).
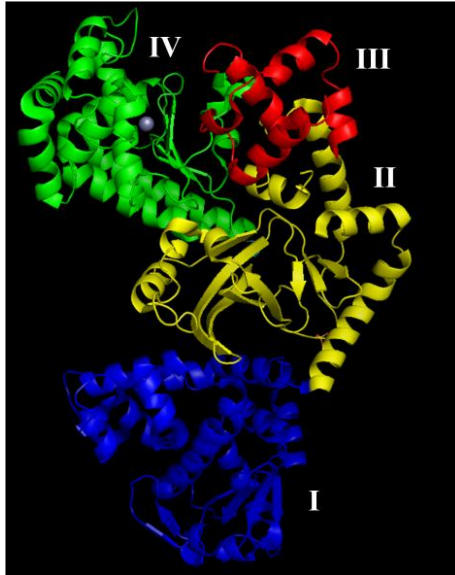
The LF active site consists of three subsites: (1) a primarily hydrophobic and sterically constrained S1′ subsite, (2) a mostly hydrophobic, but less sterically restricted S1-S2 region that is also a solvent-exposed tunnel, and (3) a poorly characterized S2′ area (**Figure 1.2**).[50,51]

LF Mechanism of Action. The first step in anthrax toxin-mediated pathogenesis is PA binding to cellular anthrax toxin receptor (ATR).[52] This binding event triggers proteolytic cleavage of PA by furin proteases to PA_{63}, which then undergoes heptamerization to form a prepore-like structure that binds three LF and/or EF units.[52–54] The LF/EF-bound complex is then endocytosed into a host target cell.[55–57] A low pH environment converts the prepore-like structure to a pore-like channel releasing EF and LF into the cytoplasm.[58–62] Once in the cytoplasm, LF cleaves mitogen-activated protein kinase kinases (MAPKKs) MEK1, MEK2, MKK3, MKK4, MKK6, and MKK7.[63] MAPKK degradation inhibits the phosphorylation of MAPKs, which effectively shuts down

cellular immune defense mechanisms.[55,63–72] In later stages of the disease, LF also invades endothelial cells causing disruption of endothelial barriers and leaky vasculature.[73–75] The exact mechanism by which LF cleaves MAPPKs is not well understood. However, current hypotheses are modeled on known mechanisms of similar catalytic zinc enzymes that contain the HEXXH signature consensus motif, such as the matrix metalloproteinases (MMPs). Based on the MMP mechanism of action, it is proposed that LF cleaves MAPKKs through activation of a zinc-coordinated water molecule.[51,76] Specifically, the activated water molecule, which is strongly H-bonded to Glu687 and Tyr728, attacks the scissile amide bond in the MAPPK substrate. Tyr728 plays a crucial role in this catalytic process by stabilizing the amino moiety of the leaving group, making it more susceptible to nucleophilic attack by the water molecule.[77–79]

Figure 1.2. Active site of anthrax toxin lethal factor (PDB ID 1YQY).[110] (PyMOL Molecular Graphics System, Version 1.7.4). Zn^{2+} is represented as a blue sphere, with Zn-coordinating residues indicated and LF subsites labeled.

9
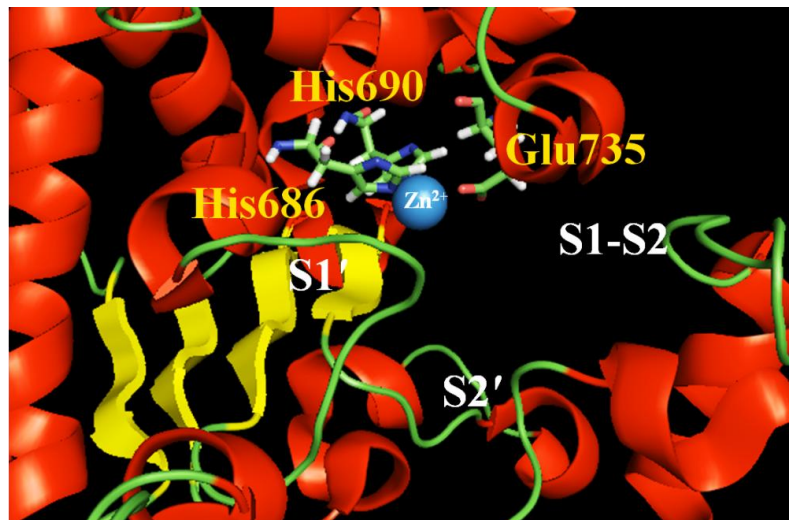
1.4 History of Lethal Factor Inhibitor Development

Numerous studies have been conducted toward the design of small molecule LF inhibitors.[53,80–88] The first reported LF inhibitors were small peptide sequences, designed as mimics of the natural MAPKK substrate, which were chemically linked to hydroxamic acid zinc-binding groups (ZBGs).[89–91] Substrate analog inhibitors were first designed by Montecucco and coworkers, and the best inhibitor from their efforts was compound **1.1** (**Figure 1.3**) with a $K_i \sim 1$ nM.[90] Following this discovery, Cantley and coworkers screened millions of peptide sequences against LF. This approach resulted in the discovery of micromolar inhibitor **GM6001** (Ilomastat, **Figure 1.3**).[91] X-ray co-crystal studies with LF showed that the isobutyl moiety of **GM6001** occupied the deep and wide S1′ subsite, while the Trp residue engaged the S2′ subsite (PDB ID 1PWU).[91] They concluded that binding to the S1′ subsite greatly contributes to potent LF inhibition.[91] By taking advantage of the deep S1′ subsite, Merck & Co. developed potent hydroxamic acid-based inhibitors of LF including **MK-702/LF1-B** with a $K_i \sim 24$ nM (**Figure 1.3**). More recently, Johnson and coworkers used **MK-702/LF1-B** to develop the most potent LF inhibitor to date with a K_i close to 40 pM (PT-8541, **Figure 1.3**).[92,93] However, development of hydroxamic acid-based inhibitors as therapeutic agents has been limited due to the poor selectivity, pharmacokinetic and toxicological liabilities of this moiety.[94–99] Hence, much attention has also been directed to develop non-hydroxamate LF inhibitors.[44,81–85,100–111] Reported inhibitors have included cationic polyamines,[83] aminoglycosides,[101] pyrazolones,[84] polyphenols,[107] tetracyclins,[102] α-defensins,[108] quinolines,[100] rhodanines,[103] and catechols.[44] The majority of these compounds exhibit micromolar activity against LF. Most notably, Pellecchia and coworkers developed several potent inhibitors of LF based on a rhodanine scaffold. Compound **1.2** was the

most active rhodanine-based inhibitor of this series with a half maximal inhibitory concentration (IC_{50}) of 190 nM (**Figure 1.3**).[81,82,103,109] X-ray studies with LF showed that the molecule primarily binds to the S1-S2 subsite, and the rhodanine heterocycle chelates zinc through the thiazolidinedione sulfur (PDB ID 1ZXV) (**1.2, Figure 1.3**).[78,79] These rhodanine-based inhibitors have limited therapeutic value because they do not target the crucial S1′ or S2′ subsites that are involved in native substrate binding. More recently, rhodanines were found to interfere with many biological assays gaining them a reputation as pan assay interference compounds (PAINS). Due to these problems, as well as a lack of selectivity, rhodanines were deemed non-optimizable for future LF inhibitor development.[112,113]

Given the liabilities of the current LF inhibitors, there is still a significant need for novel, non-hydroxamate LF inhibitors. The challenge remains to develop small-molecule inhibitors that would engage in crucial protein-ligand interactions, while displaying selectivity for LF over other metalloproteinases.

Figure 1.3. Published LF inhibitors with activities.

Ac-GY-bA-RRRRRRRRVLR-NHOH

1.1, Ki = 1 nM

GM6001, Ki = 2.1 mM

MK-702/LF-1B, Ki = 24 nM

PT-8541, Ki = 0.04 nM

1.2 IC_{50} = 190 nM
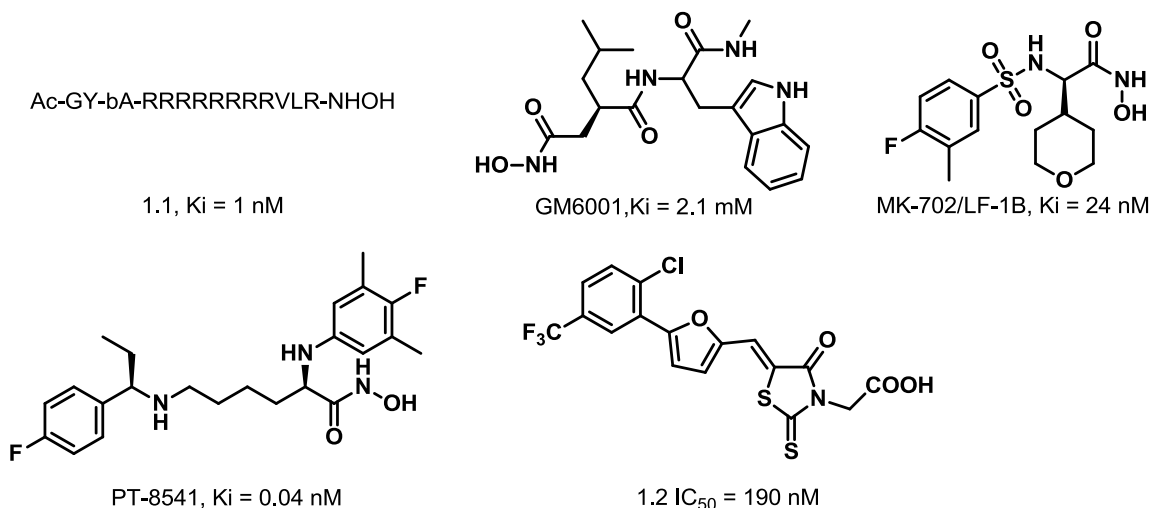
Chapter 2

PROBING THE S2′ SUBSITE OF THE ANTHRAX TOXIN LETHAL FACTOR USING N-ALKYLATED HYDROXAMATES

Adapted with permission from:

Maize, K. M.; Kurbanov, E. K.; De La Mora-Rey, T.; Geders, T. W.; Hwang, D.-J.; Walters, M. A.; Johnson, R. L.; Amin, E. A.; Finzel, B. C. Anthrax Toxin Lethal Factor Domain 3 Is Highly Mobile and Responsive to Ligand Binding. *Acta Crystallogr. Sect. D* **2014**, *70*, 2813–2822.

Kurbanov, E. K.; Chiu, T.-L.; Solberg, J.; Francis, S.; Maize, K. M; Fernandez, J.; Johnson, R. L.; Finzel, B. C.; Hawkinson, J. E.; Walters, M. A.; Amin, E. A. Probing the S2′ Subsite of the Anthrax Toxin Lethal Factor Using N-Alkylated Hydroxamates. *Submitted* **2015**.

2.1 ACKNOWLEDGEMENTS

2.2 INTRODUCTION

The LF active site consists of three subsites: a strongly hydrophobic and sterically constrained S1′ subsite, a mostly hydrophobic but less sterically restricted S1-S2 region that is also open-ended making a solvent-exposed tunnel, and a poorly characterized S2′ area (**Figure 2.1**). The S1′ and S1-S2 subsites were previously explored and characterized by several groups.[80,86] Our research objectives delineated in this and the following chapters were to characterize the S2′ subsite so that ultimately the resulting data might be exploited to optimize potent LF inhibitors based on novel non-hydroxamate scaffolds.

13

Figure 2.1. LF active site with catalytic Zn^{2+} (blue sphere) co-crystallized with **MK-702/LF1-B** (PDB ID 1YQY). (PyMOL Molecular Graphics System, Version 1.7.4).
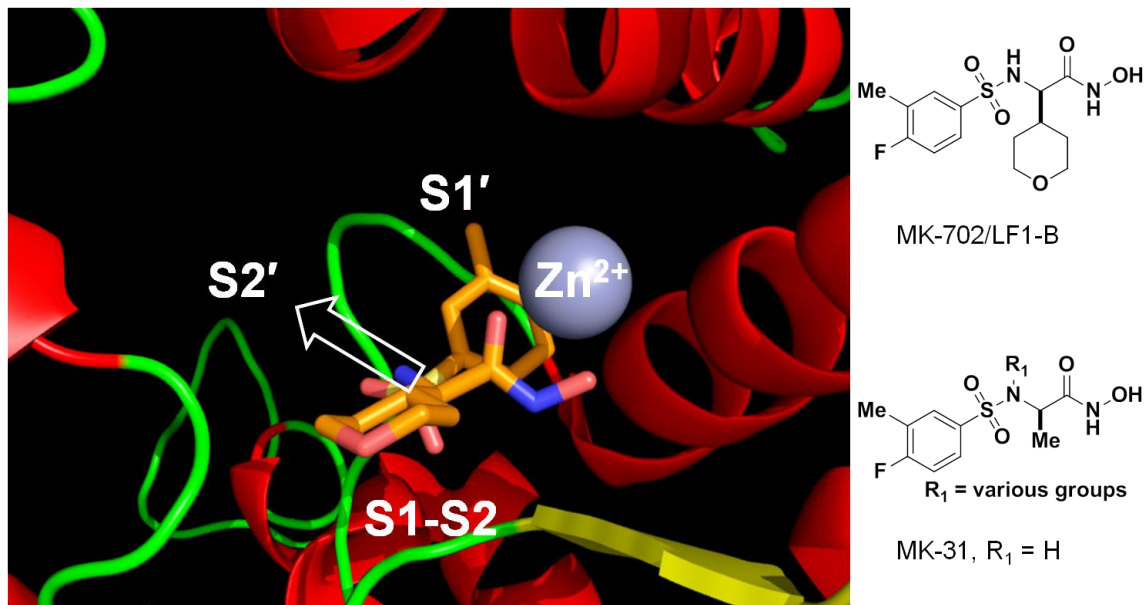
Upon examining the published X-ray structure of **MK-702/LF1-B** (**Figure 2.1**) co-crystallized with LF (PDB ID 1YQY),[110] we noted that the S2′ subsite could potentially be engaged and explored by modifying the ligand sulfonamide using various chemical functionalities (**Figure 2.1**). We ultimately selected **MK-31**,[80] an analog of **MK-702/LF1-B**, as our starting point for modification, because the tetrahydropyranyl group of **MK-702/LF1-B** does not engage in protein-ligand interactions and is not essential for S2′ exploration, and moreover, **MK-31** derivatives are more synthetically tractable than those of **MK-702/LF1-B** (**Figure 2.1**). We have also obtained the crystal structure of **MK-31** bound to LF (PDB ID 4WF6) and showed that **MK-702/LF1-B** and **MK-31** have similar binding modes with an RMSD of 0.47 Å (**Figure 2.2**). We therefore designed and synthesized a series of **MK-31** analogs that are functionalized at the sulfonamide N, thereby identifying residues in the S2′ pocket that can be engaged by small molecules. In so doing, we also found that the S2′ subsite is relatively accommodating to a variety of

14

chemical functionalities. The results presented in this and the following chapters outline how the S2′ subsite can be targeted to elucidate its binding requirements.

Figure 2.2. LF active site with co-crystallized **MK-31** (yellow) (PDB ID 4WF6) superimposed to co-crystallized **MK-702/LF1-B** (green) (PDB ID 1YQY). (Schrödinger Maestro Discovery Suite 9.4).
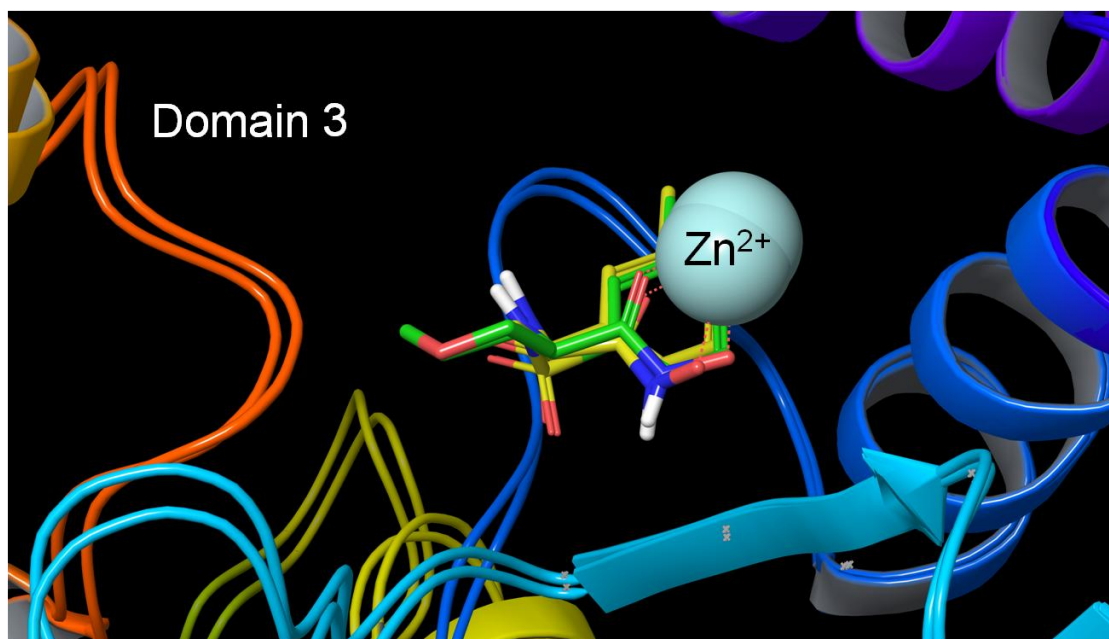
2.3 MATERIALS AND METHODS

2.3.1 Molecular Modeling

With 4WF6, we computationally docked and scored 300 N-alkylated analogs of **MK-31** (**Figure 2.1)** using Glide 5.9[114–117] with standard (SP) and then extra precision (XP) in the Schrödinger's Maestro Discovery Suite 9.4 (Schrödinger, Inc.).[118] The library of compounds utilized for the docking studies was designed from the following Sigma-Aldrich databases:[119] aliphatic bromides, benzyl bromides, and aliphatic iodides. Compounds with reactive moieties such as Michael acceptors, aldehydes, and cyanides were excluded. Interactive enumeration in the Maestro Suite was used to

computationally alkylate **MK-31** generating 300 diverse and synthesizable compounds. The structures were appropriately protonated at pH 7 using LigPrep[120] (Maestro), energy-minimized using the OPLS-AA force field,[121] and docked with the SP protocol into the crystal structure of 4WF6. Their scores were compared to that of **MK-31**. The protein for the docking studies was prepared using the protein preparation wizard[122,123] in Maestro which added hydrogens, removed co-crystallized waters, and added missing sidechains with Prime.[124] The compounds that scored higher than **MK-31** were re-docked using the XP protocol. The 26 compounds that scored higher than **MK-31** using the SP protocol were docked using the XP protocol. Of these 26 compounds, 19 were benzyl alkylated on the sulfonamide N. From these computational studies, we concluded that **MK-31** analogues with benzyl modifications were able to occupy the S2′ subsite, and were predicted to have higher predicted potencies than **MK-31**. Our observations could be rationalized by the presence of hydrophobic residues at the entrance of the S2′ subsite, such as Tyr728 and Val675, which likely interact with the benzyl group. Based on these studies, we decided to synthesize a series of benzyl analogs of **MK-31** (**Scheme 2.1**).

2.3.2 Synthesis

Synthetic modifications to the sulfonamide of **MK-31** were accomplished as outlined in **Scheme 2.1**. A generalized synthetic route was fashioned for all designed analogues as follows. Intermediate sulfonamide **2.3** was readily synthesized from commercially available D-alanine and 4-fluoro-3-methylphenyl-sulfonylchloride through nucleophilic substitution. Protection of the carboxylic acid was accomplished under Fisher esterification conditions to give **2.4** in 78% yield over the first two steps**.** Alkylation of sulfonamide **2.4** with the appropriate aryl bromides and chlorides was carried out

16

under basic conditions to afford tertiary sulfonamides **2.5-2.9** in 52-83% yield. The penultimate esters were converted to hydroxamic acids **2.10-2.14** using hydroxylamine hydrochloride and sodium methoxide in 28-73% yield.

Scheme 2.1 Synthetic modifications to the sulfonamide of **MK-31**.
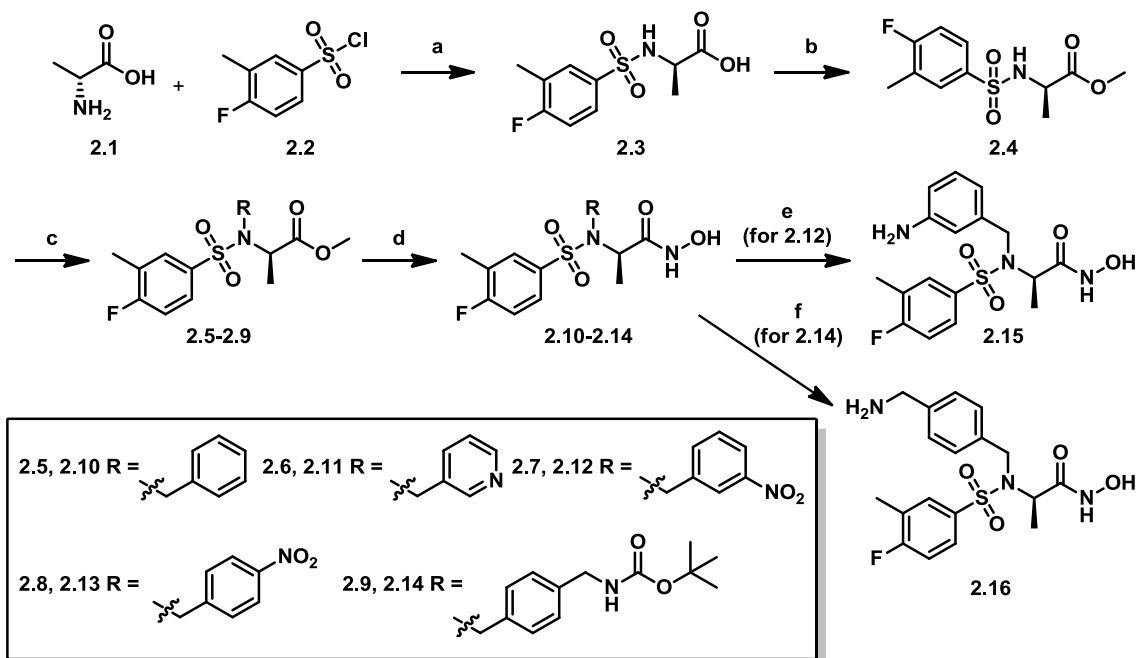
Reagents and conditions: (a) K$_2$CO$_3$, dioxane/H$_2$O (1:1), rt; (b) concd H$_2$SO$_4$, methanol, reflux (78% over two steps); (c) R-X, K$_2$CO$_3$, DMF, rt (**2.5**, 82%; **2.6**, 79%; **7**, 83%; **2.8**, 52%; **2.9**, 71%); (d) NH$_2$OH·HCl, NaOMe, methanol, 0 °C to rt (**2.10**, 54%;**2. 11**, 28%; **2.12**, 73%; **2.13**, 45%; **2.14**, 50%); (e) 10 wt. % Pd/C, DCM, rt (**2.15**, 82%); (f) 4M HCl in dioxane (**2.16**, 97%).

Further modification to **2.12** was pursued by reducing the *m*-nitrobenzyl substituent of the sulfonamide. The hydrogenation of **2.12**, to afford **2.15**, was accomplished in an 82% yield under a hydrogen atmosphere in the presence of palladium on activated carbon (Pd/C). Moreover, tert-butoxycarbonyl deprotection of analogue **2.14** was achieved with 4M HCl in dioxane to yield **2.16** as its HCl salt in 97% yield.

2.3.3 Biochemical Evaluation

2.3.3.1 FRET Lethal Factor Protease Assay

This work was completed by Jonathan Solberg and Jon E. Hawkinson at the Institute for Therapeutics Discovery and Development (ITDD) at the University of Minnesota.

The quenched fluorescence resonance energy transfer (FRET) assay measures LF enzymatic activity using an internally quenched peptide substrate derived from MAPKK.[111] LF cleavage separates the Dnp quencher from the oABz fluorescent tag, leading to an increase in fluorescence intensity.[83,110] In this assay, 10 μL of 100 nM anthrax toxin lethal factor[111] (final concentration 50 nM) in 2x assay buffer (40 mM HEPES containing 0.02% Triton X-100, pH 8.0) was added to 384-well assay plates (Corning #3677) using a MultiDrop dispensing instrument (Thermo-Fisher). The plate was pre-incubated at 37 °C for 15 min and the reaction was initiated by the addition of 10 μL of 14 μM oAbz/Dnp-labeled substrate (developed in-house at the Biomedical Genomics Center, University of Minnesota) in water (final substrate concentration 7 μM). The reaction was allowed to continue for 5 minutes at 37 °C, before being terminated with addition of 5 μL 50 mM EDTA (final concentration 10 mM). Fluorescence intensity was measured on a SpectraMax M2e microplate reader with excitation and emission wavelengths of 320 nm and 420 nm, respectively. For IC_{50} determinations, test compounds were dissolved in DMSO at 10 mM and varying volumes were added to the assay plate using a Labcyte Echo® 550 acoustic dispenser prior to LF addition to achieve 8 concentrations in duplicate (final DMSO concentration 1%). IC_{50} values were determined as a percent of control wells containing no inhibitor (following subtraction of background wells lacking enzyme) from at least 3 independent experiments using

GraphPad Prism software. IC_{50} values were also determined for the positive control compounds **MK-702/LF1-B** and **GM6001** for comparison, which were included on every plate.

2.3.4 Structural Biology

X-ray crystallography studies were done by Kimberly M. Maize, Teresa De La Mora-Rey, and Todd W. Geders under the supervision of Barry C. Finzel. Materials and methods for obtaining and processing X-ray crystal structures were previously published.[125]

2.3.4.1 Protein Purification

DNA encoding residues 265-776 (A266S) of *Bacillus anthracis* lethal factor (LF_{NT}) was cloned into pMCSG10[126–130] to produce a TEV-cleavable, N-terminal GST fusion bearing a His_6 tag. LF_{NT} was expressed using BL21(DE3) Rosetta2 pLysS cells. In 10 L scale, the cells were grown to an OD_{600} = 0.6-0.8 at 310 K, cooled to 303 K, induced with 0.2 mM IPTG for 6-8 hours and then harvested by centrifugation (15 minutes at 8,200 x *g*). Cell pellets were frozen at 253 K. Cell pellets were resuspended in 145 mL of 50 mM Tris pH 7.6, 500 mM NaCl, 10% glycerol and 1 mM DTT and lysed by sonication on ice. Lysozyme (1 mg/mL), benzonase (1 mU/mL) and $MgCl_2$ (1 mM) were added and stirred for 30 minutes on ice. Lysate was cleared by centrifugation at 40,000 x *g* for 45 minutes at 277 K and the supernatant was clarified by a 0.45 µm syringe filter prior to loading onto a 50 mL Ni-NTA column and eluted with lysis buffer containing 500 mM imidazole. Histidine-tagged tobacco etch virus (TEV) protease was added at 0.8% (w/w) and incubated at ambient temperature for 45 minutes followed by extensive dialysis overnight at 277 K against lysis buffer with 0.5 mM TCEP instead of 1 mM DTT. The

dialyzed material was passed through the Ni-NTA column and untagged LF$_{NT}$ in the flow-through was dialyzed extensively 25 mM HEPES pH 7.5 at 277 K. Light, flocculent white precipitate was isolated by centrifugation (15 min at 5000 x g) and resuspended in 50 mM Tris pH 7.6, 500 mM NaCl and 10% glycerol. The redissolved LF$_{NT}$ was applied to a HiPrep 26/60 Sephacryl S-200 HR column (GE Healthcare) equilibrated with 25 mM HEPES pH 7.5 and 150 mM NaCl and eluted as a single peak. LF$_{NT}$ was concentrated to A$_{280}$ = 25.7 and stored at 193 K. Yield was 25 mg from a 10 L batch.

2.4.3.2 Crystallization

Prior to crystallization, the protein was incubated with each compound of interest. In brief, the incubation solution (500 µL) consisted of 200 µM compound, 2 µM protein, and 10% DMSO in 25 mM HEPES pH 7.5, 150 mM NaCl. After incubation at room temperature for 30-45 min, the solution was filtered (0.22 µm) and concentrated to greater than 5 mg/mL.

Crystals were grown at 286 K using the hanging drop vapor diffusion method and microseeding to encourage the growth of fewer larger crystals. Crystallization drops consisted of post-incubation protein solution (2.0 µL), and either 2.0 µL well solution or well solution (1.5 µL) plus microseeding solution (0.5 µL). Well solutions that yielded crystals are 50 mM Bis-Tris pH 6.8, 100 mM magnesium acetate, and polyethylene glycol 8000 (PEG 8K, 11-16%). A microseeding solution was prepared by crushing crystals grown without seeding with a micropestle. Crystals appeared and grew to full size within a month. To harvest samples for data collection, crystals were quickly dipped in a 25% ethylene glycol-supplemented well solution, followed by flash vitrification in liquid nitrogen.

2.4.3.3 Crystallographic Data Collection and Processing

Diffraction data for structures 4PKQ, 4PKT, 4PKU, 4PKV, and 4PKW were collected from crystals at 100 K at beamline 17-ID-B (IMCA-CAT) using a Dectris Pilatus 6M Pixel Array Detector at the Advanced Photon Source of Argonne National Laboratories in Argonne, IL. The data were processed using XDS[131] and scaled with SCALA.[132]

For structures 4PKR and 4PKS, diffraction data were collected from crystals at 100 K using the NOIR-1 MBC detector at beamline 4.2.2 at the Advanced Light Source of Lawrence Berkeley National Laboratories in Berkeley, CA. The data were processed using d*TREK.[133]

The structures were solved using molecular replacement with atomic coordinates from structure 1YQY[134] and the program Phaser[135] in the CCP4 suite.[136] Both the Refmac5[137] and Phenix[138,139] programs were utilized for data refinement, along with the Coot modelling and visualization software.[140]

2.3.4.4 Protein Superposition
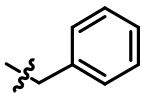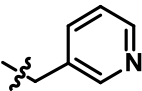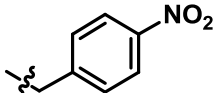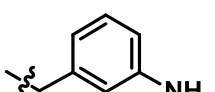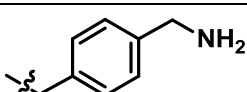
Non-isomorphous protein structures were aligned onto a common frame-of-reference using only a conserved core substructure comprised of two helical segments (residues 686-692 and 735-740) from reference structure 1YQY.[134] The segments include the Zn-coordinating histidines and glutamate. Locally centralized superposition of only this core substructure gives rise to a better alignment of the ligands[141] and simplifies recognition of that changes to protein quaternary structure relative to the fixed active site. Overlay method 'ATLF' has been shared by the Finzel group at

where web-based services exist to overlay any structures that share this core.[141]

2.4 RESULTS AND DISCUSSION

2.4.1 Biochemical Evaluation and Structural Biology

Table 2.1. LF FRET assay results for five N-benzylated analogs of **MK-31**.

GPHR #[a]	PDB ID	Cpd #	R	LF IC$_{50}$ (µM)[b]
00223332	4PKR	**2.10**		15.2 ± 0.5
00223341	4PKS	**2.11**		23.8 ± 0.9
00223590	4PKT	**2.13**		14.9 ± 2.4
00223588	4PKU	**2.15**		29.8 ± 0.7
00223596	4PKV	**2.16**		5.6 ± 0.3

[a]GPHR # is a compound number in our in-house collection of compounds
[b]IC$_{50}$ is a half maximal inhibitory concentration

Initially, benzyl and picolyl N-alkylated analogs of **MK-31** were synthesized and co-crystallized with LF (PDB IDs: 4PKR, 4PKS).[125] As can be seen from **Figure 2.3**, the binding mode of N-benzylated analog **2.10** (PDB ID 4PKR) is similar to the binding mode of **MK-31**. Specifically, the 4-fluoro-3-methylphenyl groups of **2.10** and **MK-31** occupy the S1′ site, whereas the hydroxamate groups of both chelate zinc. **Figure 2.3** also shows that the benzyl group of **2.10** occupies the S2′ subsite of LF, as predicted by

computer modeling. The N-picolyl analog **2.11** also binds in this fashion (PDB ID 4PKS). We observed that in order to accommodate the benzyl and picolyl groups, domain 3 undergoes a ligand-induced conformational change, which is depicted by the arrow in **Figure 2.3**. By combining this information with that obtained by previous crystal structures,[50,81,91,100] we observed that the proximity of domain 3 to domain 4, where domain 4 contains the active site, can vary greatly. We have classified these three different states of domain 3 as tight, open, and bioactive. The tight position is characterized by domain 3 being the closest to domain 4, and is observed in the crystal structure of LF with **MK-31** (colored green in **Figure 2.3**). An alternative domain 3 position can be induced by certain N-alkylated **MK-31** derivatives, where the proximity of domain 3 to domain 4 is intermediary to the tight and bioactive positions. This domain 3 position is termed the open position (colored magenta in **Figure 2.3**). Analogs **2.10** and **2.11** induce the open conformation of domain 3. The bioactive position is induced by binding of the peptide substrate.[125]

The advantage of promoting the open state of domain 3 is that the conformational change exposes crucial residues, such as Asp328, Lys380, and His654 to ligand binding, offering multiple novel sites to accomplish favorable charge-charge and H-bonding interactions. We employed 4PKR, our crystal structure with domain 3 in the open position, to design and synthesize analogs **2.13**, **2.15**, and **2.16**. Analogous to **MK-31**, analogs **2.13** and **2.16** induce the tight conformation of domain 3 (PDB IDs: 4PKT, 4PKV). Alternatively, analog **2.15** induces the open conformation of domain 3 (PDB ID 4PKU). Notably, the activities of **2.10** and **2.13** are similar while the analogs bind to different states of domain 3. Moreover, the activities of **2.13** and **2.16** are noticeably different while the analogs bind to the same state of domain 3. As a result of

23

these observations, we argue that binding to a specific state of domain 3 (tight vs open)

does not produce more potent inhibitors of LF.

Figure 2.**3**. LF active site with co-crystallized **MK-31** (green) (PDB ID 4WF6) superimposed to the co-crystallized N-benzyl analog **2.10** (yellow). (PDB ID 4PKR). The arrow denotes the conformational change that occurs upon binding **2.10**. (Schrödinger Maestro Discovery Suite 9.4).
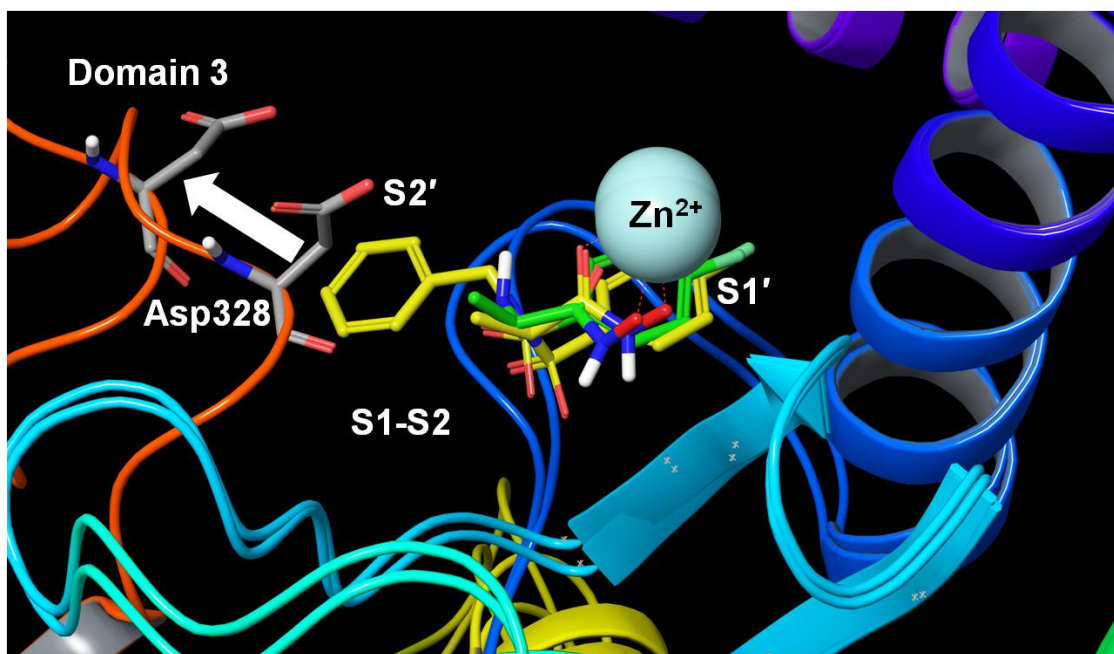
By analyzing the activities of **2.10**, **2.11**, **2.13**, **2.15**, and **2.16,** we observed that

the amino-containing analog **2.16** is more active than the rest. We analyzed the complex

4PKV, between LF and **2.16**, for further insights, but found no clear evidence for why

2.16 is more active than the other compounds. We were able, however, to answer this

question through molecular dynamics simulations discussed in the next chapter.

Additionally, we found interesting protein-ligand interactions between **2.13** and

2.15 and LF in the crystal structures 4PKT and 4PKU, accordingly. **Figure 2.4** shows

that the nitro group of **2.13** is engaged in electrostatic and H-bond interactions with the

sidechain of Lys335 in the S2′ site (see figure legend). However, the potency of **2.13** did not improve upon the potency of **2.10**. We argue that this can be due to several factors. The primary factor is that the S2′ site is solvent exposed. This, in turn, decreases the strength of the charge-charge interaction between **2.13** and the sidechain of Lys335. Moreover, desolvation and entropic penalties can also reduce the strength of this interaction. Taken together, these factors severely weaken the interactions of **2.13** with Lys335.

Figure **2**.4. Two-dimensional (2D) protein-ligand interaction map for **2.13** (PDB ID 4PKT). (Schrödinger Maestro Discovery Suite 9.4).
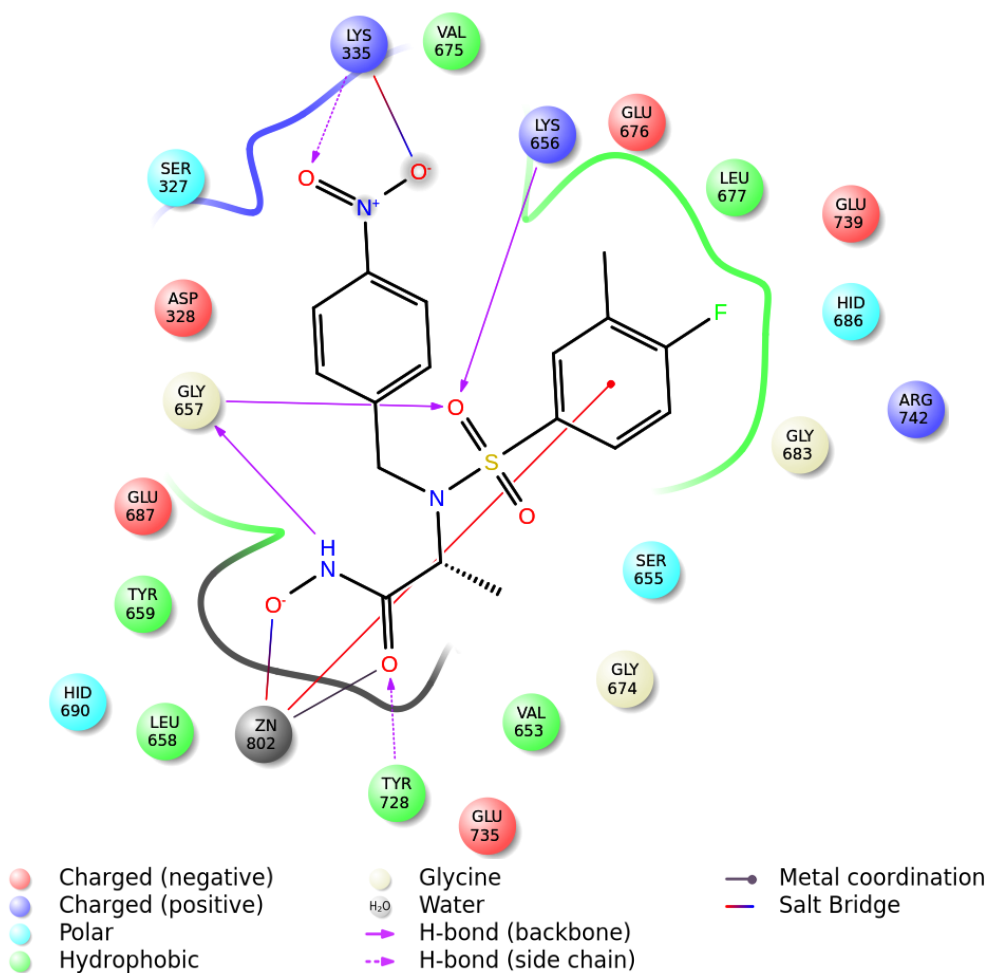
For **2.15**, we observe H-bond donating interactions between the aniline and the backbone of His654. However, the potency of **2.15** also did not improve upon the potency of **2.10**. We argue that the H-bond interaction loses strength due to desolvation and entropic penalties (**Figure 2.5**).

Figure 2.**5**. 2D protein-ligand interaction map for **2.15** (PDB ID 4PKU). (Schrödinger Maestro Discovery Suite 9.4).
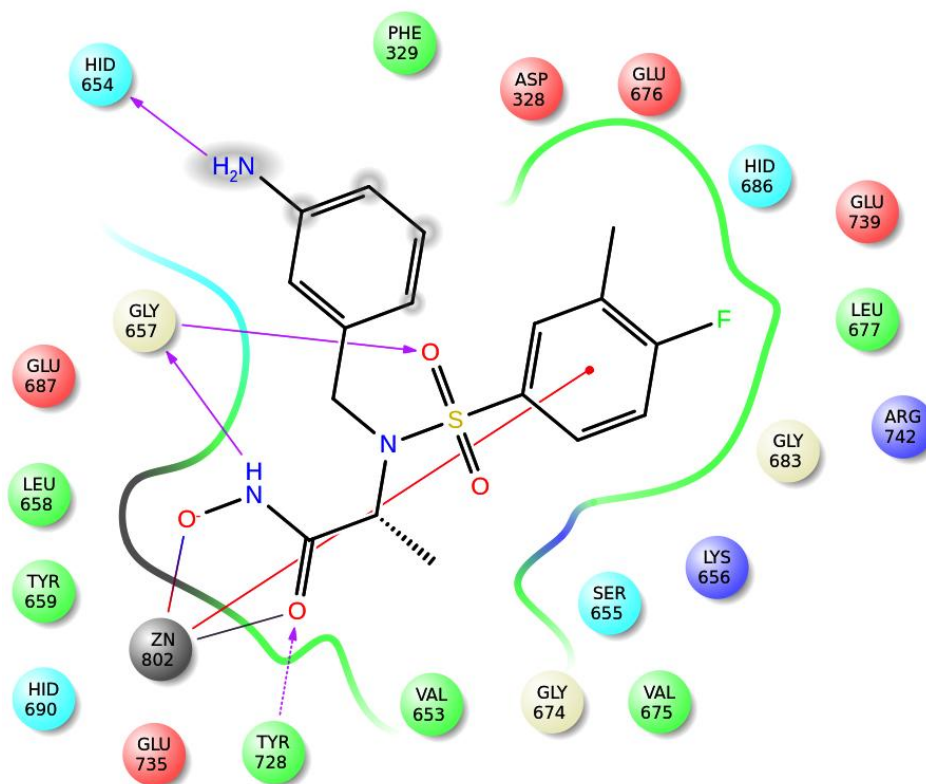
2.5 CONCLUDING REMARKS

In this study, we synthesized five N-benzylated analogs of **MK-31** to explore the S2′ site of LF. Co-crystal structures of these compounds with LF were obtained, which showed that domain 3 is highly flexible. By combining this information with that obtained by previous crystal structures,[50,81,91,100] we classified the frequently populated states of

26

domain 3 as tight, open, and bioactive. Biochemical evaluation and X-ray data for these compounds showed that binding to a specific state of domain 3 does not necessarily improve the potency of LF inhibition. However, we noted that we could design compounds that engage in additional favorable protein-ligand interactions with the newly exposed residues in the open state of domain 3. We designed **2.13**, **2.15**, and **2.16** to take advantage of crucial residues Asp328, Lys380, and His654. Surprisingly, **2.13** was found to engage Lys335. However, even though the nitro group of **2.13** was engaged in electrostatic and H-bond interactions with the sidechain of Lys335, the potency of **2.13** did not improve upon the potency of **2.10**. We argue that because the S2′ site is solvent exposed this decreases the strength of the charge-charge interaction between **2.13** and the sidechain of Lys335. For **2.15**, we observed H-bond donating interactions between the aniline and the backbone of His654. We argue that the potency of **2.15** did not improve upon the potency of **2.10** because the H-bond interaction loses strength due to desolvation and entropic penalties. In the following chapter, additional studies that thoroughly explore the S2′ site are presented.

2.6 EXPERIMENTAL

General Synthesis Information. Chemical reagents were purchased from commercial sources and used without additional purification. Bulk solvents were from Fisher Scientific and anhydrous *N,N'*-dimethylformamide (DMF) was purchased from EMD Chemicals. Reactions were performed under an atmosphere of dry N_2 unless otherwise noted. Silica gel chromatography was performed on self-packed columns with SiliaFlash 60Å silica gel (SiliCycle). Preparatory thin layer chromatography (TLC) was performed on plates with glass backed SiliaPlate 60Å silica gel (SiliCycle). Compounds used in biological testing were no less than 90% pure as determined by two-wavelength HPLC

analysis (254 and 215 nm). The reaction mixture/purified compound was analyzed on a Waters UPLC which is connected to ELSD, UV and ZQ mass spectrometers. The mass analysis was carried out in both ESI positive and ESI negative modes. Multiple wavelengths of 214 nm, 220 nm, 244 nm and 254 were used for UV detection. Elution was carried out using a gradient mobile phase of 95:5 of water: acetonitrile with 0.1% formic acid to 100% acetonitrile over 6 minutes. The purity determination was carried out using OpenLynx. Nuclear magnetic resonance (NMR) spectra were recorded in $CDCl_3$, CD_3OD, or DMSO-d_6 on a Varian instrument operating at 400 MHz (for ^1H) and 100 MHz (for ^{13}C) at ambient temperature. Chemical shifts are reported in parts per million and normalized to internal solvent peaks or tetramethylsilane (0 ppm).

(*R*)-2-(3-Fluoro-4-methylphenylsulfonamido)propanoic acid (2.3)
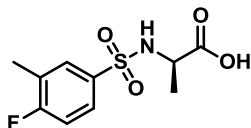
D-Alanine (**2.1**, 2.5 g, 28.1 mmol) was added to a solution of K_2CO_3 (8.3 g, 59.9 mmol) in dioxane/water (60 mL, 1:1, v/v). A solution of 4-fluoro-3-methylphenyl-sulfonylchloride (5.0 g, 24.0 mmol) in dioxane (4 mL) was added immediatly after with vigorous stirring. The mixture was stirred at rt overnight. Upon consumption of the starting material as determined by TLC, the solvent was reduced to one third the reaction volume under reduced pressure, and DCM was added to extract the organic layer. The aq. layer was acidified with conc. HCl to pH = 1 and extracted with EtOAc (3 x 20 mL).The combined organic layers were washed with brine (1 x 20 mL), dried over Na_2SO_4, filtered, and concentrated under reduced pressure to yield **2.3** as a white solid. ^1H NMR (400 MHz, $CDCl_3$) δ 9.57 (bs, 1H), 7.55-7.48 (m, 2H), 7.33 (m, 1H), 5.42 (d, *J* = 8.0 Hz, 1H, NH),

4.03 (m, 1H), 2.33 (s, 3H), 1.33 (d, $J = 6.0$ Hz, 3H).^{13}C NMR (100 MHz, CDCl$_3$) δ 177.0, 160.5 (d, $J = 49.0$ Hz), 138.8 (d, $J = 6.8$ Hz), 132.8 (d, $J = 4.5$ Hz), 131.2 (d, $J = 17.5$ Hz), 122.6 (d, $J = 3.8$ Hz), 114.1 (d, $J = 2.5$ Hz), 51.2, 19.6, 14.8.

(*R*)-Methyl 2-(4-fluoro-3-methylphenylsulfonamido)propanoate (2.4)
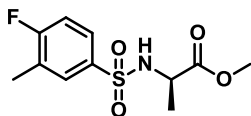
To a solution of **2.3** (4.9 g, 17.8 mmol) in MeOH (10 mL) was added a catalytic amount of conc. H$_2$SO$_4$. The reaction was heated to reflux. After 6 h, the solution was concentrated under reduced pressure and the resulting residue was dissolved in EtOAc (10 mL). This solution was washed with water (20 mL), sat. aq. NaHCO$_3$ (20 mL), brine (20 mL), dried over anhydrous MgSO$_4$, filtered, and concentrated under reduced pressure to yield **2.4** as a white solid (5.15 g, 78%, two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.72 (m, 1H), 7.67 (m, 1H), 5.54 (d, $J = 8.4$ Hz, 1H), 4.01 (m, 1H), 3.50 (s, 3H), 2.33 (s, 3H), 1.43 (d, $J = 7.4$ Hz, 0.5 H), 1.38 (d, $J = 7.4$ Hz, 2.5 H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 172.6, 163.8 (d, $J = 251.9$ Hz), 135.4 (d, $J = 3.0$ Hz), 130.9 (d, $J = 6.8$ Hz), 127.1 (d, $J = 9.8$ Hz), 126.4 (d, $J = 18.3$ Hz), 115.8 (d, $J = 24.2$ Hz), 52.7, 51.5, 19.9, 14.6.

General Procedure for N-Alkylation (2.5-2.9)

To a solution of **2.4** (110 mg, 0.4 mmol) and benzyl bromide (80 μL, 0.5 mmol) in anhydrous DMF (2 mL) was added K$_2$CO$_3$ (275.0 mg, 2.0 mmol). The reaction mixture was stirred at room temperature. After 48 h, the solvent was removed under reduced pressure and the resulting residue was taken up in H$_2$O, extracted with EtOAc (3 x 15

mL), dried over MgSO$_4$, filtered, and concentrated under reduced pressure. The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v).

(R)-methyl 2-(N-benzyl-4-fluoro-3-methylphenylsulfonamido)propanoate (2.5)
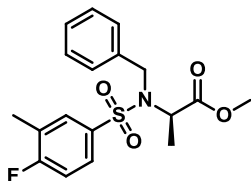
Colorless oil (120 mg, 82%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.64 (m, 2H), 7.32-7.23 (m, 5H), 7.08 (t, J = 9.2 Hz, 1H), 4.65 (q, J = 7.2 Hz, 1H), 4.57 (d, J = 16.0 Hz, 1H), 4.40 (d, J = 16.0 Hz, 1H), 3.47 (s, 3H), 2.30 (s, 3H), 1.30 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.6, 163.6 (d, J = 252 Hz), 137.1, 135.7 (d, J = 3.0 Hz), 131.1 (d, J = 6.0 Hz), 128.4, 128.1, 127.6, 127.3 (d, J = 9.1 Hz), 126.0 (d, J = 18.2 Hz), 115.6 (d, J = 23.6 Hz), 55.3, 52.1, 49.2, 16.6, 14.6 (d, J = 3.8 Hz).

(R)-methyl 2-(4-fluoro-3-methyl-N-(pyridin-3-ylmethyl)phenylsulfonamido)propanoate (2.6)
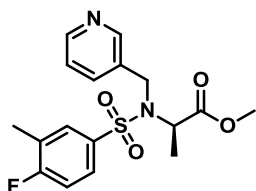
Colorless oil (126 mg, 79%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.50 (m, 2H), 7.90 (m, 1H), 7.62 (m, 2H), 7.27-7.23 (m, 1H), 7.10 (t, J = 8.8 Hz, 1H), 4.71 (q, J = 7.2 Hz, 1H), 4.60 (d, J = 16.8 Hz, 1H), 4.44 (d, J = 16.4 Hz, 1H), 3.50 (s, 3H), 2.31 (s, 3H), 1.32 (d, J = 6.8 Hz, 3.0 H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.4, 163.7 (d, J = 252.8 Hz), 149.2, 149.0,

135.9, 135.3 (d, J = 3.0 Hz), 133.2, 131.0 (d, J = 6.1 Hz), 127.3 (d, J = 9.1 Hz), 126.3 (d, J = 18.2 Hz), 123.4, 115.8 (d, J = 24.3 Hz), 55.3, 52.2, 46.6, 16.9, 14.6 (d, J = 3.1 Hz).

(*R*)-Methyl 2-(4-fluoro-3-methyl-*N*-(3-nitrobenzyl)phenylsulfonamido)propanoate (2.7)
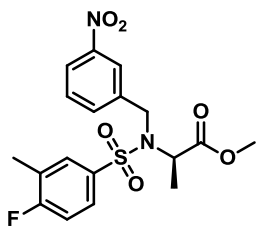
Colorless oil (147 mg, 83%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.14 (m, 2H), 7.77 (m, 1H), 7.63 (m, 2H), 7.51 (m, 1H), 7.08 (t, J = 7.6 Hz, 1H), 4.76 (m, 1H), 4.69 (d, J = 17.2Hz, 1H), 4.52 (d, J = 17.2Hz, 1H), 3.52 (s, 3H), 2.31 (s, 3H), 1.32 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.4, 163.8 (d, J = 252.7 Hz), 148.2, 140.1, 135.5 (d, J = 3.8 Hz), 133.8, 131.0 (d, J = 6.1 Hz), 129.4, 127.3 (d, J = 9.2 Hz), 126.4 (d, J = 18.2 Hz), 122.5, 122.4, 115.7 (d, J = 23.5 Hz), 55.5, 52.2, 48.3, 17.0, 14.6.

(R)-methyl 2-(4-fluoro-3-methyl-N-(4-nitrobenzyl)phenylsulfonamido)propanoate (2.8)

Colorless oil (155 mg, 52%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.17 (d, J = 8.0 Hz 2H), 7.66-7.57 (m, 4H), 7.13 (t, J = 8.0 Hz, 1H), 4.77-4.71 (m, 2H), 4.52 (d, J = 16.0 Hz, 1H), 3.50 (s, 3H), 2.33 (s, 3H), 1.30 (d, J = 4.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.4, 163.9

(d, J = 253.0 Hz), 147.32, 145.82, 134.9 (d, J = 3.0 Hz), 131.1 (d, J = 7.0 Hz), 128.35, 127.3 (d, J = 9.0 Hz), 126.4 (d, J = 19.0 Hz), 123.6, 115.8 (d, J = 24.0 Hz), 55.5, 52.2, 48.5, 17.1, 14.6 (d, J = 3 Hz).

(R)-methyl-2-(N-(4-(((tert-butoxycarbonyl)amino)methyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (2.9)

Colorless oil (380 mg, 71%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65-7.62 (m, 2H), 7.28 (d, J = 8.4 Hz, 2H), 7.20 (d, J = 8.0 Hz, 2H), 7.09 (t, J = 8.4 Hz, 1H), 4.91 (s, 1H), 4.63 (q, J = 7.6 Hz, 1H), 4.55 (d, J = 16.8 Hz, 1H), 4.38 (d, J = 16.4 Hz, 1H), 4.28 (s, 2H), 3.48 (s, 3H), 2.32 (s, 3H), 1.46 (s, 9H), 1.30 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.57, 163.65 (d, J = 252.0 Hz), 155.89, 138.47, 136.29, 136.10 (d, J = 3.7 Hz), 131.12 (d, J = 6.1 Hz), 128.21, 127.46, 127.33 (d, J = 9.9 Hz), 126.08 (d, J = 18.2 Hz), 115.63 (d, J = 23.6 Hz), 79.55, 55.27, 52.06, 48.94, 44.27, 28.40, 16.75, 14.59 (d, J = 3.0 Hz).

General Procedure for conversion of esters to hydroxamic acids (2.10-2.14)

To a solution of **2.5** in MeOH (1.0 mL) was added hydroxylamine hydrochloride (46 mg, 0.7 mmol) and NaOMe (225 μL, 1.0 mmol, 25 wt %) in MeOH at 0 °C. The reaction was allowed to gradually warm to rt and was stirred overnight. After 16 h, the solvent was removed under reduced pressure and the resulting residue suspended in brine. The aqueous layer was extracted with EtOAc (3 x 10 mL), and the combined organic phases

were dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude residue was purified by preparatory TLC using DCM/MeOH (19/1, v/v).

(R)-2-(N-benzyl-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (2.10)
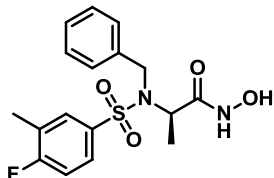
White foam (65 mg, 54%). ^{1}H NMR (400 MHz, CD_3OD) δ 7.63-7.58 (m, 2H), 7.35-7.20 (m, 5H), 7.13 (t, J = 8.8 Hz, 1H), 4.69 (m, 2H), 4.53 (q, J = 6.8 Hz 1H), 2.26 (s, 3H), 1.21 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD_3OD) δ 168.68, 163.56 (d, J = 250.5 Hz), 138.40, 135.75 (d, J = 3.1 Hz), 130.80 (d, J = 6.1 Hz), 127.82, 127.60, 127.09 (d, J = 9.9 Hz), 126.81, 125.97 (d, J = 19.0 Hz), 115.23 (d, J = 23.5 Hz), 53.1, 16.4, 13.07 (d, J = 3.8 Hz). MS (ESI) 367.22 [M + H]$^+$.

(R)-2-(4-fluoro-3-methyl-N-(pyridin-3-ylmethyl)phenylsulfonamido)-N-hydroxypropanamide (2.11)
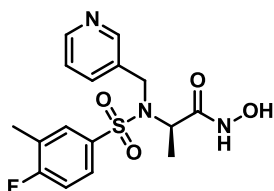
White foam (35 mg, 28%). ^{1}H NMR (400 MHz, CD_3OD) δ 8.53-8.40 (m, 2H), 7.88-7.86 (m, 1H), 7.66-7.64 (m, 2H), 7.37-7.35 (m, 1H), 7.17 (t, J = 9.2 Hz, 1H), 4.80-4.69 (m, 2H), 4.60-4.55 (m, 1H), 2.30 (s, 3H), 1.22 (d, J = 7.2 Hz, 3.0 H). ^{13}C NMR (100 MHz, CD_3OD) δ 168.4, 163.7 (d, J = 250.5 Hz), 148.2, 147.2, 136.6, 135.3 (d, J = 3.8 Hz),

130.8 (d, J = 6.0 Hz), 127.2 (d, J = 9.9 Hz), 126.3 (d, J = 19.0 Hz), 123.5, 115.4 (d, J = 24.3 Hz), 52.9, 45.5, 16.1, 13.0 (d, J = 3.0 Hz). MS (ESI) 368.26 [M + H]$^+$.

(*R*)-2-(4-Fluoro-3-methyl-*N*-(3-nitrobenzyl)phenylsulfonamido)-*N*-hydroxypropanamide (2.12)

White foam (92 mg, 73%). ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 10.95 (s, 1H), 8.87 (s, 1H), 8.17 (s, 1H), 8.07 (d, J = 8.4 Hz, 1H), 7.78 (d, J = 8 Hz, 1H), 7.70 (m, 1H), 7.65 (m, 1H), 7.58 (t, J = 8 Hz, 1H), 7.30 (t, J = 9.0 Hz, 1H), 4.80 (d, J = 17.4 Hz, 1H), 4.72 (d, J = 17.4 Hz, 1H), 4.46 (m, 1H), 2.25 (s, 3H), 1.10 (d, J = 7.2 Hz, 3.0 H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.8, 164.1 (d, J = 252.8 Hz), 148.2, 139.1, 134.6 (d, J = 3.2 Hz), 134.2, 130.8 (d, J = 6.1 Hz), 129.6, 127.1 (d, J = 9.9 Hz), 127.0, 122.9, 122.8, 116.3 (d, J = 23.5 Hz), 53.0, 47.5, 14.9, 14.6. MS (ESI) 412.22 [M + H]$^+$.

(R)-2-(4-fluoro-3-methyl-N-(4-nitrobenzyl)phenylsulfonamido)-N-hydroxypropanamide (2.13)

White foam (70 mg, 45%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.12 (d, J = 8.4 Hz 2H), 7.68-7.58 (m, 4H), 7.17 (t, J = 9.2 Hz, 1H), 4.88-4.72 (m, 2H), 4.57 (q, J = 6.4 Hz, 1H), 2.29 (s, 3H), 1.19 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.39, 163.80 (d, J =

250.4 Hz), 147.01, 146.84, 135.13 (d, J = 3.8 Hz), 130.86 (d, J = 6.0 Hz), 128.24, 127.25

(d, J = 9.1 Hz), 126.30 (d, J = 18.2 Hz), 122.87, 115.50 (d, J = 23.6 Hz), 52.89, 16.02,

13.04 (d, J = 3.8 Hz). MS (ESI) 412.17 [M + H]$^+$.

(R)-tert-butyl-4-((4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-

methylphenylsulfonamido)methyl)benzylcarbamate (2.14)

White foam (190 mg, 50 %). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.1 (s, 1H), 7.60-7.58 (m, 2H),

7.27 (d, J = 8.4 Hz, 2H), 7.21 (d, J = 7.6 Hz, 2H), 7.11 (t, J = 8.0 Hz, 1H), 4.99 (s, 1H),

4.62 (d, J = 15.6 Hz, 1H), 4.43 (brs, 1H), 4.28 (brs, 3H), 2.30 (s, 3H), 1.46 (s, 9H), 1.19

(d, J = 5.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.98, 163.86 (d, J = 253.5 Hz),

156.10, 138.80, 135.31, 135.04 (d, J = 3.7 Hz), 130.86 (d, J = 6.1 Hz), 128.80, 127.58,

127.06 (d, J = 9.1 Hz), 126.70 (d, J = 18.2 Hz), 116.06 (d, J = 23.5 Hz), 79.73, 53.17,

48.12, 44.30, 28.42, 14.61 (d, J = 3.0 Hz), 14.42.

(R)-2-(N-(3-aminobenzyl)-4-fluoro-3-methylphenylsulfonamido)-N-

hydroxypropanamide (2.15)
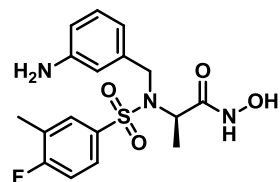
To a solution of **2.12** (130 mg, 0.3 mmol) in DCM (10 mL) under an H_2 atmosphere was added palladium on activated carbon (Pd/C) (130.0 mg, 10 wt. %). The reaction was shaken in a Parr apparatus at 4 atm at rt. After 4 h, the reaction mixture was filtered through a Celite pad and the filtrate was concentrated under reduced pressure. The crude residue was purified by flash column chromatography on SiO_2 using acetone/hexane (1/3, v/v) to yield the **2.15** as white foam (98 mg, 82%). ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 10.61 (s, 1H), 8.78 (s, 1H), 7.62 (m, 2H), 7.26 (t, J = 9.2 Hz, 1H), 6.90 (t, J = 7.6 Hz, 1H), 6.53 (s, 1H), 6.40 (m, 2H), 5.30-4.95 (bs, 2H, NH$_2$), 4.55 (d, J = 16.8 Hz, 1H), 4.46 (d, J = 16.8 Hz, 1H), 4.37 (m, 1H), 1.17 (d, J = 6.8 Hz, 3.0 H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 168.2, 163.8 (d, J = 263.5 Hz), 146.2, 137.4, 135.1, 130.9 (d, J = 6.0 Hz), 129.5, 127.1 (d, J = 9.1 Hz), 126.6 (d, J = 17.4 Hz), 119.1, (d, J = 24.3 Hz), 53.5, 48.5, 14.6, 14.1. MS (ESI) 382.33 [M + H]$^+$.

(R)-2-(N-(4-(aminomethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide hydrochloride (2.16)
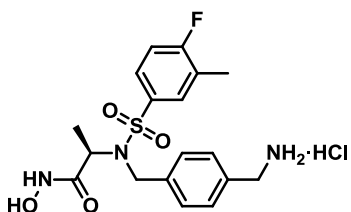
Compound **2.14** (190 mg, 0.4 mmol) was treated with 4N HCl in dioxane (2 mL) at rt. After 30 min, the solvent was removed under reduced pressure , and the resulting residue was triturated with diethyl ether (3 x 3 mL) to yield **2.16** as white solid (160.0 mg, 97%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.72-7.65 (m, 2H), 7.48 (d, J = 8.0 Hz, 2H), 7.40 (d, J = 8.0 Hz, 2H), 7.20 (t, J = 8.8 Hz, 1H), 4.79 (d, J = 16.8 Hz, 1H), 4.65 (d, J = 16.8 Hz, 1H), 4.51 (q, J = 7.2 Hz, 1H), 4.10 (s, 2H), 2.32 (s, 3H), 1.16 (d, J = 7.2 Hz, 3H). ^{13}C

NMR (100 MHz, CD$_3$OD) δ 168.56, 163.75 (d, J = 250.5 Hz), 140.27, 135.33 (d, J = 3.0 Hz), 131.79, 130.81 (d, J = 6.0 Hz), 128.50, 128.11, 127.20 (d, J = 9.9 Hz), 126.26 (d, J = 18.2 Hz), 115.42 (d, J = 23.5 Hz), 52.88, 42.64, 16.09, 13.05 (d, J = 3.8 Hz). MS (ESI) 396.36 [M + H]$^+$.

Chapter 3

MODELING, SYNTHESIS, AND *IN VITRO* EVALUATION OF SECOND-GENERATION LF INHIBITORS

Kurbanov, E. K.; Chiu, T.-L.; Solberg, J.; Francis, S.; Maize, K. M; Fernandez, J.; Johnson, R. L.; Finzel, B. C.; Hawkinson, J. E.; Walters, M. A.; Amin, E. A. Probing the S2′ Subsite of the Anthrax Toxin Lethal Factor Using N-Alkylated Hydroxamates. *Submitted* **2015.**

3.1 ACKNOWLEDGEMENTS

This chapter includes a description of work done in collaboration with Kimberly M. Maize, Barry C. Finzel, Rodney L. Johnson, Subhashree Francis, Jenna Fernandez, Michael A. Walters, Jonathan Solberg, and Jon E. Hawkinson. Subhashree Francis characterized the synthesized compounds by LC-MS. Jonathan Solberg tested the compounds in *in vitro* assays and Jon E. Hawkinson analyzed the data. X-ray crystallization studies with these compounds were completed by Kimberly M. Maize under the supervision of Barry C. Finzel. Jenna Fernandez did the work presented on **Table 3.2**. Synthetic guidance was provided by Rodney L. Johnson and Michael A. Walters.

3.2 INTRODUCTION

Based on the studies reported in chapter 2, we have concluded that domain 3 of LF is highly dynamic. This domain 3 mobility exposes crucial residues, such as Asp328, Lys380, and His654, to additional protein-ligand interactions. In this chapter, we will describe our efforts to take advantage of these important interactions.

3.3 MATERIALS AND METHODS

3.3.1 Molecular Modeling

The 4PKR co-crystal structure was used for further computational studies to target Asp328, Lys380, and His654 of LF with next generation **MK-31** analogs. This strategy included the design and docking and scoring of ~90 diverse analogs of **MK-31** using Glide 5.9 with the XP protocol in Schrödinger's Maestro Discovery Suite 9.4 (Schrödinger, Inc.).[118] Designed compounds were minimized, appropriately protonated

with LigPrep (Maestro), and docked with the XP protocol into the 4PKR co-crystal structure. Docking and scoring revealed that incorporating polar functionalities, such as amino, amido, or carboxyl groups, into the N-alkyl substituent would successfully engage the side chains of Lys380 and Asp328 in favorable charge-charge interactions and the backbone of His654 and Asp328 through H-bonding interactions. Upon analyzing the results of docking and scoring, we decided to synthesize a library of compounds bearing such polar functionalities.

3.3.2 Synthesis of Second-Generation Inhibitors

Scheme 3.1. Synthesis of Compounds 3.2a-3.2r.
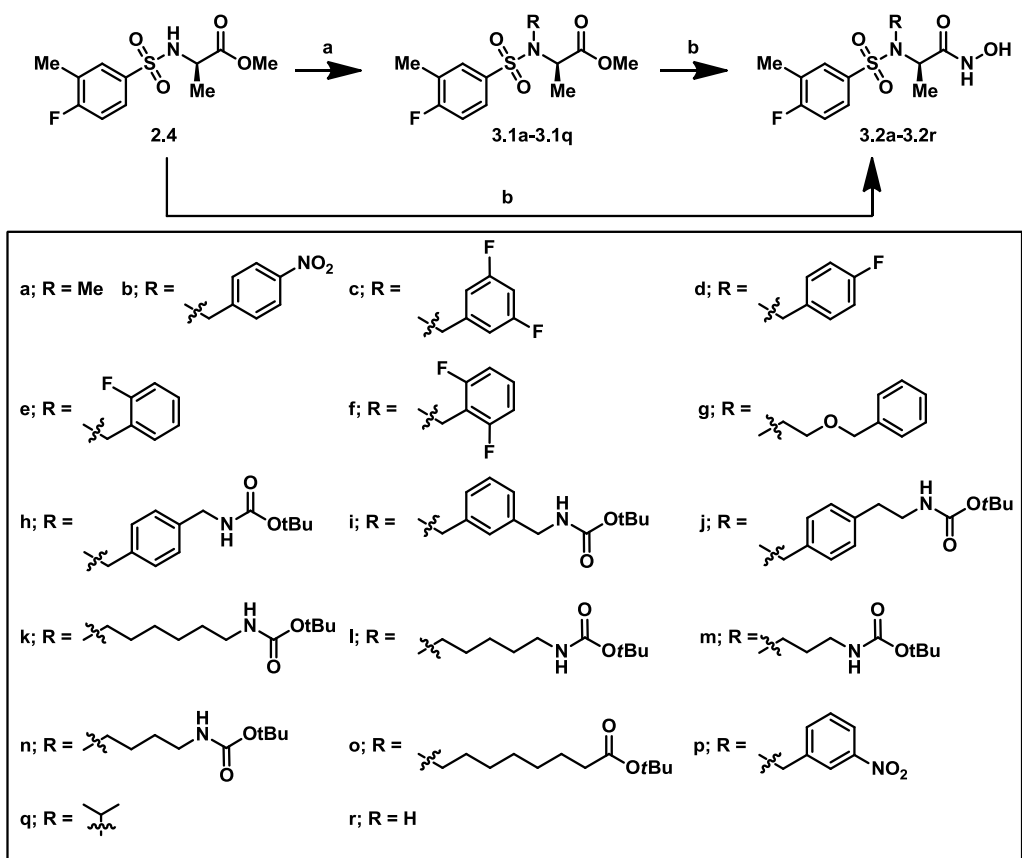
Reagents and Conditions: (a) R-Br, K_2CO_3, DMF, rt, 2 days; for **3.1q**: 2-propanol, PPh_3, DIAD, THF (b) $NH_2OH \cdot HCl$, NaOMe, MeOH, 0 °C to rt, 16 h.

40

Synthetic modifications to the sulfonamide of **MK-31** were accomplished from advanced intermediate **2.4**, as shown in **Scheme 3.1**. Intermediate **2.4** was alkylated using various bromides and K_2CO_3 in DMF to give **3.1a-3.1p** in 52%-88% yield (**Scheme 3.1a**). Intermediate **3.1q** was preferentially alkylated under Mitsunobu conditions employing PPh_3 and DIAD, which increased the yield to 54% from 17% with the S_N2 methodology. Esters **2.4** and **3.1a-3.1q** were converted to their corresponding hydroxamic acids **3.2a-3.2q** using hydroxylamine hydrochloride and NaOMe in MeOH in 27%–87% yield (**Scheme 3.1b**). Compound **3.2b** was further hydrogenated to **3.3a** using 10% wt. Pd/C in 43% yield. Compounds **3.2i-3.2o** were deprotected using 4 N HCl in 1,4-dioxane or a solution of TFA/DCM (1:2) to their corresponding hydrochloride or trifluoroacetate salts (**3.4a-3.4g**, **Scheme 3.2**).

Scheme 3.2. Synthesis of Compounds 3.4a-3.4g.
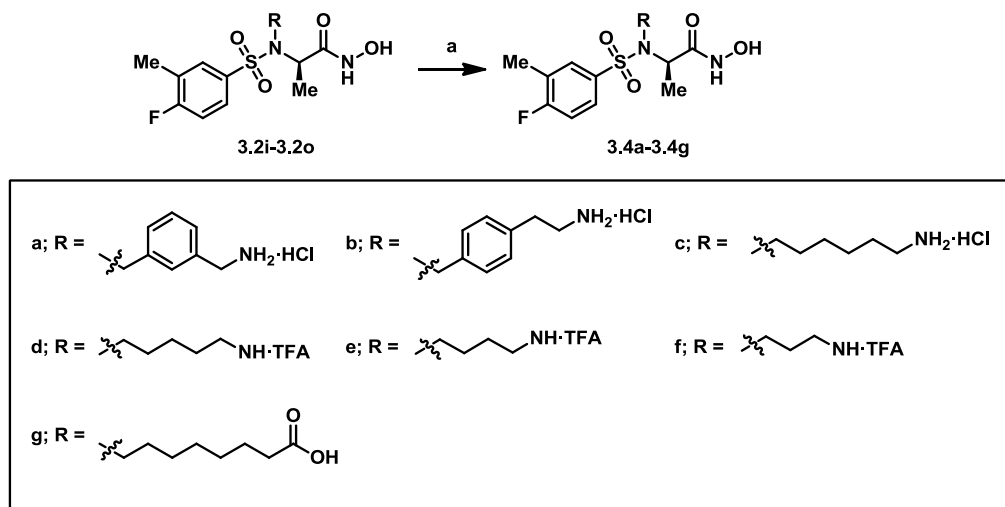
Reagents and Conditions: (a) 4 N HCl in dioxane, rt, 1 h; or TFA/DCM (1:2), TES, 0 °C.

Acetylated analogues **3.6a-3.6f** were synthesized through a three-step procedure from **3.1h-3.1m** (**Scheme 3.3**). Initially, Boc-protected intermediates **3.1h-3.1m** were

deprotected using 4 N HCl in 1,4-dioxane, and the resulting amines were acetylated with Ac$_2$O, TEA, and DMAP to give **3.5a-3.5f** in 33%-95% yield.

Scheme 3.3. Synthesis of Compounds 3.6a-3.6f
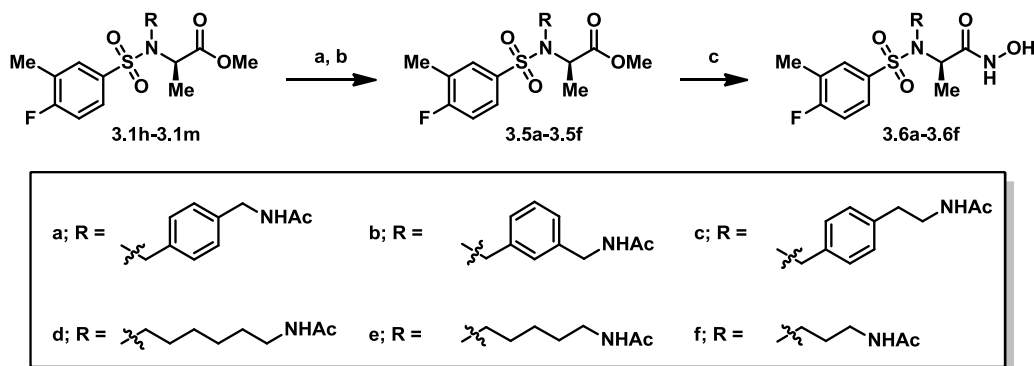
Reagents and Conditions: (a) 4 N HCl in dioxane, rt, 1 h; (b) Ac$_2$O, TEA, DMAP, THF, 0 °C, 16 h; (c) NH$_2$OH·HCl, NaOMe, MeOH, 0 °C to rt, 16 h.

Finally, the penultimate esters were converted to their corresponding hydroxamic acids using hydroxylamine hydrochloride and NaOMe in MeOH as described above. Methyl amide **3.8** was synthesized in parallel fashion over three steps (**Scheme 3.4**). Initially, *t*Bu-ester **3.1o** was deprotected with TFA/DCM (1:3). The resulting carboxylic acid was coupled with methylamine hydrochloride using EDCI, HOBt, and NMM to yield methyl ester **3.7** in 71% yield. The desired hydroxamate was accomplished using the previously described procedure in 30% yield.

Compound **3.11** was synthesized according to **Scheme 3.5**. The *m*-nitrobenzyl substituent was first hydrogenated to the *m*-amine using 10% wt. Pd/C in MeOH in 77% yield. The free amine was acetylated with AcCl to yield **3.10,** which was subsequently converted to hydroxamic acid **3.11** under the previously described conditions.

42

Scheme 3.4. Synthesis of Compound 3.8
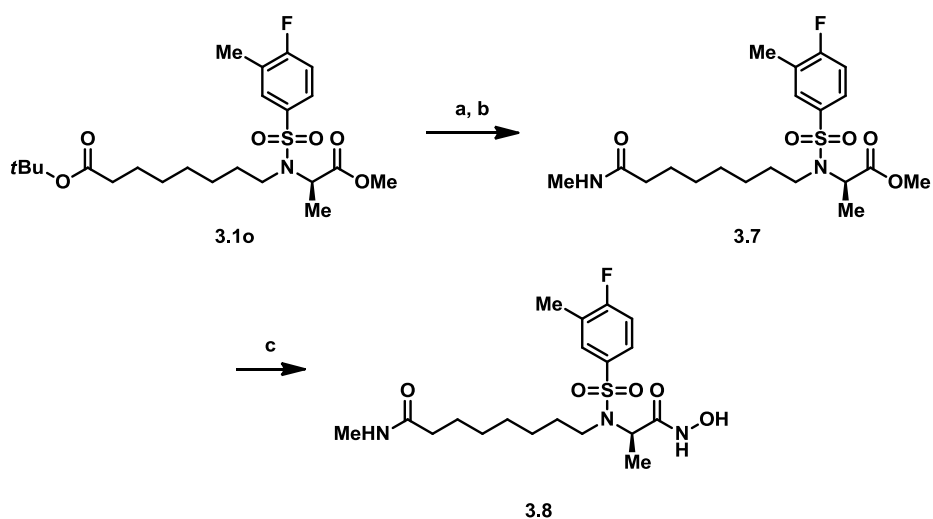
Reagents and Conditions: (a) TFA/DCM (1:3), rt, 2.5 h; (b) CH₃NH₂·HCl, NMM, HOBt, EDC, THF, rt, 16 h; (c) NH₂OH·HCl, NaOMe, MeOH, 0 °C to rt, 16 h.

Scheme 3.5. Synthesis of Compound 3.11
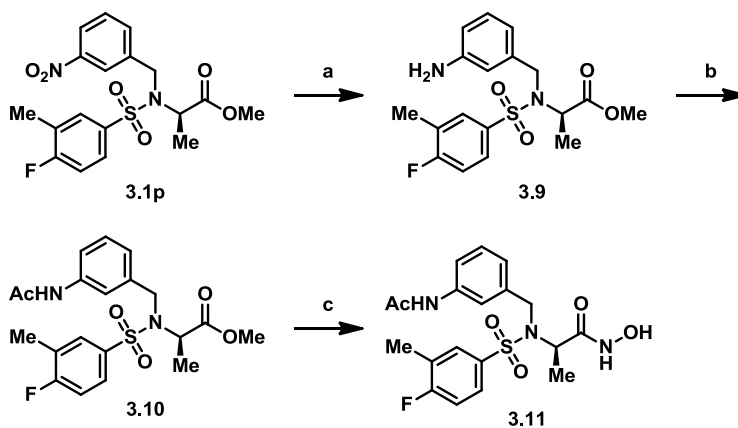
Reagents and Conditions: (a) 10 wt. % Pd/C, MeOH, rt, 4h; (b) AcCl, TEA, DCM, 0 °C to rt, 16 h; (c) NH₂OH·HCl, NaOMe, MeOH, 0 °C to rt, 16 h.

43

3.3.3 Biochemical Evaluation

3.3.3.1 Mobility Shift Protease (MSA) and FRET Assays

The off-chip mobility-shift protease assay uses a microfluidic chip to measure the conversion of fluorescent substrate to fluorescent product using a Caliper LC3000 (PerkinElmer). The terminated reaction mixture is introduced through a capillary sipper onto the chip where substrate and product are separated by electrophoresis and detected via laser-induced fluorescence. Briefly, 10 µL of 100 nM anthrax toxin lethal factor[111] (final concentration 50 nM) in 2X assay buffer (40 mM HEPES containing 0.02 % Triton X-100 (to prevent compound aggregation), pH 8.0) was added to 384-well assay plates (Corning #3677) using a MultiDrop (Thermo-Fisher). The plate was pre-incubated at 37 °C for 15 min and the reaction was initiated by the addition of 10 µL of 8 µM FITC-substrate (Celtek Peptides, #RK-10-4) in water (final substrate concentration 4 µM). The peptide sequence of the substrate is identical to the FRET assay substrate. The reaction was allowed to continue for 10 minutes at 37 °C, and then terminated with addition of 4 µL 0.5 mM phenanthroline/32.5 µM EDTA solution (final concentration of 0.1 mM phenanthroline /6.5 µM EDTA). Samples were then analyzed via the LabChip 3000 software. For IC_{50} determinations, test compounds were dissolved in DMSO at 10 mM and varying volumes were added to the assay plate using a Labcyte Echo® 550 acoustic dispenser prior to LF addition to achieve 8 concentrations in duplicate (final DMSO concentration 1 %). IC_{50} values were determined as a percent of control wells containing no inhibitor (following subtraction of background wells lacking enzyme) from at least 3 independent experiments using GraphPad Prism software. IC_{50} values were determined for the positive control compounds **MK-702/LF1-B** and **GM6001** included on every plate. FRET assay was performed as described earlier.

44

3.3.4 Structural Biology

Structural biology studies were performed as described earlier. Briefly, lethal factor protein (residues 265-776, A266S) was prepared and small molecule ligands were co-crystallized as reported previously.[125] Diffraction data for structure 4WF6 was collected at 100 K using a Saturn 944+ detector and a Rigaku Miromax-007FHM source at the University of Minnesota. The data were processed using HKL2000.[142]

Diffraction data for structures 5D1S, 5D1T, and 5D1U were collected at 100 K on beamline 17-ID-B (IMCA-CAT) using a Dectris PILATUS 6M pixel-array detector at the Advanced Photon Source of Argonne National Laboratories in Argonne IL. The data were processed using XDS[131] and scaled with SCALA.[132]

The structures were solved using molecular replacement with the atomic coordinates from either 1YQY or 4PKR using Phaser[135] in the CCP4 suite.[136] Structural refinement was done using both Refmac5[137] and Phenix,[138,139] while the Coot modeling software[140] was used for visualization and model building.

3.3.5 Molecular Dynamics Simulations

We performed molecular dynamics simulations on amine-containing analog **2.16** using Desmond 3.7[143–146] in Schrödinger's Maestro Discovery Suite 9.7 (Schrödinger, Inc.).[147] The 4PKV system was solvated with a TIP3P water model and simulated for 1.2 ns at constant pressure and temperature ensemble (NPT) at 300 K.
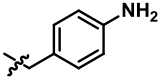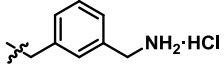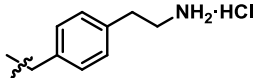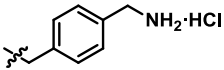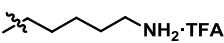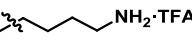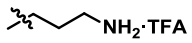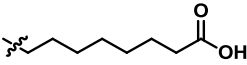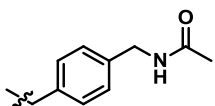
3.4 RESULTS AND DISCUSSION
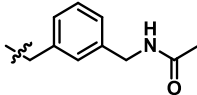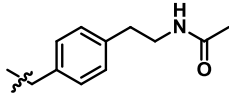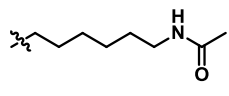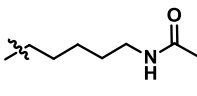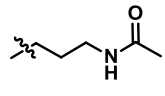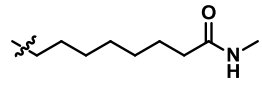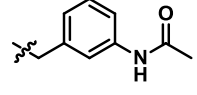
3.4.1 Biochemical Evaluation and Structural Biology

Table 3.1. Activities of novel **MK-31**-based LF inhibitors bearing substitutions at R_1.
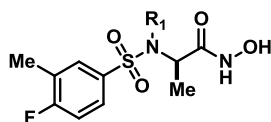
GPHR #[a]	Compound	R^1	MSA[b] IC$_{50}$ (µM)[c]	FRET IC$_{50}$ (µM)[c]
00223593	**3.2a**	Me	9.6 ± 0.1	37 ± 5
00225280	**3.2c**		7.0 ± 2	26 ± 5
00225273	**3.2d**		11 ± 3	29 ± 5
00225274	**3.2e**		9.5 ± 2	33 ± 3
00225275	**3.2f**		13 ± 2	42 ± 1
00223693	**3.2g**		14 ± 0.6	57 ± 6
00223595	**3.2h**		3.4 ± 0.1	15 ± 0.9
00223694	**3.2j**		2.1 ± 0.1	6.9 ± 1
00225282	**3.2o**		3.2 ± 0.3	8.0 ± 0.9

00223659	**3.2q**	isopropyl	>100	>100
00223407	**3.2r**	H	0.5 ± 0.1	1.6 ± 0.1
00223591	**3.3a**		6.4 ± 0.7	29 ± 6
00223599	**3.4a**		1.7 ± 0.1	10 ± 2
00223692	**3.4b**		1.2 ± 0.1	8.9 ± 0.6
00223596	**2.16***		1.3 ± 0.1	5.6 ± 0.3
00223664	**3.4c**		1.1 ± 0.1	4.4 ± 0.2
00223709	**3.4d**		1.3 ± 0.2	8.2 ± 0.3
00225278	**3.4e**		1.8 ± 0.1	7.4 ± 0.6
00225279	**3.4f**		3.7 ± 0.8	12 ± 3
00225281	**3.4g**		47 ± 2	91 ± 4
00223655	**3.6a**		3.5 ± 0.3	16 ± 4

00223657	**3.6b**		4.7 ± 0.5	15 ± 2
00223695	**3.6c**		3.2 ± 0.3	17 ± 2
00223696	**3.6d**		3.9 ± 0.5	14 ± 1
00223706	**3.6e**		4.3 ± 0.5	19 ± 3
00225277	**3.6f**		30 ± 2	117 ± 18
00225276	**3.8**		9.3 ± 1	19 ± 3
00223601	**3.11**		4.8 ± 0.2	21 ± 2

[a]GPHR # is a compound number in our in-house collection of compounds
[b]MSA = mobility shift assay
[c]IC$_{50}$ is the mean of 3 independent experiments.
*- previously published[125]

To gain insight into the SAR of our compounds, we employed the five co-crystal structures discussed in chapter 2 (PDB IDs 4PKR, 4PKS, 4PKT, 4PKU, and 4PKV) in further docking studies. Docking validation was performed by removing the co-crystallized ligands, then docking the ligands back into the LF active site using Surflex,[148,149] Glide,[114–117] AutoDock,[150] and MOE.[151] A docking program's accuracy is defined by how well the native pose of the ligand is reproduced by the highest ranked pose. An average root mean square deviation (RMSD) between the docked and

crystallized poses in each of the LF structures was used to compare the programs. **Table 3.2** shows RMSDs for each ligand. An effective docking tool should reliably identify the most favorable binding pose and highly rank this mode of interaction. Thus, programs that best reproduce the experimentally determined structure would be of most use in further docking and scoring studies. As we can see from **Table 3.2**, Glide reproduces the conformations observed in our crystal structures more accurately than the other programs. Thus, we selected Glide 5.9 since it performs better than Surflex,[148,149] AutoDock,[150] and MOE.[151] (**Table 3.2**).

Table 3.2. Comparison of RMSD values (Å) for Surflex, Glide, AutoDock, and MOE.

	Surflex 2.1	Surflex	Glide 5.9	AutoDock	MOE
4PKR	5.39	5.35	0.59	2.5	6.31
4PKS	5.39	1.36	1.00	0.49	5.92
4PKT	0.66	0.64	0.30	2.07	1.02
4PKU	1.02	1.04	0.98	3.93	1.42
4PKV	4.86	5.55	0.68	3.15	6.49
Average	3.46	2.79	**0.71**	2.43	4.23

To select the best crystal structures for further docking studies, we cross-docked the five co-crystallized ligands into each of the structures. The RMSD values between the docked ligands and the co-crystallized ligands were measured and are presented in **Table 3.3**. Structures 4PKS and 4PKV were selected for further docking studies since they exhibit the lowest RMSD values, of 0.91 Å and 0.93 Å respectively (**Table 3.3**).

Table 3.3. RSMD (Å) values for Glide 5.9 XP docked structures.

	4PKR	4PKS	4PKT	4PKU	4PKV
4PKR ligand	0.67	0.58	0.50	0.54	0.84
4PKS ligand	1.00	1.01	0.38	1.19	1.04
4PKT ligand	0.38	0.77	0.47	5.32	0.47
4PKU ligand	1.98	1.00	1.80	0.99	1.66
4PKV ligand	0.96	1.20	5.10	5.12	0.65
Average	1.00	**0.91**	1.65	2.63	**0.93**

Based on our previous studies, which are summarized in chapter 2,[125] we hypothesized that binding to a specific conformational state of domain 3 does not guarantee a more potent LF inhibitor. For example, the potent LF inhibitors **MK-702/LF1-B** (1YQY) and **MK-31** (4WF6) bind to domain 3 in the tight position and exhibit IC_{50}s in the nanomolar range; however, later generation N-alkylated inhibitors, which also bind to the tight domain 3 state (PDB ID 4PKT, previously discussed in chapter 2 and 5D1S, a complex with analog **3.2a** (**Figure 3.1**)), show drastically reduced potency. Moreover, N-alkylated analogues that bind to the open domain 3 state demonstrated better potencies than some tight binders. For example, structures 5D1T, a complex with compound **3.4a**, and 5D1U, a complex with **3.4c**, exhibit domain 3 in the open position, yet the co-crystallized ligands are among the more potent analogs. Based on these results, we have concluded that inhibitors are not required to bind the tight conformation of LF to achieve high potency, thus corroborating our conclusions in chapter 2.

From the crystallographic[125] and activity data, we can ascertain that the active site of LF is able to accommodate many diverse N-alkylated analogs of **MK-31**. The IC_{50}s of all synthesized analogs in the mobility shift assay (MSA) range from 1.1 µM to

50

47 µM, which can be compared to the 0.5 µM IC_{50} value of **MK-31**. Because the loss of inhibitory potency for N-alkylated analogs is not likely due to the ligand-induced conformational change of domain 3 as described above, we investigated the effect of losing an H-bond interaction with the sidechain of Tyr728. We were able to provide support for this hypothesis by synthesizing and co-crystallizing **3.2a** (PDB ID 5D1S), which contained an N-methyl modification. Based on the activity and crystal structure data for **3.2a** (**Figure 3.1**), we rationalized that the reduced inhibitory potency of N-alkylated **MK-31** analogs results from the loss of a key H-bond with the sidechain of Tyr728, which is part of the S1′ site.

Figure 3.1. LF active site with co-crystallized **MK-31** (yellow) (PDB ID 4WF6) superimposed onto the co-crystallized N-methyl analog **3.2a** (green) (PDB ID 5D1S). The circle denotes the sufonamide H. (Schrödinger Maestro Discovery Suite 9.4).
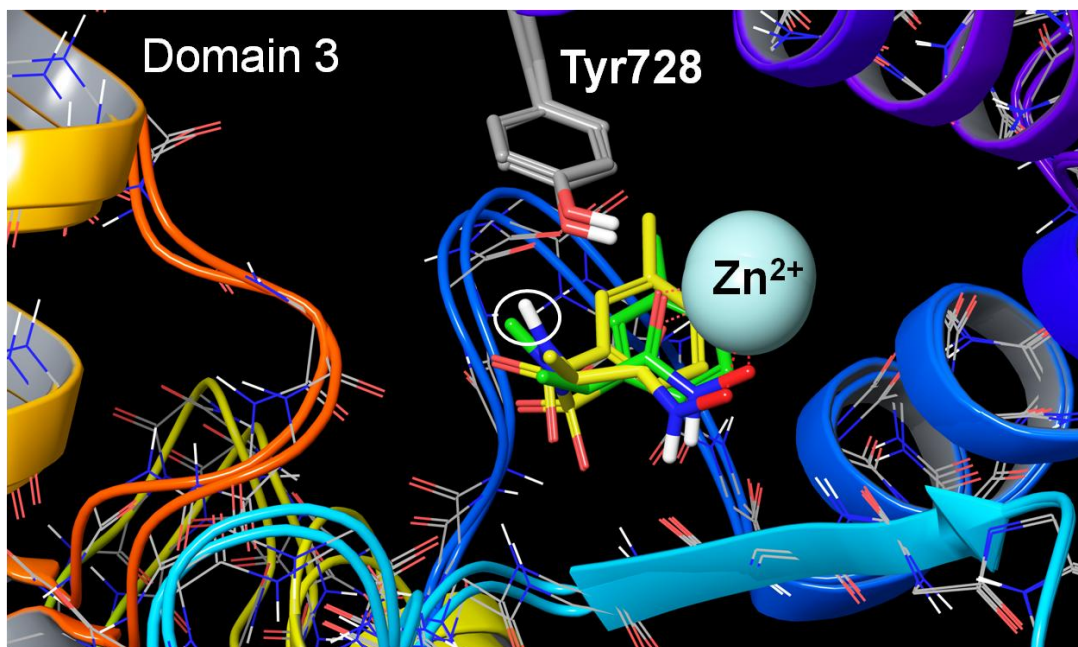
In both structures, domain 3 is observed in the tight position (**Figure 3.1**). The ligands align well, with the 4-fluoro-3-methylphenyl groups of **3.2a** and **MK-31** occupying

the S1′ site and the hydroxamate groups of both chelating zinc. Compound **3.2a**, however, exhibits a 20-fold loss in activity compared to **MK-31**, providing support for our hypothesis that the loss of an H-bond interaction with the sidechain of Tyr728 is the main reason for the loss of **3.2a**'s activity.

An inactive analog incorporated an isopropyl modification (**3.2q**). We hypothesize that the isopropyl analog **3.2q** is inactive due to a steric clash with the sidechain of Tyr728 (**Figure 3.2**). Our docking studies similarly predicted **3.2q** to be the least active. The steric clash with Tyr728 could be relieved by synthesizing the *n*-propyl analog, which shows inhibitory activity of 6.4 µM in the MSA.

Figure 3.2. LF active site with modeled **3.2q** (yellow). The circle denotes the steric clash between the isopropyl group (grey) and the sidechain of Tyr728 (yellow). (Schrödinger Maestro Discovery Suite 9.4).
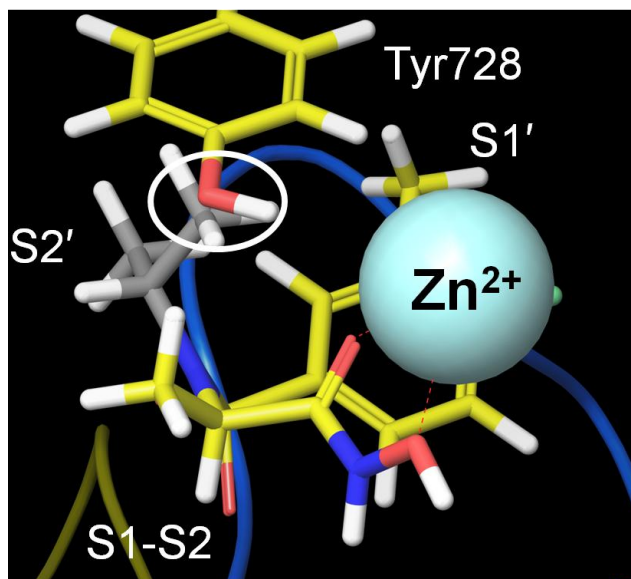
The most active analogs incorporated primary amines into the N-alkyl substituents, with the notable exception of **3.4f**. The docking studies rationalized that the primary amine-containing compounds are more active due to the introduction of

additional electrostatic interactions with the sidechain of Asp328, or H-bond donating interactions with the backbones of Asp325, Ser326 or Ser327. Analogue **3.4f**, however, exhibited a completely different orientation in the active site when docked into both crystal structures and was not predicted to have any interactions with the four aforementioned residues. As analog **3.4f** is characterized by the shortest N-alkyl substituent, a propyl linker, the substituent is likely too short to reach the interacting residues.

The amine-containing analogs **3.4a-3.4f** were more active than their Boc-protected or acetylated counterparts **3.2h**, **3.2j**, and **3.6a-3.6f**. Docking studies revealed that upon Boc-protection or acetylation, the analogs lose H-bond donating interactions with the backbone of Asp325, Ser236, and Ser327, or electrostatic interactions with Asp328.

Carboxylic acid-containing compound **3.4g**, butyl protected analog **3.2o,** and amide-containing analog **3.8** are all less active than the amine-containing analogues. This is not surprising since these compounds cannot engage in H-bond donating interactions with Asp325, Ser236, and Ser327, or electrostatic interactions with Asp328.

Hydrophobic compounds **3.2c-3.2g** do not engage in crucial protein-ligand electrostatic or H-bond donating interactions with Asp325, Ser236, Ser327, and Asp328, and are therefore less active than amines **3.4a-3.4e**.

Since 4PKV contains one of the most potent amine-containing analogs (**2.16**), we selected this co-crystal structure for molecular dynamics simulation studies. The molecular dynamics simulation demonstrated that the amino group was engaged in H-bond donating and electrostatic interactions with Asp325, and water-bridging H-bond interactions with Ser326 and Ser327 for 30% of the simulation time. Hence, electrostatic and H-bond donating interactions with Asp325, and water-bridging H-bond interactions

53

with Ser326 and Ser327 are responsible for the activity associated with primary amine-containing analogs **3.4a-3.4e**.

3.5 CONCLUDING REMARKS

The S2′ binding site of LF is a large, solvent exposed tunnel that undergoes a key conformational change upon binding to **MK-31** analogs alkylated at the sulfonamide N. We modified **MK-31** analogs to take advantage of the residues that were exposed as a result of the conformational change. We observed that steric clashes with Tyr728, near the entrance of the S2′ subsite, drastically decreased the potency of the compounds and should be taken into account during future inhibitor optimization. The biological activity and X-ray data indicate that the reduced inhibitory activity is not due to the conformational change of domain 3, but likely results from the loss of a key ligand-receptor H-bond to the sidechain of Tyr728. Partial inhibitory activity can be recovered by installing amine-containing substituents on the sulfonamide N, which can interact with the key residues Asp325, Ser326 and Ser327. Hence, it may be possible to improve the potency of future scaffolds, other than those based on **MK-31**, by incorporating amino-containing groups that specifically target the S2′ site residues Asp325, Ser326, and Ser327. Because of this observation, next generation compounds with improved potency may be obtained by targeting the S2′ subsite.

3.6 EXPERIMENTAL

FRET and MSA Correlation

As shown in **Figure 3.3**, there was a high correlation between the quenched FRET and mobility shift assays indicating that both assays provide reliable inhibitory

potency data for these compounds. On average, the compounds were 3.9-fold more potent in the mobility shift assay relative to the quenched FRET assay. This apparent potency difference is likely due to: 1) a 1.75-fold higher substrate concentration in the FRET relative to the mobility shift assay, and 2) the substrate K_m may be higher in the mobility shift assay relative to the FRET assay. Although the amino acid sequence is identical for the substrates used in both assays, the N-terminal and C-terminal modifications are different (FITC-substrate in MSA and Dnp/oAbz-substrate in FRET), which may affect cleavage efficiency by LF. Unfortunately, it is not technically feasible to determine the substrate K_m in the mobility shift assay because signal linearity is lost at high concentrations of substrate. The higher substrate concentration in the FRET assay coupled with a higher K_m in the mobility shift assay would decrease the apparent potency of these competitive inhibitors in the FRET assay.

Figure 3.3. Correlation between the quenched LF FRET and mobility shift assays.
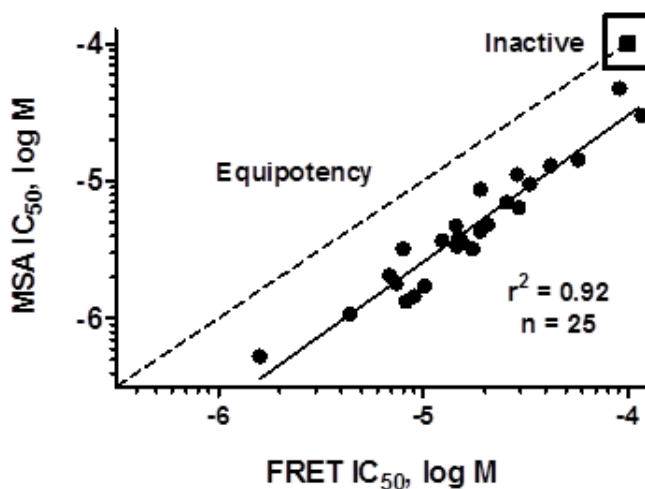
General Synthesis Information. Chemical reagents were purchased from commercial sources and used without additional purification. Bulk solvents were purchased from Fisher Scientific and anhydrous *N,N'*-dimethylformamide (DMF) was purchased from EMD Chemicals. Reactions were performed under an atmosphere of dry N_2 unless otherwise noted. Silica gel chromatography was performed on self-packed columns with SiliaFlash 60Å silica gel (SiliCycle). Preparatory thin layer chromatography (TLC) was performed on glass-backed SiliaPlate 60Å silica gel plates (SiliCycle). Compounds used in biological testing were no less than 95% pure as determined by two-wavelength HPLC analysis (254 and 215 nm) and nuclear magnetic resonance (NMR). The reaction mixture/purified compound was analyzed on a Waters UPLC which is connected to ELSD, UV and ZQ mass spectrometers. The mass analysis was carried out in both ESI positive and ESI negative modes. Multiple wavelengths of 214 nm, 220 nm, 244 nm and 254 were used for UV detection. Elution was carried out using a gradient mobile phase of 95:5 of water: acetonitrile with 0.1% formic acid to 100% acetonitrile over 6 minutes. The purity determination was carried out using OpenLynx. NMR spectra were recorded in $CDCl_3$, CD_3OD, or DMSO-d_6 on a Varian instrument operating at 400 MHz (for ^1H) and 100 MHz (for ^{13}C) at ambient temperature. Chemical shifts are reported in parts per million and normalized to internal solvent peaks or tetramethylsilane (0 ppm).

Procedure A for N-Alkylation (3.1a-3.1p)

To a solution of **2.4** (1 equiv.) and appropriate bromide (1.25 equiv.) in anhydrous DMF (0.4 M) was added K_2CO_3 (5 equiv.). The reaction mixture was stirred at rt. After 48 h, the solvent was removed under reduced pressure, and the resulting residue was taken up in H_2O (15 mL) and extracted with EtOAc (3 x 15 mL). The combined organic layers

were washed with brine (5 mL), dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure.

(R)-Methyl 2-(4-fluoro-N,3-dimethylphenylsulfonamido)propanoate (3.1a)
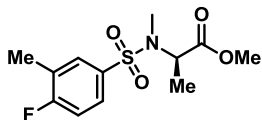
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/3, v/v) to yield the desired product as a colorless oil (110 mg, 96%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.70-7.63 (m, 2H), 7.14 (t, J = 8.8 Hz, 1H), 4.76 (q, J = 8.0 Hz, 1H), 3.57 (s, 3H), 2.84 (s, 3H), 2.34 (s, 3H), 1.36 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 170.91, 163.2 (d, J = 251.0 Hz), 134.5 (d, J = 3.0 Hz), 130.5 (d, J = 7.0 Hz), 126.8 (d, J = 9.0 Hz), 125.8 (d, J = 18.0 Hz), 115.2 (d, J = 24.0 Hz), 54.3, 51.6, 29.5, 15.0, 14.1 (d, J = 4.0 Hz).

(R)-Methyl 2-(4-fluoro-3-methyl-N-(4-nitrobenzyl)phenylsulfonamido)propanoate (3.1b)

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v) to yield the desired product as a colorless oil (155 mg, 52%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.17 (d, J = 8.0 Hz 2H), 7.66-7.57 (m, 4H), 7.13 (t, J = 8.0 Hz, 1H), 4.77-4.71 (m, 2H), 4.52 (d, J = 16.0 Hz, 1H), 3.50 (s, 3H), 2.33 (s, 3H), 1.30 (d, J = 4.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.4, 163.9 (d, J = 253.0 Hz), 147.32, 145.82, 134.9 (d, J = 3.0

Hz), 131.1 (d, J = 7.0 Hz), 128.35, 127.3 (d, J = 9.0 Hz), 126.4 (d, J = 19.0 Hz), 123.6, 115.8 (d, J = 24.0 Hz), 55.5, 52.2, 48.5, 17.1, 14.6 (d, J = 3 Hz).

(R)-Methyl 2-(N-(3,5-difluorobenzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1c)
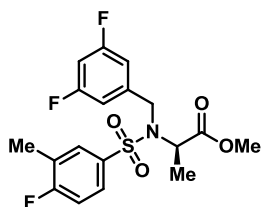
The crude product was purified over SiO_2 using an eluent of EtOAc/hexane (1/5, v/v) to yield the desired product as a colorless oil (71 mg, 68%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.64-7.60 (m, 2H), 7.11 (t, J = 8.4 Hz, 1H), 6.92-6.88 (m, 2H), 6.71-6.66 (m, 1H), 4.71 (q, J = 7.2 Hz, 1H), 4.60 (d, J = 17.2 Hz, 1H), 4.38 (d, J = 16.8 Hz, 1H), 3.51 (s, 3H), 2.33 (s, 3H), 1.31 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.47, 163.83 (d, J = 252.7 Hz), 163.08 (d, J = 247.4 Hz), 162.95 (d, J = 247.4 Hz), 142.18 (t, J = 9.1 Hz), 135.13 (d, J = 3.8 Hz), 131.11 (d, J = 6.1 Hz), 127.34 (d, J = 9.1 Hz), 126.34 (d, J = 18.2 Hz), 115.76 (d, J = 23.5 Hz), 110.49 (d, $^1J_{C-F}$ = 7.6 Hz, d, $^2J_{C-F}$ = 18.9 Hz), 102.91 (t, J = 25.0 Hz), 55.42, 52.14, 48.43, 17.0, 14.57 (d, J = 3.8 Hz).

(R)-Methyl 2-(4-fluoro-N-(4-fluorobenzyl)-3-methylphenylsulfonamido)propanoate (3.1d)
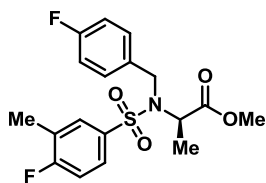
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/5, v/v) to yield the desired product as a colorless oil (75 mg, 64%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.64-7.61 (m, 2H), 7.33-7.29 (m, 2H), 7.10 (t, J = 8.0 Hz, 1H), 6.98 (t, J = 8.0 Hz, 2H), 4.67 (q, J = 8.0 Hz, 1H), 4.53 (d, J = 16.0 Hz, 1H), 4.37 (d, J = 16.0 Hz, 1H), 3.49 (s, 3H), 2.31 (s, 3H), 1.30 (d, J = 6.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.53, 163.65 (d, J = 252.0 Hz), 162.23 (d, J = 245 Hz), 135.60 (d, J = 3.0 Hz), 134.93 (d, J = 3.0 Hz), 131.04 (d, J = 6.0 Hz), 129.77 (d, J = 8.0 Hz), 127.22 (d, J = 9.0 Hz), 126.13 (d, J = 18.0 Hz), 115.67 (d, J = 23.0 Hz), 115.23 (d, J = 21.0 Hz), 55.23, 52.10, 48.41, 16.73, 14.58 (d, J = 4.0 Hz).

(R)-Methyl 2-(4-fluoro-N-(2-fluorobenzyl)-3-methylphenylsulfonamido)propanoate (3.1e)

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v) to yield the desired product as a colorless oil (55 mg, 71%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65-7.62 (m, 2H), 7.56 (t, J = 7.6 Hz, 1H), 7.27-7.2 (m, 1H), 7.14-7.05 (m, 2H), 6.96 (t, J = 9.2 Hz, 1H), 4.70 (q, J = 7.2 Hz, 1H), 4.59 (d, J = 16.8 Hz, 1H), 4.53 (d, J = 16.4 Hz, 1H), 3.48 (s, 3H), 2.31 (s, 3H), 1.33 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.44, 163.69 (d, J = 252.0 Hz), 160.28 (d, J = 244.4 Hz), 135.38 (d, J = 3.8 Hz), 131.09 (d, J = 6.0 Hz), 130.50 (d, J = 3.8 Hz), 129.24 (d, J = 8.4 Hz), 127.30 (d, J = 9.1 Hz), 126.11 (d, J = 18.2 Hz), 124.42 (d, J = 12.9 Hz), 124.16 (d, J = 3.8 Hz), 115.64 (d, J = 24.3 Hz), 115.01 (d, J = 21.3 Hz), 55.40, 52.09, 42.33, 16.33, 14.59 (d, J = 3.8 Hz).

(R)-Methyl 2-(N-(2,6-difluorobenzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1f)

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/5, v/v) to yield the desired product as a colorless oil (95 mg, 73%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65-7.63 (m, 2H), 7.28-7.19 (m, 1H), 7.06 (t, *J* = 9.2 Hz, 1H), 6.80 (t, *J* = 7.6 Hz, 2H), 4.62 (q, *J* = 7.2 Hz, 1H), 4.53 (s, 2H), 3.53 (s, 3H), 2.29 (s, 3H), 1.47 (d, *J* = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.44, 163.63 (d, *J* = 252.0 Hz), 161.89 (d, *J* = 249.7 Hz), 161.81 (d, *J* = 249.7 Hz), 135.42 (d, *J* = 3.8 Hz), 131.20 (d, *J* = 6.9 Hz), 130.24 (t, *J* = 9.8 Hz), 127.43 (d, *J* = 9.1 Hz), 125.83 (d, *J* = 18.2 Hz), 115.48 (d, *J* = 23.5 Hz), 112.14 (t, *J* = 17.5 Hz), 111.27 (d, $^{1}J_{C-F}$ = 19.7 Hz , d, $^{2}J_{C-F}$ = 6.1 Hz), 55.55, 52.18, 37.17, 15.42, 14.55 (d, *J* = 3.8 Hz).

(R)-Methyl 2-(N-(2-(benzyloxy)ethyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1g)

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/5, v/v) to yield the desired product as a colorless oil (222 mg, 71%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.68-7.61 (m, 2H), 7.35-7.25 (m, 5H), 7.06 (t, *J* = 8.8 Hz, 1H), 4.63 (q, *J* = 7.6 Hz, 1H), 4.47 (s, 2H), 3.72-3.61 (m, 2H), 3.53-3.46 (m, 4H), 3.40- 3.33 (m, 1H), 2.29 (s, 3H), 1.43

(d, *J* = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl₃) δ 171.67, 163.63 (d, *J* = 252.0 Hz),

137.98, 135.39 (d, *J* = 3.8 Hz), 131.01 (d, *J* = 6.8 Hz), 128.38, 127.68, 127.59, 127.28

(d, *J* = 9.1 Hz), 126.07 (d, *J* = 18.9 Hz), 115.57 (d, *J* = 23.6 Hz), 73.20, 69.91, 55.61,

52.04, 45.09, 16.77, 14.54 (d, *J* = 3.1 Hz).

(R)-Methyl 2-(N-(4-(((tert-butoxycarbonyl)amino)methyl)benzyl)-4-fluoro-3-

methylphenylsulfonamido)propanoate (3.1h)

The crude product was purified over SiO₂ using an eluent of EtOAc/hexane (1/3, v/v) to

yield the desired product as a colorless oil (380 mg, 71%). ^{1}H NMR (400 MHz, CDCl₃) δ

7.65-7.62 (m, 2H), 7.28 (d, *J* = 8.4 Hz, 2H), 7.20 (d, *J* = 8.0 Hz, 2H), 7.09 (t, *J* = 8.4 Hz,

1H), 4.91 (s, 1H), 4.63 (q, *J* = 7.6 Hz, 1H), 4.55 (d, *J* = 16.8 Hz, 1H), 4.38 (d, *J* = 16.4

Hz, 1H), 4.28 (s, 2H), 3.48 (s, 3H), 2.32 (s, 3H), 1.46 (s, 9H), 1.30 (d, *J* = 7.2 Hz, 3H).

^{13}C NMR (100 MHz, CDCl₃) δ 171.57, 163.65 (d, *J* = 252.0 Hz), 155.89, 138.47, 136.29,

136.10 (d, *J* = 3.7 Hz), 131.12 (d, *J* = 6.1 Hz), 128.21, 127.46, 127.33 (d, *J* = 9.9 Hz),

126.08 (d, *J* = 18.2 Hz), 115.63 (d, *J* = 23.6 Hz), 79.55, 55.27, 52.06, 48.94, 44.27,

28.40, 16.75, 14.59 (d, *J* = 3.0 Hz).

(R)-Methyl 2-(N-(3-(((tert-butoxycarbonyl)amino)methyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1i)
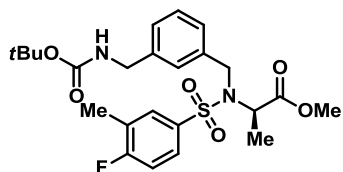
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/3, v/v) to yield the desired product as a colorless oil (350 mg, 73%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.64-7.62 (m, 2H), 7.24-7.15 (m, 4H), 7.09 (t, J = 8.0 Hz, 1H), 5.0 (bs, 1H), 4.64 (q, J = 7.6 Hz, 1H), 4.56 (d, J = 16.0 Hz, 1H) 4.38 (d, J = 16.0 Hz, 1H), 4.26-4.22 (m, 2H), 3.48 (s, 3H), 2.31 (s, 3H), 1.46 (s, 9H), 1.30 (d, J = 8.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.50, 163.55 (d, J = 252.0 Hz), 155.85, 139.24, 137.52, 135.61 (d, J = 4.0 Hz), 131.05 (d, J = 6.0 Hz), 128.60, 127.27 (d, J = 9.0 Hz), 126.87, 126.83, 126.61, 126.0 (d, J = 19.0 Hz), 115.67 (d, J = 23.0 Hz), 79.38, 62.78, 55.29, 52.01, 49.06, 28.35, 16.69, 14.51 (d, J = 4.0 Hz).

(R)-Methyl 2-(N-(4-(2-((tert-butoxycarbonyl)amino)ethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1j)
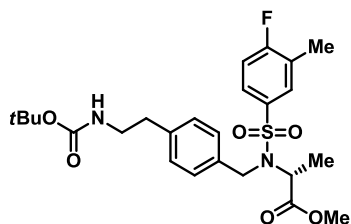
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v) to yield the desired product as a colorless oil (244 mg, 78%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.66-7.63 (m, 2H), 7.25 (d, J = 7.2 Hz, 2H), 7.13-7.07 (m, 3H), 4.66-4.61 (m, 2H), 4.53 (d, J = 16.0 Hz, 1H), 4.38 (d, J = 16.0 Hz, 1H), 3.48 (s, 3H), 3.35 (m, 2H), 2.78 (t, J = 6.8

Hz, 2H), 2.31 (s, 3H), 1.43 (s, 9H), 1.31 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$)

δ 171.57, 163.63 (d, J = 251.9 Hz), 155.85, 138.47, 135.67 (d, J = 3.8 Hz), 135.19,

131.12 (d, J = 6.1 Hz), 128.81, 128.24, 127.33 (d, J = 9.1 Hz), 126.06 (d, J = 19.0 Hz),

115.61 (d, J = 23.6 Hz), 79.2, 55.21, 52.04, 48.93, 41.77, 35.84, 28.4, 16.63, 14.58 (d, J

= 3.8 Hz).

(R)-Methyl 2-(N-(6-((tert-butoxycarbonyl)amino)hexyl)-4-fluoro-3-

methylphenylsulfonamido)propanoate (3.1k)
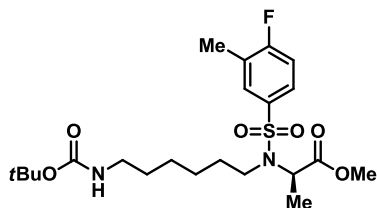
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/5, v/v) to

yield the desired product as a colorless oil (960 mg, 88%). ^{1}H NMR (400 MHz, CDCl$_3$) δ

7.69-7.63 (m, 2H), 7.11 (t, J = 8.4 Hz, 1H), 4.69 (s, 1H), 4.61 (q, J = 7.6 Hz 1H), 3.56 (s,

3H), 3.26-3.18 (m, 1H), 3.12-3.04 (m, 3H) 2.33 (s, 3H), 1.74-1.53 (m, 2H), 1.52-1.42 (m,

14H), 1.34-1.28 (m, 4H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.80, 163.56 (d, J = 251.3 Hz),

156.03, 135.67 (d, J = 3.8 Hz), 130.95 (d, J = 6.0 Hz), 127.22 (d, J = 9.1 Hz), 126.02 (d,

J = 19.0 Hz), 115.56 (d, J = 23.5 Hz), 78.91, 55.20, 52.10, 45.81, 40.39, 30.90, 29.98,

28.43, 26.56, 26.33, 16.74, 14.59 (d, J = 3.8 Hz).

(R)-Methyl 2-(N-(5-((tert-butoxycarbonyl)amino)pentyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1l)
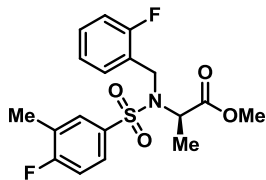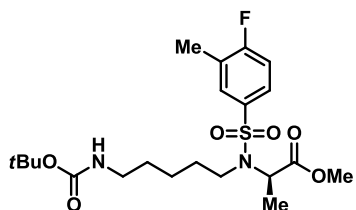

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v) to yield the desired product as a colorless oil (300 mg, 69%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.69-7.64 (m, 2H), 7.12 (t, *J* = 8.8 Hz, 1H), 4.81 (s, 1H), 4.61 (q, *J* = 7.2 Hz 1H), 3.56 (s, 3H), 3.26-3.06 (m, 4H), 2.34 (s, 3H), 1.74-1.24 (m, 18H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.82, 163.58 (d, *J* = 251.2 Hz), 156.07, 135.70 (d, *J* = 3.8 Hz), 130.95 (d, *J* = 6.8 Hz), 127.25 (d, *J* = 9.1 Hz), 126.05 (d, *J* = 18.3 Hz), 115.60 (d, *J* = 23.5 Hz), 78.87, 55.24, 52.10, 45.79, 40.38, 30.67, 29.70, 28.43, 24.11, 16.73, 14.57 (d, *J* = 3.0 Hz).

(R)-Methyl 2-(N-(3-((tert-butoxycarbonyl)amino)propyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1m)
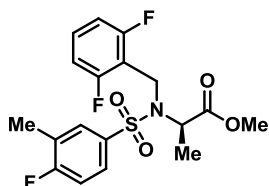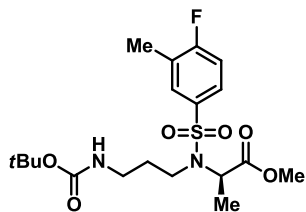

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/3, v/v) to yield the desired product as a colorless oil (315 mg, 81%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.69-7.64 (m, 2H), 7.12 (t, *J* = 8.8 Hz, 1H), 5.08 (s, 1H), 4.60 (q, *J* = 7.2 Hz 1H), 3.56 (s, 3H), 3.32-3.16 (m, 4H), 2.34 (s, 3H), 1.89-1.76 (m, 2H), 1.44-1.40 (m, 12H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.58, 163.50 (d, *J* = 252.0 Hz), 155.97, 135.28 (d, *J* = 3.0 Hz),

130.80 (d, J = 6.0 Hz), 127.11 (d, J = 9.8 Hz), 126.05 (d, J = 18.3 Hz), 115.55 (d, J = 23.5 Hz), 78.84, 55.15, 52.02, 43.09, 37.55, 31.02, 28.27, 16.37, 14.42 (d, J = 3.0 Hz).

(R)-Methyl 2-(N-(4-((tert-butoxycarbonyl)amino)butyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.1n)
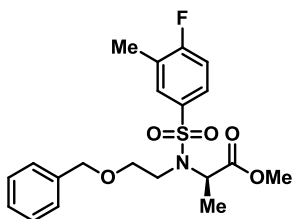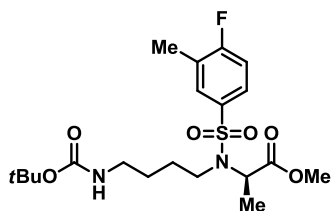

The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/3, v/v) to yield the desired product as a colorless oil (240 mg, 73%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.69-7.62 (m, 2H), 7.11 (t, J = 8.4 Hz, 1H), 4.77 (s, 1H), 4.61 (q, J = 7.2 Hz 1H), 3.55 (s, 3H), 3.30-3.09 (m, 4H), 2.34 (s, 3H), 1.76-1.62 (m, 1H), 1.62-1.58 (m, 1H), 1.49-1.41 (m, 14H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.69, 163.50 (d, J = 252.0 Hz), 155.94, 135.45 (d, J = 3.0 Hz), 130.85 (d, J = 6.0 Hz), 127.12 (d, J = 9.9 Hz), 126.0 (d, J = 18.2 Hz), 115.52 (d, J = 23.5 Hz), 78.93, 55.14, 52.14, 45.38, 39.84, 28.33, 28.11, 27.27, 16.62, 14.48 (d, J = 3.8 Hz).

(R)-tert-Butyl 8-(4-fluoro-N-(1-methoxy-1-oxopropan-2-yl)-3-methylphenylsulfonamido)octanoate (3.1o)
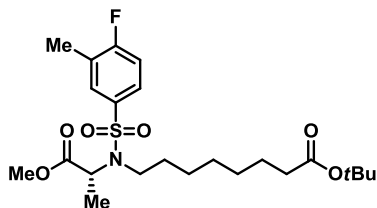
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/5, v/v) to yield the desired product as a colorless oil (270 mg, 66%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.70-7.64 (m, 2H), 7.11 (t, J = 8.8 Hz, 1H), 4.62 (q, J = 6.8 Hz, 1H), 3.56 (s, 3H), 3.26-3.19 (m, 1H), 3.12-3.05 (m, 1H), 2.33 (s, 3H), 2.19 (t, J = 7.6 Hz, 2H), 1.71-1.28 (m, 22H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 173.08, 171.83, 163.54 (d, J = 252.0 Hz), 135.80 (d, J = 3.8 Hz), 130.98 (d, J = 6.1 Hz), 127.27 (d, J = 9.1 Hz), 125.96 (d, J = 18.3 Hz), 115.54 (d, J = 23.5 Hz), 79.85, 55.20, 52.05, 45.90, 35.47, 30.89, 29.97, 28.87, 28.11, 26.76, 24.97, 16.73, 14.56 (d, J = 3.0 Hz).

(*R*)-Methyl 2-(4-fluoro-3-methyl-*N*-(3-nitrobenzyl)phenylsulfonamido)propanoate (3.1p)
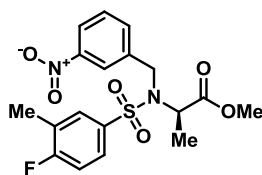
The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v) to yield the desired product as a colorless oil (147 mg, 83%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.17-8.07 (m, 2H), 7.79 (m, 1H), 7.67-7.63 (m, 2H), 7.53-7.51 (m, 1H), 7.12 (t, J = 7.6 Hz, 1H), 4.79-4.69 (m, 2H), 4.54 (d, J = 17.2Hz, 1H), 3.52 (s, 3H), 2.32 (s, 3H), 1.34 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.27, 163.71 (d, J = 252.7 Hz), 148.10, 140.15, 135.03 (d, J = 3.1 Hz), 133.86, 130.96 (d, J = 6.0 Hz), 129.38, 127.30 (d, J = 9.1 Hz), 126.37 (d, J = 18.2 Hz), 122.39, 122.34, 115.71 (d, J = 23.5 Hz), 55.49, 52.11, 48.22, 16.76, 14.4 (d, J = 3.8 Hz).

Procedure B for N-Alkylation (3.1q)

(R)-Methyl 2-(4-fluoro-N-isopropyl-3-methylphenylsulfonamido)propanoate (3.1q)
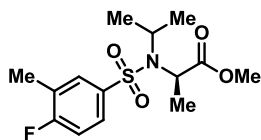
To a solution of **2.4** (100 mg, 0.36 mmol), 2-propanol (140 µL, 1.82 mmol), and triphenylphosphine (478 mg, 1.82 mmol) in THF (2 mL) at 0 °C was slowly added DIAD (360 µL, 1.82 mmol). The mixture was stirred at room temperature overnight. Upon consumption of the starting material as determined by TLC, the solvent was concentrated under reduced pressure. The crude product was purified over SiO_2 using a gradient elution of EtOAc/hexane (1/9 to 1/5, v/v) to yield **3.1q** as a colorless oil (62 mg, 54%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81-7.76 (m, 2H), 7.10 (t, J = 8.8 Hz, 1H), 4.09 (q, J = 7.2 Hz, 1H), 3.77-3.73 (m, 4H), 2.34 (s, 3H), 1.64 (d, J = 6.8 Hz, 3H), 1.19 (m, 6H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.3, 163.47 (d, J = 252.0 Hz), 136.92 (d, J = 3.8 Hz), 131.41 (d, J = 6.0 Hz), 127.56 (d, J = 9.1 Hz), 125.86 (d, J = 18.2 Hz), 115.53 (d, J = 23.5 Hz), 52.59, 52.46, 50.22, 21.61, 20.93, 18.21, 14.63 (d, J = 3.8 Hz).

General Procedure for the Conversion of Esters to Hydroxamic acids (3.2a-3.2r, 3.6a-3.6f, 3.8, 3.11)

To a solution of ester (1 equiv.) in anhydrous MeOH (0.38 M) was added hydroxylamine hydrochloride (2 equiv.) and NaOMe (3 equiv., 25 wt. % in MeOH) at 0 °C. The reaction was allowed to gradually warm to rt and was stirred overnight. After 16 h, the solvent was removed under reduced pressure, and the resulting residue was suspended in brine (5 mL). The aqueous layer was extracted with EtOAc (3 x 10 mL), and the combined

organic phases were dried over Na_2SO_4, filtered, and concentrated under reduced pressure.

(R)-2-(4-fluoro-N,3-dimethylphenylsulfonamido)-N-hydroxypropanamide (3.2a)
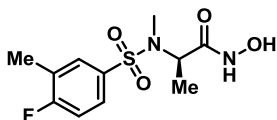
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (19/1, v/v) to yield the desired product as a white foam (55 mg, 50%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.75-7.67 (m, 2H), 7.22 (t, J = 8.0 Hz, 1H), 4.52 (q, J = 8.0 Hz, 1H), 2.87 (s, 3H), 2.32 (s, 3H), 1.13 (d, J = 8.0 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 169.83, 165.11 (d, J = 250.0 Hz), 135.73 (d, J = 3.0 Hz), 132.10 (d, J = 7.0 Hz), 128.58 (d, J = 9.0 Hz), 127.82 (d, J = 19.0 Hz), 116.98 (d, J = 25.0 Hz), 54.3, 30.84, 14.67, 14.55 (d, J = 3.0 Hz). MS (ESI) 291.0 [M + H]$^+$.

(R)-2-(4-Fluoro-3-methyl-N-(4-nitrobenzyl)phenylsulfonamido)-N-hydroxypropanamide (3.2b)
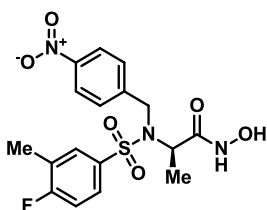
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (19/1, v/v) to yield the desired product as a white foam (70 mg, 45%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.12 (d, J = 8.4 Hz 2H), 7.68-7.58 (m, 4H), 7.17 (t, J = 9.2 Hz, 1H), 4.88-4.72 (m, 2H), 4.57 (q, J = 6.4 Hz, 1H), 2.29 (s, 3H), 1.19 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.39, 163.80 (d, J = 250.4 Hz), 147.01, 146.84, 135.13 (d,

J = 3.8 Hz), 130.86 (d, *J* = 6.0 Hz), 128.24, 127.25 (d, *J* = 9.1 Hz), 126.30 (d, *J* = 18.2

Hz), 122.87, 115.50 (d, *J* = 23.6 Hz), 52.89, 16.02, 13.04 (d, *J* = 3.8 Hz). MS (ESI) 412.1

[M + H]$^+$.

(R)-2-(N-(3,5-Difluorobenzyl)-4-fluoro-3-methylphenylsulfonamido)-N-

hydroxypropanamide (3.2c)
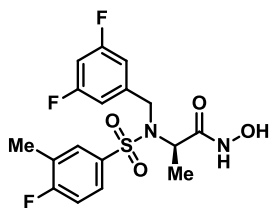
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH

(20/1, v/v) to yield the desired product as a colorless oil (46 mg, 68%). ^{1}H NMR (400

MHz, CDCl$_3$) δ 7.68-7.64 (m, 2H), 7.17 (t, *J* = 8.8 Hz, 1H), 6.97-6.95 (m, 2H), 6.77 (t, *J* =

7.2 Hz, 1H), 4.60 (d, *J* = 17.2 Hz, 1H), 4.38 (d, *J* = 17.2 Hz, 1H), 4.54 (q, *J* = 7.2 Hz, 1H),

2.30 (s, 3H), 1.21 (d, *J* = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 168.45, 163.77 (d, *J*

= 251.2 Hz), 163.00 (d, *J* = 245.9 Hz), 162.87 (d, *J* = 245.9 Hz), 143.78 (t, *J* = 9.1 Hz),

135.27 (d, *J* = 3.8 Hz), 130.85 (d, *J* = 6.8 Hz), 127.23 (d, *J* = 9.8 Hz), 126.24 (d, *J* = 18.2

Hz), 115.40 (d, *J* = 24.3 Hz), 110.13 (d, $^1J_{C-F}$ = 6.8 Hz, d, $^2J_{C-F}$ = 19.0 Hz), 101.70 (t, *J* =

25.8 Hz), 52.91, 16.21, 13.03 (d, *J* = 3.8 Hz). MS (ESI) 403.0 [M + H]$^+$.

(R)-2-(4-Fluoro-N-(4-fluorobenzyl)-3-methylphenylsulfonamido)-N-

hydroxypropanamide (3.2d)
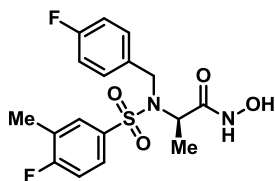
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (19/1, v/v) to yield the desired product as a white foam (32 mg, 40%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.63-7.58 (m, 2H), 7.37-7.34 (m, 2H), 7.14 (t, J = 8.8 Hz, 1H), 6.98 (t, J = 8.8 Hz, 2H), 4.65 (d, J = 16.4 Hz, 1H), 4.58 (d, J = 16.4 Hz, 1H), 4.54 (q, J = 7.2 Hz, 1H), 2.27 (s, 3H), 1.21 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.64, 163.60 (d, J = 250.4 Hz), 162.05 (d, J = 242.9 Hz), 135.75 (d, J = 3.0 Hz), 134.36 (d, J = 3.0 Hz), 130.76 (d, J = 6.1 Hz), 129.59 (d, J = 8.3 Hz), 127.07 (d, J = 9.9 Hz), 126.02 (d, J = 18.2 Hz), 115.28 (d, J = 24.3 Hz), 114.42 (d, J = 21.3 Hz), 53.07, 16.29, 13.02 (d, J = 3.8 Hz). MS (ESI) 385.1 [M + H]$^+$.

(R)-2-(4-Fluoro-N-(2-fluorobenzyl)-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.2e)
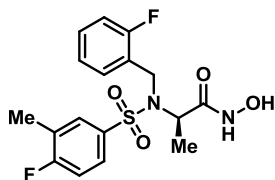
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (17 mg, 31%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.67-7.64 (m, 2H), 7.51 (t, J = 7.4 Hz, 1H), 7.27-7.09 (m, 3H), 6.99 (t, J = 8.8 Hz, 1H), 4.64-4.52 (m, 3H), 2.30 (s, 3H), 1.23 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.64, 163.71 (d, J = 250.5 Hz), 160.13 (d, J = 243.6 Hz), 135.32 (d, J = 3.0 Hz), 130.82 (d, J = 6.8 Hz), 129.85 (d, J = 3.8 Hz), 128.60 (d, J = 7.6 Hz), 127.17 (d, J = 9.9 Hz), 126.12 (d, J = 18.2 Hz), 125.11 (d, J = 12.9 Hz), 123.69 (d, J = 3.8 Hz), 115.34 (d, J = 23.5 Hz), 114.43 (d, J = 21.3 Hz), 52.91, 41.42, 15.87, 13.04 (d, J = 3.8 Hz). MS (ESI) 385.0 [M + H]$^+$.

(R)-2-(N-(2,6-Difluorobenzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.2f)

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (40 mg, 42%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.59-7.51 (m, 2H), 7.26 (p, J = 8.0 Hz, 1H), 7.11 (t, J = 8.8 Hz, 1H), 6.82 (t, J = 8.4 Hz, 2H), 4.69-4.57 (m, 3H), 2.25 (s, 3H), 1.40 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.63, 163.57 (d, J = 249.7 Hz), 161.75 (d, J = 248.2 Hz), 161.68 (d, J = 248.2 Hz), 135.46 (d, J = 3.0 Hz), 130.62 (d, J = 6.8 Hz), 129.82 (t, J = 10.6 Hz), 127.02 (d, J = 9.1 Hz), 125.83 (d, J = 18.2 Hz), 115.22 (d, J = 23.5 Hz), 112.92 (t, J = 16.7 Hz), 110.85 (d, $^1J_{C-F}$ = 19.0 Hz , d, $^2J_{C-F}$ = 6.1 Hz), 53.92, 36.63, 14.82, 13.06 (d, J = 3.8 Hz). MS (ESI) 403.1 [M + H]$^+$.

(R)-2-(N-(2-(Benzyloxy)ethyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.2g)

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (18/1, v/v) to yield the desired product as a white foam (100 mg, 44%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.68-7.61 (m, 2H), 7.35-7.26 (m, 5H), 7.05 (t, J = 8.8 Hz, 1H), 4.54-4.48 (m, 3H), 3.82-3.80 (m, 1H), 3.58-3.56 (m, 2H), 3.24-3.22 (m, 1H), 2.26 (s, 3H), 1.22 (d, J

71

= 5.6 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 169.06, 163.71 (d, J = 250.5 Hz), 137.92, 135.20 (d, J = 3.0 Hz), 130.74 (d, J = 6.1 Hz), 128.01, 127.51, 127.36, 127.21 (d, J = 9.1 Hz), 126.34 (d, J = 18.2 Hz), 115.49 (d, J = 24.3 Hz), 72.71, 68.98, 53.54, 43.97, 14.90, 13.09 (d, J = 3.8 Hz). MS (ESI) 409.37 [M - H]$^-$.

(R)-tert-Butyl 4-((4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)methyl)benzylcarbamate (3.2h)

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (190 mg, 50%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.1 (s, 1H), 7.60-7.58 (m, 2H), 7.27 (d, J = 8.4 Hz, 2H), 7.21 (d, J = 7.6 Hz, 2H), 7.11 (t, J = 8.0 Hz, 1H), 4.99 (s, 1H), 4.62 (d, J = 15.6 Hz, 1H), 4.43 (brs, 1H), 4.28 (brs, 3H), 2.30 (s, 3H), 1.46 (s, 9H), 1.19 (d, J = 5.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.98, 163.86 (d, J = 253.5 Hz), 156.10, 138.80, 135.31, 135.04 (d, J = 3.7 Hz), 130.86 (d, J = 6.1 Hz), 128.80, 127.58, 127.06 (d, J = 9.1 Hz), 126.70 (d, J = 18.2 Hz), 116.06 (d, J = 23.5 Hz), 79.73, 53.17, 48.12, 44.30, 28.42, 14.61 (d, J = 3.0 Hz), 14.42. (ESI) 496.34 [M + H]$^+$.

(R)-tert-Butyl 3-((4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)methyl)benzylcarbamate (3.2i)
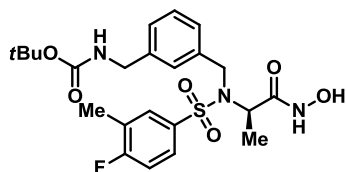
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a colorless oil (216 mg, 62%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.40 (brs, 1H), 7.59 (m, 2H), 7.23-7.06 (m, 5H), 5.22 (bs, 1H), 4.63 (d, J = 16.0 Hz, 1H) 4.46 (m, 1H), 4.26-4.23 (m, 3H), 2.28 (s, 3H), 1.45 (s, 9H), 1.14 (m, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.25, 163.74 (d, J = 252.0 Hz), 156.31, 139.30, 136.65, 135.10 (d, J = 4.0 Hz), 130.80 (d, J = 6.0 Hz), 128.67, 127.27, 127.01 (d, J = 9.0 Hz), 126.75, 126.54 (d, J = 19.0 Hz), 115.96 (d, J = 24.0 Hz), 79.77, 60.41, 53.14, 48.25, 28.38, 21.01, 14.53 (d, J = 3.0 Hz).

(R)-tert-Butyl 4-((4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)methyl)phenethylcarbamate (3.2j)
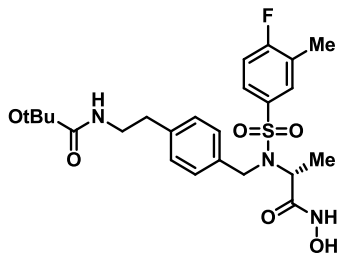
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (14/1, v/v) to yield the desired product as a white foam (80 mg, 27%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.60-7.59 (m, 2H), 7.24 (d, J = 8.0 Hz, 2H), 7.14-7.09 (m, 3H), 6.53 (s, 1H), 4.66 (d, J = 16.4 Hz, 1H), 4.60 (d, J = 16.4 Hz, 1H), 4.49 (q, J = 7.2 Hz, 1H), 3.21 (t,

J = 7.2 Hz, 2H), 2.71 (t, *J* = 7.2 Hz, 2H), 2.26 (s, 3H), 1.40 (s, 9H), 1.19 (d, *J* = 6.8 Hz,

3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.70, 163.55 (d, *J* = 250.5 Hz), 138.26, 136.24,

135.79 (d, *J* = 3.8 Hz), 130.79 (d, *J* = 6.1 Hz), 128.29, 127.75, 127.11 (d, *J* = 9.9 Hz),

125.95 (d, *J* = 19.0 Hz), 115.21 (d, *J* = 24.3 Hz), 78.54, 53.13, 41.58, 35.41, 27.36,

16.34, 13.05 (d, *J* = 3.8 Hz). MS (ESI) 510.45 [M + H]$^+$.

(R)-tert-Butyl (6-(4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-

methylphenylsulfonamido)hexyl)carbamate (3.2k)
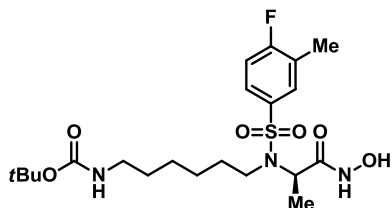
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH

(24/1, v/v) to yield the desired product as a yellow foam (306 mg, 87%). ^{1}H NMR (400

MHz, CD$_3$OD) δ 7.76-7.69 (m, 2H), 7.21 (t, *J* = 8.0 Hz, 1H), 4.44-4.43 (m, 1H), 3.43-3.31

(m, 1H), 3.31-3.15 (m, 1H) 3.03-3.00 (m, 2H), 2.33 (s, 3H), 1.67-1.59 (m, 2H), 1.42 (m,

11H), 1.28-1.21 (m, 7H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.82, 163.61 (d, *J* = 250.4

Hz), 157.09, 135.66 (d, *J* = 3.0 Hz), 130.66 (d, *J* = 6.1 Hz), 127.12 (d, *J* = 9.1 Hz),

126.24 (d, *J* = 18.2 Hz), 115.48 (d, *J* = 24.3 Hz), 78.39, 52.76, 44.72, 39.86, 30.65,

29.46, 27.50, 26.23, 26.03, 14.78, 13.20 (d, *J* = 3.7 Hz). MS (ESI) 476.33 [M + H]$^+$.

(R)-tert-Butyl (5-(4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)pentyl)carbamate (3.2l)
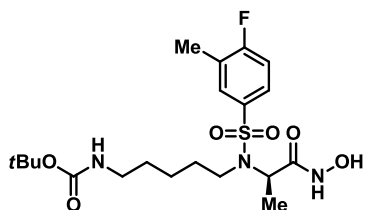
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (24/1, v/v) to yield the desired product as a white foam (60 mg, 30%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.76-7.69 (m, 2H), 7.21 (t, *J* = 8.4 Hz, 1H), 4.43-4.42 (m, 1H), 3.42 (brs, 1H), 3.18 (brs, 1H) 3.01 (brs, 2H), 2.33 (s, 3H), 1.67-1.21 (m, 18H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.80, 163.65 (d, *J* = 250.5 Hz), 157.12, 135.50 (d, *J* = 3.1 Hz), 130.67 (d, *J* = 6.1 Hz), 127.12 (d, *J* = 9.1 Hz), 126.27 (d, *J* = 18.2 Hz), 115.46 (d, *J* = 24.3 Hz), 78.42, 52.76, 44.68, 39.81, 30.41, 29.14, 27.44, 23.71, 14.76, 13.12 (d, *J* = 3.8 Hz). MS (ESI) 462.38 [M + H]$^+$.

(R)-tert-Butyl (3-(4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)propyl)carbamate (3.2m)
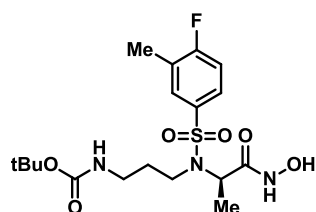
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (120 mg, 50%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.76-7.70 (m, 2H), 7.20 (t, *J* = 8.8 Hz, 1H), 4.45-4.43 (m, 1H), 3.48-3.46 (m, 1H), 3.23-3.22 (m, 1H), 3.05 (brs, 2H), 2.33 (s, 3H), 1.82-1.81 (m, 2H), 1.42 (s, 9H),

1.21 (d, *J* = 6.0 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.81, 163.68 (d, *J* = 250.5 Hz), 157.05, 135.60 (d, *J* = 3.8 Hz), 130.71 (d, *J* = 6.0 Hz), 127.16 (d, *J* = 9.9 Hz), 126.32 (d, *J* = 18.2 Hz), 115.51 (d, *J* = 23.5 Hz), 78.61, 52.88, 42.36, 37.49, 30.96, 27.47, 14.64, 13.2 (d, *J* = 3.0 Hz).

(R)-tert-Butyl (4-(4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)butyl)carbamate (3.2n)
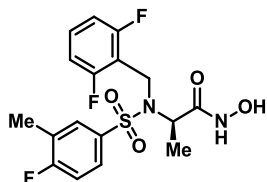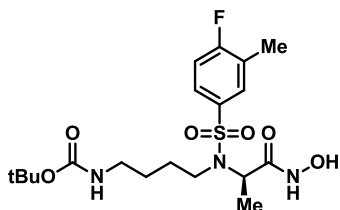

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (150 mg, 63%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.77-7.70 (m, 2H), 7.20 (t, *J* = 8.8 Hz, 1H), 4.43-4.42 (m, 1H), 3.43-3.42 (m, 1H), 3.21-3.17 (m, 1H), 3.04-2.98 (m, 2H), 2.33 (s, 3H), 1.66-1.58 (m, 2H), 1.42-1.38 (m, 11H), 1.21 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.77, 163.65 (d, *J* = 250.5 Hz), 157.11, 135.60 (d, *J* = 3.8 Hz), 130.69 (d, *J* = 6.0 Hz), 127.13 (d, *J* = 9.1 Hz), 126.28 (d, *J* = 18.2 Hz), 115.47 (d, *J* = 24.3 Hz), 78.49, 52.78, 44.45, 39.50, 28.03, 27.46, 26.86, 14.77, 13.13 (d, *J* = 3.8 Hz).

(R)-tert-Butyl 8-(4-fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)octanoate (3.2o)
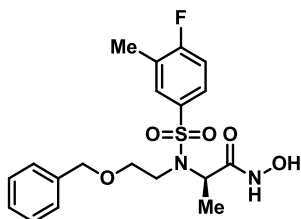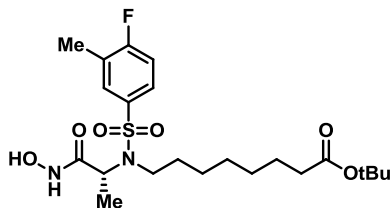

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (130 mg, 48%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.76-7.70 (m, 2H), 7.21 (t, J = 8.8 Hz, 1H), 4.43 (m, 1H), 3.43-3.37 (m, 1H), 3.21-3.14 (m, 1H), 2.33 (s, 3H), 2.20 (t, J = 7.2 Hz, 2H), 1.66-1.53 (m, 4H), 1.43 (s, 9H), 1.29-1.21 (m, 9H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 173.66, 168.81, 163.62 (d, J = 250.5 Hz), 135.72 (d, J = 3.1 Hz), 130.61 (d, J = 6.1 Hz), 127.18 (d, J = 9.1 Hz), 126.23 (d, J = 19.0 Hz), 115.47 (d, J = 24.3 Hz), 79.91, 52.81, 44.78, 35.02, 30.68, 28.66, 28.56, 27.11, 26.42, 24.74, 14.8 (d, J = 3.0 Hz), 13.19. MS (ESI) 475.1 [M + H]$^+$.

(*R*)-2-(4-Fluoro-3-methyl-*N*-(3-nitrobenzyl)phenylsulfonamido)-*N*-hydroxypropanamide (3.2p)

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (19/1, v/v) to yield the desired product as a white foam (92 mg, 73%). ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 10.95 (s, 1H), 8.87 (s, 1H), 8.17 (s, 1H), 8.07 (d, J = 8.4 Hz, 1H), 7.78 (d, J = 8 Hz, 1H), 7.70 (m, 1H), 7.65 (m, 1H), 7.58 (t, J = 8 Hz, 1H), 7.30 (t, J = 9.0 Hz, 1H), 4.80 (d, J = 17.4 Hz, 1H), 4.72 (d, J = 17.4 Hz, 1H), 4.46 (m, 1H), 2.25 (s, 3H),

1.10 (d, $J = 7.2$ Hz, 3.0 H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.8, 164.1 (d, $J = 252.8$ Hz), 148.2, 139.1, 134.6 (d, $J = 3.2$ Hz), 134.2, 130.8 (d, $J = 6.1$ Hz), 129.6, 127.1 (d, $J = 9.9$ Hz), 127.0, 122.9, 122.8, 116.3 (d, $J = 23.5$ Hz), 53.0, 47.5, 14.9, 14.6. MS (ESI) 412.22 [M + H]$^+$.

(R)-2-(4-Fluoro-N-isopropyl-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.2q)

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white foam (25 mg, 30%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.85-7.77 (m, 2H), 7.20 (t, $J = 8.8$ Hz, 1H), 4.19 (q, $J = 7.2$ Hz, 1H), 3.94 (s, $J = 6.8$ Hz, 1H), 2.33 (s, 3H), 1.41 (d, $J = 6.8$ Hz, 3H), 1.27 (m, 6H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 169.3, 163.4 (d, $J = 249.7$ Hz), 137.3 (d, $J = 3.8$ Hz), 131.0 (d, $J = 6.1$ Hz), 127.5 (d, $J = 9.1$ Hz), 125.9 (d, $J = 18.3$ Hz), 115.2 (d, $J = 23.5$ Hz), 52.4, 50.1, 21.0, 20.6, 15.8, 13.0 (d, $J = 3.0$ Hz). MS (ESI) 319.34 [M + H]$^+$.
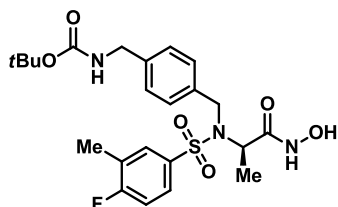
(R)-2-(4-Fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.2r)
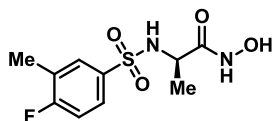
The crude residue was purified by preparatory TLC using DCM/MeOH (12/1, v/v) to yield white solid (45 mg, 45%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.77-7.70 (m, 2H), 7.19 (t, $J = 8.8$ Hz, 1H), 3.76 (brs, 1H), 2.32 (s, 3H), 1.19 (d, $J = 5.2$ Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 169.38, 163.53 (d, $J = 249.7$ Hz), 136.32 (d, $J = 3.0$ Hz), 130.45 (d, $J = 6.1$

Hz), 126.82 (d, J = 9.8 Hz), 126.09 (d, J = 18.9 Hz), 115.32 (d, J = 23.5 Hz), 50.15, 18.05, 13.06 (d, J = 3.1 Hz). MS (ESI) 277.24 [M + H]$^+$.

## (R)-2-(N-(4-(Acetamidomethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.6a)
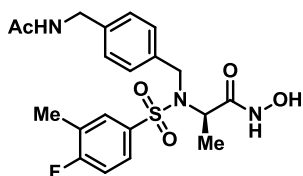
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white solid (10 mg, 7%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.39 (s, 1H), 7.62-7.59 (m, 2H), 7.31 (d, J = 8.0 Hz, 2H), 7.19 (d, J = 8.0 Hz, 2H), 7.14 (t, J = 9.2 Hz, 1H), 4.70 (d, J = 16.4 Hz, 1H), 4.64 (d, J = 16.4 Hz, 1H), 4.51 (q, J = 7.2 Hz, 1H), 4.33 (s, 2H), 2.29 (s, 3H), 1.98 (s, 3H), 1.20 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 171.66, 168.68, 163.60 (d, J = 250.5 Hz), 137.54, 137.48, 135.73 (d, J = 3.8 Hz), 130.78 (d, J = 6.1 Hz), 127.78, 127.17, 127.09, 126.02 (d, J = 18.2 Hz), 115.23 (d, J = 23.5 Hz), 53.09, 42.48, 21.13, 16.34, 13.03 (d, J = 3.8 Hz). MS (ESI) 438.34 [M + H]$^+$.

(R)-2-(N-(3-(Acetamidomethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.6b)
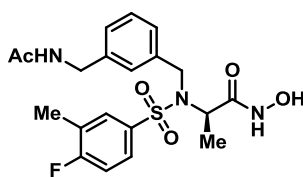
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a white solid (60 mg, 43%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.63-7.60 (m, 2H), 7.26-7.12 (m, 5H), 4.70 (d, J = 16.4 Hz, 1H), 4.63 (d, J = 16.4 Hz, 1H), 4.53 (q, J = 6.8 Hz, 1H), 4.30 (s, 2H), 2.28 (s, 3H), 1.99 (s, 3H), 1.20 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 171.79, 168.66, 163.59 (d, J = 250.5 Hz), 138.68, 138.57, 135.77 (d, J = 3.0 Hz), 130.79 (d, J = 6.1 Hz), 128.00, 127.12 (d, J = 9.8 Hz), 126.71, 126.44, 126.03, 126.03 (d, J = 19.0 Hz), 115.27 (d, J = 24.3 Hz), 53.07, 42.70, 21.16, 16.09, 13.05 (d, J = 3.8 Hz). MS (ESI) 438.34 [M + 1]$^+$.

(R)-2-(N-(4-(2-Acetamidoethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.6c)
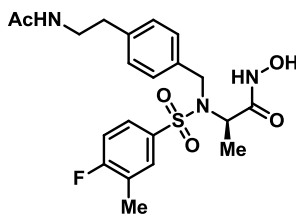
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (10/1, v/v) to yield the desired product as a white solid (40 mg, 20%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.63-7.61 (m, 2H), 7.27 (d, J = 8.0 Hz, 2H), 7.17-7.12 (m, 3H), 4.68 (d, J = 16.4 Hz, 1H), 4.60 (d, J = 16.4 Hz, 1H), 4.51 (q, J = 7.2 Hz, 1H), 3.40-3.35 (m, 2H), 2.75 (t, J = 7.2 Hz, 2H), 2.28 (s, 3H), 1.90 (s, 3H), 1.20 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 171.81, 168.66, 163.57 (d, J = 250.5 Hz), 138.11, 136.32, 135.77 (d, J = 3.8 Hz), 130.78 (d, J = 6.1 Hz), 128.25, 127.80, 127.11 (d, J = 9.2 Hz), 125.99 (d, J = 18.2 Hz), 115.23 (d, J = 23.5 Hz), 53.08, 47.88, 40.59, 34.71, 21.08, 16.12, 13.02 (d, J = 3.8 Hz). MS (ESI) 452.30 [M + H]$^+$.

### (R)-2-(N-(6-Acetamidohexyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.6d)
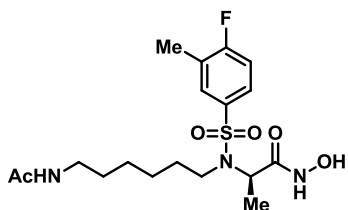
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (12/1, v/v) to yield the desired product as a white solid (60 mg, 30%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.79-7.69 (m, 2H), 7.23 (t, J = 8.8 Hz, 1H), 4.44 (q, J = 7.2 Hz, 1H), 3.47-3.39 (m, 1H), 3.23-3.14 (m, 3H), 2.35 (s, 3H), 1.94 (s, 3H), 1.72-1.59 (m, 2H), 1.53-1.46 (m, 2H), 1.37-1.28 (m, 4H), 1.23 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 171.76, 168.82, 163.62 (d, J = 250.5 Hz), 135.67 (d, J = 3.0 Hz), 130.64 (d, J = 6.0 Hz), 127.09 (d, J = 9.1 Hz), 126.25 (d, J = 18.2 Hz), 115.44 (d, J = 23.5 Hz), 52.73, 44.66, 38.97, 30.58, 28.78, 26.14, 26.09, 21.18, 14.72, 13.09 (d, J = 3.8 Hz). MS (ESI) 418.32 [M + H]$^+$.

(R)-2-(N-(5-Acetamidopentyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.6e)
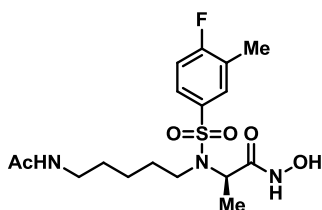
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (12/1, v/v) to yield the desired product as a white solid (15 mg, 18%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.96 (brs, 1H), 7.76-7.68 (m, 2H), 7.21 (t, J = 8.8 Hz, 1H), 4.40 (q, J = 7.2 Hz,

1H), 3.46-3.38 (m, 1H), 3.21-3.12 (m, 3H), 2.34 (s, 3H), 1.92 (s, 3H), 1.71-1.58 (m, 2H), 1.53-1.45 (m, 2H), 1.38-1.26 (m, 2H), 1.21 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 171.81, 168.79, 163.65 (d, J = 249.8 Hz), 135.67 (d, J = 3.1 Hz), 130.65 (d, J = 6.1 Hz), 127.09 (d, J = 9.1 Hz), 126.27 (d, J = 18.3 Hz), 115.43 (d, J = 24.2 Hz), 52.72, 44.56, 38.90, 30.34, 28.47, 23.76, 21.13, 14.64, 13.03 (d, J = 3.8 Hz). MS (ESI) 404.3 [M + H]$^+$.

(R)-2-(N-(3-Acetamidopropyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.6f)
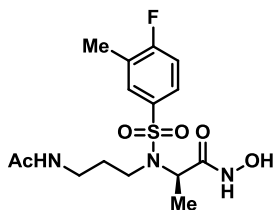
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (10/1, v/v) to yield the desired product as a white solid (5 mg, 8%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.0 (brs, 1H), 7.77-7.68 (m, 2H), 7.22 (t, J = 8.8 Hz, 1H), 4.41 (q, J = 7.2 Hz, 1H), 3.51-3.44 (m, 1H), 3.25-3.16 (m, 3H), 2.34 (s, 3H), 1.94 (s, 3H), 1.89-1.82 (m, 2H), 1.21 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 171.89, 168.76, 163.71 (d, J = 250.5 Hz), 135.45 (d, J = 3.8 Hz), 130.68 (d, J = 6.8 Hz), 127.12 (d, J = 9.1 Hz), 126.33 (d, J = 18.2 Hz), 115.46 (d, J = 24.3 Hz), 52.75, 42.35, 36.62, 30.46, 21.16, 14.58, 13.02 (d, J = 3.8 Hz). MS (ESI) 376.2 [M + H]$^+$.

(R)-8-(4-Fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)-N-methyloctanamide (3.8)
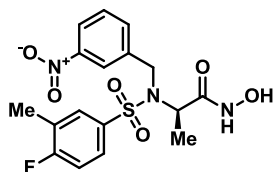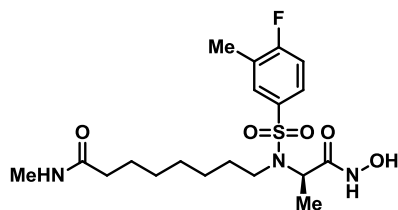

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (12/1, v/v) to yield the desired product as a white solid (15 mg, 30%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.87 (brs, 1H), 7.76-7.67 (m, 2H), 7.22 (t, J = 8.8 Hz, 1H), 4.41 (q, J = 6.8 Hz, 1H), 3.45-3.37 (m, 1H), 3.20-3.12 (m, 1H), 2.71 (s, 3H), 2.34 (s, 3H), 2.16 (t, J = 7.2 Hz, 2H), 1.69-1.57 (m, 4H), 1.30-1.21 (m, 9H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 175.46, 168.83, 163.63 (d, J = 250.4 Hz), 135.70 (d, J = 3.0 Hz), 130.64 (d, J = 6.8 Hz), 127.09 (d, J = 9.9 Hz), 126.24 (d, J = 18.2 Hz), 115.41 (d, J = 23.6 Hz), 52.71, 44.72, 35.54, 30.67, 28.72, 28.50, 26.34, 25.47, 24.88, 14.75, 13.05 (d, J = 3.8 Hz). MS (ESI) 432.1.0 [M + H]$^+$.

(R)-2-(N-(3-Acetamidobenzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.11)
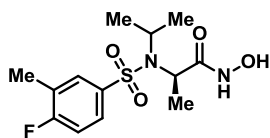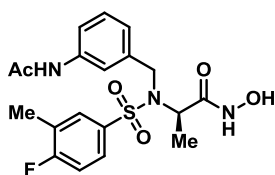

The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (10/1, v/v) to yield the desired product as a white solid (65 mg, 33%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.57 (s, 1H), 8.49 (s, 1H), 7.62-7.61 (m, 2H), 7.16-6.95 (m, 4H), 4.60 (d, J = 16.0 Hz, 1H), 4.48 (m, 1H), 4.25 (d, J = 16.0 Hz, 1H), 2.27 (s, 3H), 1.98 (s, 3H), 1.11 (d,

J = 6.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 170.25, 167.93, 163.87 (d, J = 252.7 Hz), 138.12, 137.34, 134.84 (d, J = 3.0 Hz), 130.82 (d, J = 6.1 Hz), 128.88, 127.08 (d, J = 9.1 Hz), 126.75 (d, J = 18.2 Hz), 123.97, 120.14, 119.79, 116.09 (d, J = 23.5 Hz), 53.09, 48.16, 24.14, 14.55 (d, J = 3.8 Hz), 13.88. MS (ESI) 424.26 [M + H]$^+$.

Procedure A for Boc Deprotection (3.4a-3.4c)

Boc-protected amines (1 equiv.) were treated with 4 N HCl in dioxane (2 mL) at rt. After 1 h, the solvent was removed under reduced pressure and the resulting residue was triturated with diethyl ether (3 x 3 mL) to yield desired product.

(R)-2-(N-(3-(Aminomethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide hydrochloride (3.4a)
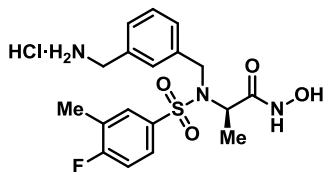
White solid (180 mg, 96%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.70-7.68 (m, 2H), 7.46-7.36 (m, 4H), 7.18 (t, J = 8.8 Hz, 1H), 4.76 (d, J = 16.4 Hz, 1H), 4.64-4.57 (m, 2H), 4.08 (s, 2H), 2.29 (s, 3H), 1.19 (d, J = 6.4 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.78, 163.71 (d, J = 250.5 Hz), 139.69, 135.37 (d, J = 3.0 Hz), 133.0, 130.84 (d, J = 6.8 Hz), 128.76, 128.26, 128.09, 127.53, 127.27 (d, J = 9.9 Hz), 126.27 (d, J = 18.2 Hz), 115.52 (d, J = 23.5 Hz), 53.04, 42.97, 15.83, 13.19 (d, J = 3.8 Hz). MS (ESI) 396.29 [M + H]$^+$.

(R)-2-(N-(4-(2-Aminoethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide hydrochloride (3.4b)
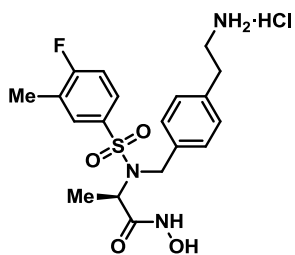
White solid (69 mg, 99%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.67-7.64 (m, 2H), 7.34 (d, J = 6.8 Hz, 2H), 7.22-7.15 (m, 3H), 4.72 (d, J = 16.4 Hz, 1H), 4.60-4.53 (m, 2H), 3.18-3.14 (m, 2H), 2.96-2.92 (m, 2H), 2.30 (s, 3H), 1.19 (d, J = 6.0 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.67, 163.67 (d, J = 249.7 Hz), 137.55, 135.56 (d, J = 3.1 Hz), 135.37, 130.79 (d, J = 6.9 Hz), 128.29, 128.16, 127.17 (d, J = 9.1 Hz), 126.16 (d, J = 18.2 Hz), 115.37 (d, J = 24.3 Hz), 52.92, 40.56, 32.73, 15.75, 13.06 (d, J = 3.0 Hz). MS (ESI) 410.31 [M + H]$^+$.

(R)-2-(N-(6-Aminohexyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide hydrochloride (3.4c)
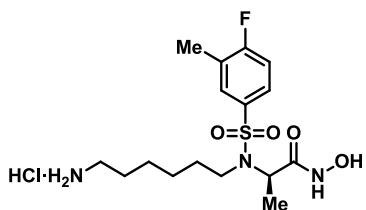
White solid (157 mg, 85%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.74-7.66 (m, 2H), 7.21 (t, J = 8.8 Hz, 1H), 4.38 (q, J = 7.6 Hz, 1H), 3.46-3.11 (m, 2H) 2.90 (t, J = 7.6 Hz, 2H), 2.32 (s, 3H), 1.70-1.62 (m, 4H), 1.42-1.33 (m, 4H), 1.17 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.92, 163.68 (d, J = 250.5 Hz), 135.59 (d, J = 3.8 Hz), 130.62 (d, J = 6.1 Hz), 127.09 (d, J = 9.1 Hz), 126.35 (d, J = 19.0 Hz), 115.52 (d, J = 23.6 Hz), 52.68,

44.20, 39.17, 30.04, 26.64, 23.11, 14.54, 13.08 (d, J = 3.8 Hz). MS (ESI) 376.33 [M + H]$^+$.

Procedure B for Boc or tButyl Deprotection (3.4d-3.4g)

To a solution of protected amine (1 equiv.) in DCM (0.2 M) at 0 °C was added a solution of TFA and triethylsilane (5/1, v/v). The TFA and triethylsilane were premixed at 0 °C prior to addition. After 1 h, the solvent was removed under reduced pressure yielding the desired products.

(R)-2-(N-(5-Aminopentyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide TFA salt (3.4d)
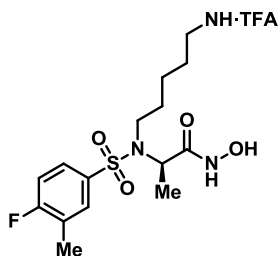
White foam (26 mg, 50%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.76-7.70 (m, 2H), 7.68 (t, J = 8.8 Hz, 1H), 4.41 (q, J = 7.2 Hz, 1H), 3.49-3.41 (m, 1H), 3.24-3.17 (m, 1H), 2.92 (t, J = 7.6 Hz, 2H), 2.34 (s, 3H), 1.74-1.63 (m, 4H), 1.43-1.36 (m, 2H), 1.19 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.89, 163.65 (d, J = 250.5 Hz), 135.56 (d, J = 3.0 Hz), 130.58 (d, J = 6.1 Hz), 127.04 (d, J = 9.1 Hz), 126.35 (d, J = 18.3 Hz), 115.49 (d, J = 24.3 Hz), 52.63, 44.15, 39.11, 30.01, 26.63, 23.08, 14.45, 13.03 (d, J = 3.8 Hz). MS (ESI) 362.2 [M + H]$^+$.

(R)-2-(N-(4-Aminobutyl)-4-fluoro-3-methylphenylsulfonamido)-N-

hydroxypropanamide TFA salt (3.4e)
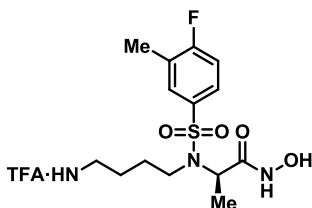
White foam (15 mg, 17%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.76-7.68 (m, 2H), 7.22 (t, J = 8.8 Hz, 1H),4.38 (q, J = 6.8 Hz, 1H), 3.51-3.45 (m, 1H), 3.28-3.21 (m, 1H), 2.97-2.93 (m, 2H), 2.34 (s, 3H), 1.77-1.64 (m, 4H), 1.19 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.82, 163.74 (d, J = 250.5 Hz), 135.41 (d, J = 3.8 Hz), 130.59 (d, J = 6.1 Hz), 127.04 (d, J = 9.9 Hz), 126.40 (d, J = 18.3 Hz), 115.53 (d, J = 24.3 Hz), 52.60, 43.68, 38.94, 27.56, 24.32, 14.50, 13.03 (d, J = 3.8 Hz). MS (ESI) 348.1 [M + H]$^+$.

(R)-2-(N-(3-Aminopropyl)-4-fluoro-3-methylphenylsulfonamido)-N-

hydroxypropanamide TFA salt (3.4f)
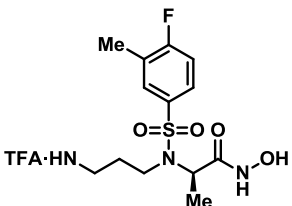
White foam (52 mg, 42%).^{1}H NMR (400 MHz, CD$_3$OD) δ 7.79-7.70 (m, 2H), 7.24 (t, J = 8.4 Hz, 1H),4.39 (q, J = 6.8 Hz, 1H), 3.60-3.53 (m, 1H), 3.39-3.31 (m, 1H), 3.07-3.01 (m, 2H), 2.34 (s, 3H), 2.09-2.01 (m, 2H), 1.19 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 168.91, 163.86 (d, J = 250.5 Hz), 134.98 (d, J = 3.8 Hz), 130.66 (d, J = 6.9 Hz), 127.11 (d, J = 9.1 Hz), 126.54 (d, J = 18.2 Hz), 115.63 (d, J = 24.3 Hz), 52.83, 41.32, 36.88, 28.53, 14.54, 13.02 (d, J = 3.8 Hz). MS (ESI) 334.1 [M + H]$^+$.

**(R)-8-(4-Fluoro-N-(1-(hydroxyamino)-1-oxopropan-2-yl)-3-methylphenylsulfonamido)octanoic acid (3.4g)**
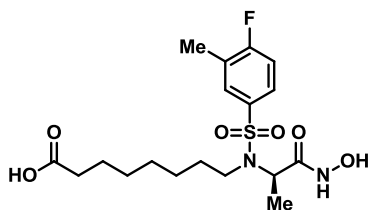
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH/AcOH (19/1/0.1, v/v) to yield the desired product as a white solid (19 mg, 20%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.67-7.62 (m, 2H), 7.14 (t, J = 8.8 Hz, 1H), 4.42 (m, 1H), 3.39-3.31 (m, 1H), 3.08-3.02 (m, 1H), 2.36-2.32 (m, 5H), 1.61-1.57 (m, 4H), 1.30-1.26 (m, 6H), 1.14 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 179.19, 169.01, 163.88 (d, J = 252.8 Hz), 134.91 (d, J = 3.8 Hz), 130.66 (d, J = 6.8 Hz), 126.91 (d, J = 9.1 Hz), 126.75, 116.17 (d, J = 24.3 Hz), 52.77, 44.87, 33.88, 29.69, 28.52, 28.44, 26.41, 24.41, 14.65 (d, J = 3.1 Hz), 13.06. MS (ESI) 419.0 [M + H]$^+$.

General Procedure for Hydrogenation (3.9, 3.3a)

To a solution of nitroarene (1 equiv.) in MeOH or DCM (0.2 M) under H$_2$ was added palladium on activated carbon (Pd/C) (200 mg, 10 wt. %). The reaction was shaken in a Parr apparatus at 4 atm at rt. After 4 h, the reaction mixture was filtered through a Celite pad, and the filtrate was concentrated under reduced pressure.

(R)-Methyl 2-(N-(3-aminobenzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.9)
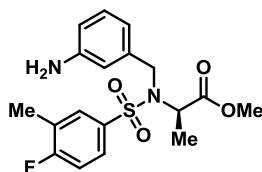
The crude residue was purified over SiO_2 using an eluent of EtOAc/hexane (1/4, v/v) to give the desired product as a colorless oil. (277 mg, 77%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.66-7.65 (m, 2H), 7.11-7.01 (m, 2H), 6.67-6.54 (m, 3H), 4.61 (q, J = 7.6 Hz, 1H), 4.48 (d, J = 16.0 Hz, 1H), 4.29 (d, J = 16.0 Hz, 1H), 3.67 (brs, 2H), 3.49 (s, 3H), 2.31 (s, 3H), 1.31 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.64, 163.60 (d, J = 251.9 Hz), 146.73, 138.29, 135.74 (d, J = 3.8 Hz), 131.15 (d, J = 6.0 Hz), 129.24, 127.32 (d, J = 9.1 Hz), 126.02 (d, J = 18.2 Hz), 117.93, 115.57 (d, J = 23.5 Hz), 114.542, 114.29, 55.29, 52.03, 49.31, 16.64, 14.56 (d, J = 3.8 Hz).

**(R)-2-(N-(4-Aminobenzyl)-4-fluoro-3-methylphenylsulfonamido)-N-hydroxypropanamide (3.3a)**
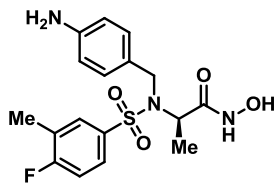
The crude residue was purified by preparatory TLC using an eluent of DCM/MeOH (19/1, v/v) to give the desired product as yellow solid (40 mg, 43%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.60-7.56 (m, 2H), 7.11-7.08 (m, 3H), 6.62-6.60 (m, 2H), 4.51-4.41 (m, 2H), 4.12-4.09 (m, 1H), 2.29 (s, 3H), 1.18 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 168.42, 163.77 (d, J = 252.7 Hz), 146.10, 135.33, 130.86 (d, J = 6.8 Hz), 130.10, 127.0

(d, J = 9.1 Hz), 126.62 (d, J = 18.2 Hz), 125.62, 116.0 (d, J = 23.5 Hez), 115.45, 53.28, 48.20, 14.58 (d, J = 3.8 Hz), 14.0. MS (ESI) 380.27 [M - H]⁻.

Procedure A for Acetylation (3.10)

To a solution of **3.9** (1 equiv.) in DCM (0.15 M) was added TEA (10 equiv.), then AcCl (10 equiv.) at 0 °C. The reaction was gradually allowed to warm to rt. After 16 h, the reaction mixture was concentrated under reduced pressure, and the resulting residue was re-suspended in EtOAc (15 mL). The organic layer was washed with sat. NH₄Cl (5 mL), brine (5 mL), dried over Na₂SO₄, filtered, and concentrated under reduced pressure.

(R)-Methyl 2-(N-(3-acetamidobenzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.10)
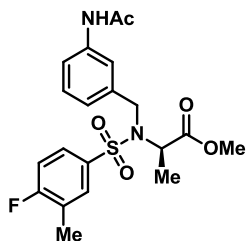
The crude product was purified over SiO₂ using an eluent of EtOAc/Hex (1/1, v/v) to yield the desired product as a yellow solid (200 mg, 65%). ¹H NMR (400 MHz, CDCl₃) δ7.65-7.63 (m, 3H), 751-7.46 (m, 2H), 7.22 (t, J = 7.6 Hz, 1H), 7.12-7.03 (m, 2H), 4.65 (q, J = 7.2 Hz, 1H), 4.57 (d, J = 16.4 Hz, 1H), 4.36 (d, J = 16.4 Hz, 1H), 3.48 (s, 3H), 2.31 (s, 3H), 2.15 (s, 3H), 1.30 (d, J = 7.6 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 171.57, 168.60, 163.70 (d, J = 252.0 Hz), 138.35, 138.31, 135.43 (d, J = 3.7 Hz), 131.10 (d, J = 6.1 Hz),

129.01, 127.33 (d, J = 9.9 Hz), 126.18 (d, J = 18.2 Hz), 123.46, 119.04, 119.0, 115.66 (d, J = 23.5 Hz), 55.45, 52.10, 48.09, 24.55, 16.86, 14.58 (d, J = 3.8 Hz).

Procedure B for Acetylation (3.5a-3.5f)

To a solution of amine (1 equiv.) in THF (0.3 M) was added TEA (1.2 equiv.), followed by DMAP (0.1 equiv.) and acetic anhydride (1.2 equiv) at 0 °C. The flask was allowed to gradually warm to rt, and the reaction was stirred overnight. After 16 h, the reaction mixture was evaporated under reduced pressure. The resulting residue was re-suspended in EtOAc (15 mL), and the organic solution was washed with 1 N HCl (5 mL), sat. aq. NaHCO$_3$ (5 mL), and brine (5 mL), dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure. The crude product was purified over SiO$_2$ using an eluent of DCM/MeOH/TEA (33/1/0.3, v/v) to yield the desired product.

(R)-Methyl 2-(N-(4-(acetamidomethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.5a)
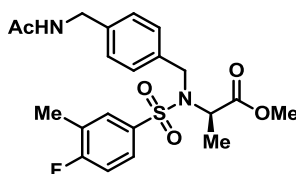
Colorless oil (204 mg, 95% over two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.64-7.6 (m, 2H), 7.27 (d, J = 8 Hz, 2H), 7.19 (d, J = 8 Hz, 2H), 7.1 (t, J = 8.4 Hz, 1H), 6.48 (bs, 1H), 4.61 (q, J = 7.2 Hz, 1H), 4.55 (d, J = 16.4 Hz, 1H) 4.38-4.35 (m, 3H), 3.47 (s, 3H), 2.32 (s, 3H), 1.98 (s, 3H), 1.29 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.5, 170.3, 163.7 (d, J = 252.8 Hz), 137.9, 136.5, 135.5 (d, J = 3.8 Hz), 131.1 (d, J = 6.1 Hz),

128.1, 127.8, 127.3 (d, J = 9.1 Hz), 126.2 (d, J = 18.2 Hz), 115.7 (d, J = 23.6 Hz), 55.3, 52.1, 49.0, 43.2, 23.1, 16.8, 14.6 (d, J = 3.0 Hz).

(R)-Methyl 2-(N-(3-(acetamidomethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.5b)
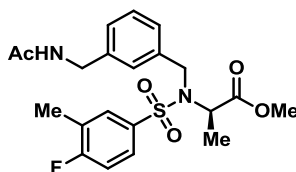
Colorless oil (91 mg, 80% over two steps)^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65-7.61 (m, 2H), 7.25-7.23 (m, 3H), 7.17 (m, 1H), 7.1 (t, J = 8 Hz, 1H), 6.1 (bs, 1H), 4.63 (q, J = 7.6 Hz, 1H), 4.54 (d, J = 16.4 Hz, 1H) 4.41-4.36 (m, 3H), 3.47 (s, 3H), 2.32 (s, 3H), 2.03 (s, 3H), 1.30 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.6, 170.0, 163.7 (d, J = 252.0 Hz), 138.6, 137.7, 135.5 (d, J = 3.0 Hz), 131.1 (d, J = 6.1 Hz), 128.8, 127.3 (d, J = 9.1 Hz), 127.17, 126.99, 126.1 (d, J = 18.2 Hz), 115.7 (d, J = 23.5 Hz), 55.3, 52.1, 49.0, 43.5, 23.2, 16.7, 14.6 (d, J = 3.8 Hz).

(R)-Methyl 2-(N-(4-(2-acetamidoethyl)benzyl)-4-fluoro-3-methylphenylsulfonamido)propanoate (3.5c)
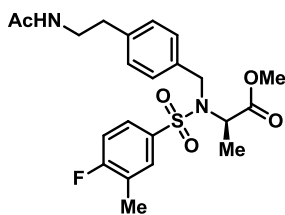
Colorless oil (195 mg, 90% over two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.67-7.63 (m, 2H), 7.25 (d, J = 7.6 Hz, 2H), 7.14-7.09 (m, 3H), 6.19 (s, 1H), 4.63 (q, J = 7.2 Hz, 1H), 4.54 (d, J = 16.0 Hz, 1H), 4.39 (d, J = 16.0 Hz, 1H), 3.48-3.43 (m, 5H), 2.79 (t, J = 7.6

Hz, 2H), 2.32 (s, 3H), 1.93 (s, 3H), 1.31 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$)

δ 171.49, 170.27, 163.64 (d, J = 252.0 Hz), 138.45, 135.54 (d, J = 3.8 Hz), 135.27,

131.06 (d, J = 6.8 Hz), 128.72, 128.19, 127.29 (d, J = 9.1 Hz), 126.13 (d, J = 18.2 Hz),

115.67 (d, J = 23.5 Hz), 55.20, 52.03, 48.93, 40.70, 35.26, 23.17, 16.60, 14.58 (d, J =

3.0 Hz).

(R)-Methyl 2-(N-(6-acetamidohexyl)-4-fluoro-3-

methylphenylsulfonamido)propanoate (3.5d)
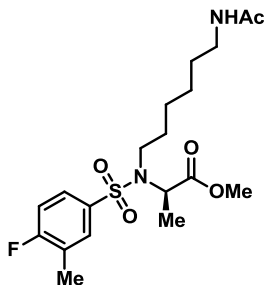
Colorless oil (190 mg, 89% over two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.69-7.64 (m,

2H), 7.12 (t, J = 8.8 Hz, 1H), 6.25 (s, 1H), 4.60 (m, 1H), 3.56 (s, 3H), 3.26-3.08 (m, 4H),

2.34 (s, 3H), 1.97 (s, 3H), 1.70-1.24 (m, 11H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.84,

170.31, 163.59 (d, J = 251.2 Hz), 135.56 (d, J = 3.8 Hz), 130.89 (d, J = 6.8 Hz), 127.18

(d, J = 9.1 Hz), 126.10 (d, J = 18.3 Hz), 115.63 (d, J = 23.6 Hz), 55.20, 52.13, 45.78,

39.39, 30.85, 29.40, 26.47, 26.41, 23.20, 16.72, 14.60 (d, J = 3.0 Hz).

(R)-Methyl 2-(N-(5-acetamidopentyl)-4-fluoro-3-

methylphenylsulfonamido)propanoate (3.5e)
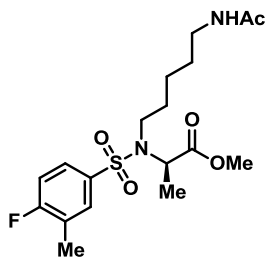
Colorless oil (85 mg, 80% over two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.68-7.61 (m, 2H), 7.11 (t, J = 8.8 Hz, 1H), 5.92 (s, 1H), 4.60 (q, J = 7.2 Hz, 1H), 3.56 (s, 3H), 3.26-3.19 (m, 3H), 3.15-3.09 (m, 1H), 2.34 (s, 3H), 1.97 (s, 3H), 1.74-1.47 (m, 4H), 1.41 (d, J = 7.2 Hz, 3H), 1.35-1.30 (m, 2H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.81, 170.21, 163.62 (d, J = 252.0 Hz), 135.48 (d, J = 3.1 Hz), 130.90 (d, J = 6.0 Hz), 127.16 (d, J = 9.1 Hz), 126.14 (d, J = 19.0 Hz), 115.63 (d, J = 23.5 Hz), 55.19, 52.15, 45.63, 39.39, 30.57, 29.07, 24.07, 23.28, 16.71, 14.61 (d, J = 3.1 Hz).

(R)-Methyl 2-(N-(3-acetamidopropyl)-4-fluoro-3-

methylphenylsulfonamido)propanoate (3.5f)
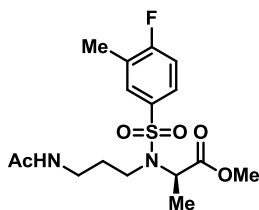
Colorless oil (50 mg, 33% over two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.67-7.62 (m, 2H), 7.13 (t, J = 8.8 Hz, 1H), 6.37 (s, 1H), 4.56 (q, J = 7.2 Hz, 1H), 3.57 (s, 3H), 3.32-3.25 (m, 4H), 2.34 (s, 3H), 1.99 (s, 3H), 1.89-1.78 (m, 2H), 1.39 (d, J = 7.6 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.65, 170.31, 163.66 (d, J = 252.0 Hz), 135.10 (d, J = 3.0

Hz), 130.79 (d, J = 6.1 Hz), 127.08 (d, J = 9.1 Hz), 126.29 (d, J = 18.2 Hz), 115.73 (d, J = 23.5 Hz), 55.14, 52.20, 43.13, 36.47, 30.35, 23.22, 16.34, 14.56 (d, J = 3.0 Hz).

Synthesis of Methyl Amide (3.7).

(R)-Methyl 2-(4-fluoro-3-methyl-N-(8-(methylamino)-8-oxooctyl)phenylsulfonamido)propanoate (3.7)
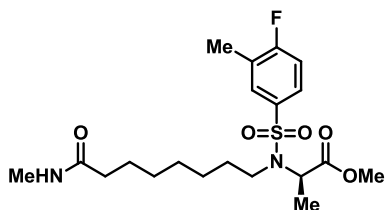
To a solution of **3.1n** (97 mg, 0.2 mmol) in DCM (3 mL) was added TFA (1 mL). The reaction mixture was stirred at room temperature. After 2.5 h, the solvent was removed under reduced pressure, and the resulting residue was dissolved in DCM (15 mL). The organic layer was washed with 1 N HCl (5 mL), and brine (5 mL), dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure to give crude product which was subjected to the next reaction without further purification. To a solution of the crude product, 1-hydroxybenzotriazole (25 mg, 0.19 mmol), 4-methylmorpholine (95 μL, 0.85 mmol), and methylamine hydrochloride (35 mg, 0.5 mmol) in THF (2 mL) was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (46 mg, 0.24 mmol), and the reaction mixture was stirred overnight. After 16 h, the reaction mixture was diluted with water (10 mL) and extracted with DCM (3 x 15 mL). The combined organic extracts are washed with brine (5 mL), dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure. The crude product was purified over SiO$_2$ using an eluent of DCM/MeOH (50/1, v/v) to give **3.7** as colorless oil (50 mg, 71%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.68-7.63 (m, 2H), 7.11 (t, J = 8.8 Hz, 1H), 5.89 (brs, 1H), 4.61 (q, J = 6.8 Hz, 1H), 3.56 (s,

3H), 3.26-3.18 (m, 1H), 3.12-3.04 (m, 1H), 2.8 (s, 3H), 2.34 (s, 3H), 2.16 (t, J = 7.6 Hz, 2H), 1.71-1.52 (m, 4H), 1.42 (d, J = 7.2 Hz, 3H), 1.28-1.25 (m, 6H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 173.86, 171.87, 163.57 (d, J = 252.0 Hz), 135.55 (d, J = 3.8 Hz), 130.91 (d, J = 6.8 Hz), 127.18 (d, J = 9.1 Hz), 126.06 (d, J = 18.2 Hz), 115.59 (d, J = 23.5 Hz), 55.17, 52.14, 45.87, 36.50, 30.89, 29.11, 28.83, 26.67, 26.24, 25.60, 16.74, 14.62 (d, J = 3.8 Hz).

Synthesis of Non-Commercial Starting Materials

tert-Butyl 4-(hydroxymethyl)phenethylcarbamate (3.12)
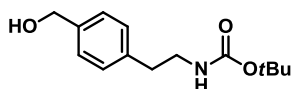
To a solution of [4-(2-Amino-ethyl)-phenyl]-methanol hydrochloride (200 mg, 1.06 mmol) in DMF (1 mL) containing TEA (163 µL, 1.17 mmol) was added Boc anhydride (256 mg, 1.17 mmol) at 0 °C. The reaction was allowed to gradually warm to rt, and was stirred overnight. After 16 h, the reaction was concentrated under reduced pressure. The crude residue was purified over SiO$_2$ using an eluent of EtOAc/Hexane (1/3, v/v) to give the desired product as colorless oil (234 mg, yield 88%). ^{1}H NMR (400 MHz, CDCl$_3$) 7.26 (d, J = 8.0 Hz, 2H), 7.13 (d, J = 8.0 Hz, 2H), 4.74 (m, 1H), 4.6 (s, 2H), 3.31 (m, 2H), 2.98 (s, 1H), 2.75 (t, J = 6.8 Hz, 2H), 1.42 (s, 9H). ^{13}C NMR (100 MHz, CDCl$_3$) 156.00, 139.24, 138.14, 128.83, 127.24, 79.24, 64.68, 41.78, 35.78, 28.38.

tert-Butyl 4-(bromomethyl)phenethylcarbamate (3.13).
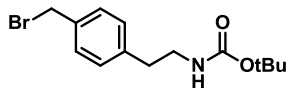
To a solution of **3.12** (234 mg, 0.93 mmol) and triphenylphosphine (367 mg, 1.4 mmol) in DCM (2 mL) was added carbon tetrabromide (463 mg, 1.4 mmol) at 0°C. After 1 h, the reaction was concentrated under reduced pressure. The crude residue was purified over SiO_2 using an eluent of EtOAc/Hexane (1/5, v/v) to give the desired product as a white solid (247 mg, yield 84%).^{1}H NMR (400 MHz, $CDCl_3$) 7.32 (d, J = 8.0 Hz, 2H), 7.16 (d, J = 8.0 Hz, 2H), 4.64 (m, 1H), 4.48 (s, 2H), 3.36 (m, 2H), 2.79 (t, J = 6.8 Hz, 2H), 1.44 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) 155.81, 139.46, 135.85, 129.23, 79.20, 41.62, 35.92, 33.43, 28.39.

tert-Butyl 8-bromooctanoate (3.14).
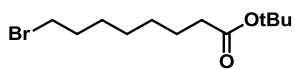
To a solution of 8-bromooctanoic acid (200 mg, 0.9 mmol) in DCM (2 mL) was added TFAA (280 μL, 2.0 mmol) dropwise at 0 °C. After 2.5 h, tBuOH (310 μL, 3.2 mmol) was slowly added. After 1 h reaction warmed to rt. After 2.5 h the reaction was quenched with H_2O (5 mL) and extracted with Et_2O (4 x 15 mL). The combined organic layers were washed with brine and dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified over SiO_2 using an eluent of EtOAc/Hex (1/10, v/v) to give **3.14** as colorless oil (240 mg, 96%).^{1}H NMR (400 MHz, $CDCl_3$) δ 3.39 (t, J = 6.8 Hz, 2H), 2.2 (t, J = 7.2 Hz, 2H), 1.85 (p, J = 6.8 Hz, 2H), 1.58 (t, J = 6.8 Hz, 2H), 1.44 (s, 11H), 1.33 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.91, 79.73, 35.36, 33.64, 32.65, 28.77, 28.34, 28.12, 27.93, 24.86.

Chapter 4

**IDENTIFICATION OF A LIGAND-INDUCED BINDING POCKET IN ANTHRAX TOXIN
LETHAL FACTOR**

Adapted with permission from:

4.1 ACKNOWLEDGEMENTS

This chapter includes a description of work performed in collaboration with Kimberly M. Maize, Barry C. Finzel, Jonathan Solberg, Jon E. Hawkinson, and Rodney L. Johnson. Jonathan Solberg tested the compounds *in vitro* and Jon E. Hawkinson analyzed the data. X-ray crystallization studies were completed by Kimberly M. Maize under the direction of Barry C. Finzel. Synthetic guidance was provided by Rodney L. Johnson.

4.2 INTRODUCTION

As a consequence of the work reported in chapter 3, we produced analog **4.17**, which shows a unique binding mode compared to the rest of the compounds. Instead of occupying the S2' subsite like the other N-alkyl substituents, the isobutyl substituent occupies the S1′ subsite, causing the 4-fluoro-3-methylphenyl group to position in a new pocket designated S1′*. The S1′* site is created by the movement of Lys656 and Leu677. This conformational change results in a solvent exposed tunnel that may be used to develop potent and novel LF inhibitors (**Figure 4.1**). I have synthesized analogs of **4.17**, with modifications to the 4-fluoro-3-methylphenyl substituent, in an effort to take advantage of this solvent exposed tunnel.

Figure 4.1. LF active site with co-crystallized **4.17** (PDB ID 4XM6). The circle denotes the site of substitution in **4.17**. (Schrödinger Maestro Discovery Suite 9.4). Surface colored by residue charge. Orange = neutral residues, Red = negatively charged residues, and Blue = positively charged residues.

4.17; IC$_{50}$ = 90 μM
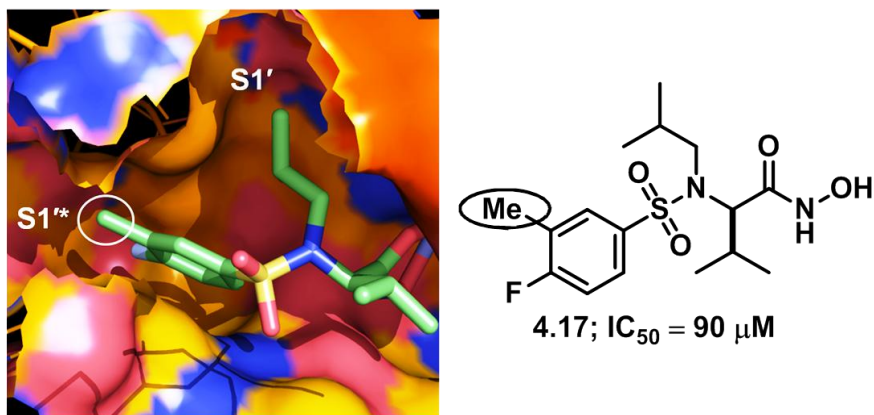
4.3 MATERIALS AND METHODS

4.3.1 Synthesis, Biochemical Evaluation, and Structural Biology

Analog synthesis was accomplished as outlined in **Scheme 4.1**. A generalized synthetic route was fashioned for all the designed analogs. Intermediate sulfonamides **4.5-4.7** were readily synthesized from commercially available D-valine *tert*-butyl ester hydrochloride and commercially available sulfonyl chlorides **4.2-4.4** through nucleophilic substitution. Alkylation of sulfonamides **4.5-4.7** with 1-iodo-2-methylpropane was carried out under basic conditions to afford tertiary sulfonamides **4.8-4.10** in 56-66% yield. tert-Butoxycarbonyl deprotection of analogues **4.8-4.10** was achieved with TFA/DCM (1:2, v/v) in nearly quantitative yields. The resulting carboxylic acids **4.11-4.13** were submitted to EDC coupling to yield THP-protected hydroxamic acids **4.14-4.16** in 30-97% yield. The THP group was removed with TFA/DCM (1:2, v/v) to yield **4.17-4.19** in 31-57% yield. Synthesized compounds were evaluated for LF inhibition utilizing the previously

described FRET assay. Crystal structures were obtained as previously described in Chapter 2.

Scheme 4.1. Synthesis of compounds **4.17-4.19**
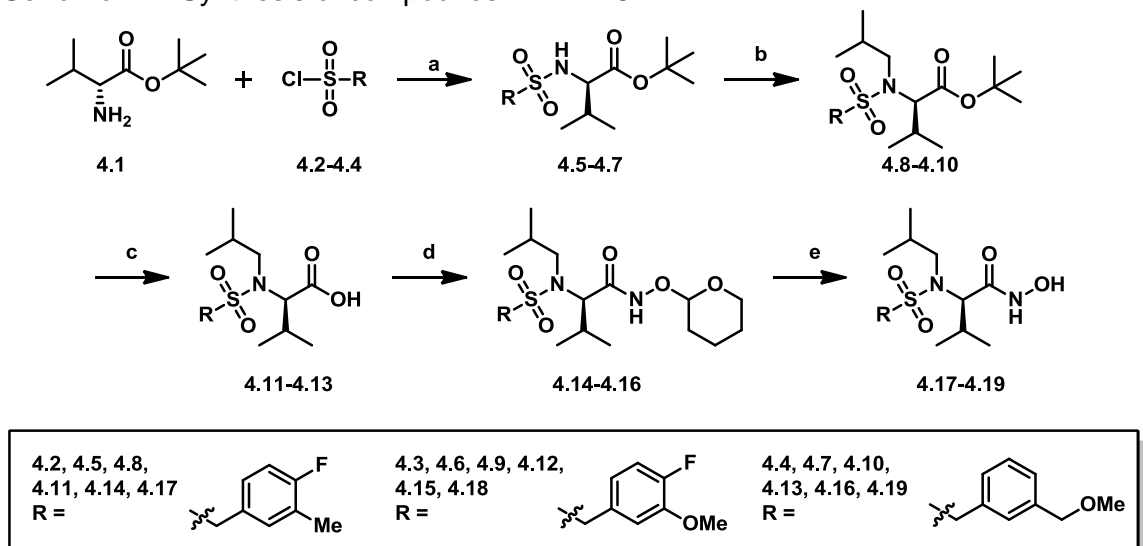
Reagents and conditions: (a) K_2CO_3, dioxane/H_2O (1:1), rt (**4.5**, 94%; **4.6**, 29%; **4.7**, 79%); (b) 1-iodo-2-methylpropane, K_2CO_3, DMF, rt (**4.8**, 59%; **4.9**, 66%; **4.10**, 56%) (c) TFA/DCM (1:2, v/v) (**4.11**, 97%; **4.12**, 98%; **4.13**, 99%); (d) THPONH$_2$, NMM, HOBt, EDC, DMF, rt (**4.14**, 30%; **4.15**, 90%; **4.16**, 97%); (e) TFA/DCM (1:2, v/v) (**4.17**, 50%; **4.18**, 31%; **4.19**, 57%).

4.4 RESULTS AND DISCUSSION

4.4.1 Biochemical Evaluation and Structural Biology

Both analogs of **4.17** (**4.18** and **4.19**) were co-crystallized with LF and show similar binding modes to **4.17** (**Figures 4.2** and **4.3**). Specifically, we can see that the 3-methoxyphenyl group of **4.18** and the 3-methoxymethylphenyl group of **4.19** target our newly discovered, ligand-induced S1'* tunnel formation.

101

Figure 4.2. LF active site with co-crystallized **4.18** (PDB ID 4XM7). The circle denotes site of modification in **4.18**. (Schrödinger Maestro Discovery Suite 9.4). Surface colored by residue charge. Orange = neutral residues, Red = negatively charged residues, and Blue = positively charged residues.

4.18; IC$_{50}$ = 600 μM
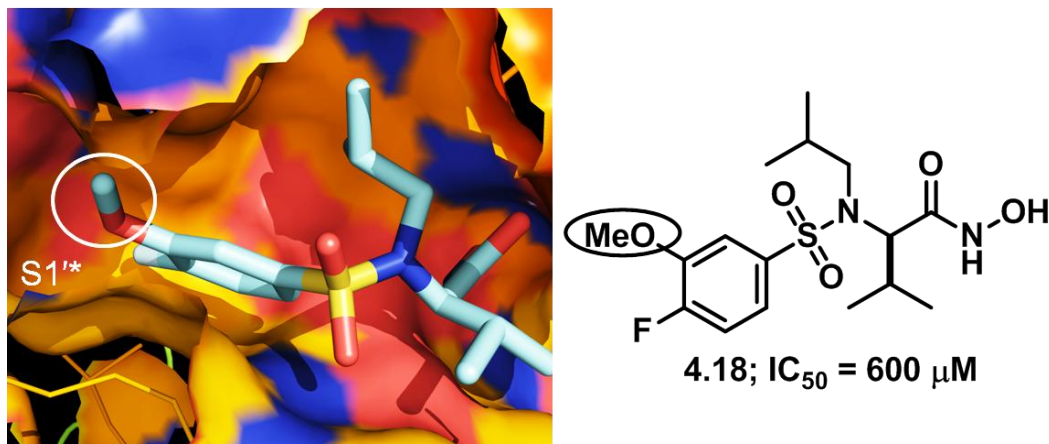
Figure 4.3. LF active site with co-crystallized **4.19** (PDB ID 4XM8). The circle denotes site of modification in **4.19**. (Schrödinger Maestro Discovery Suite 9.4). Surface colored by residue charge. Orange = neutral residues, Red = negatively charged residues, and Blue = positively charged residues.

4.19; IC$_{50}$ = 740 μM
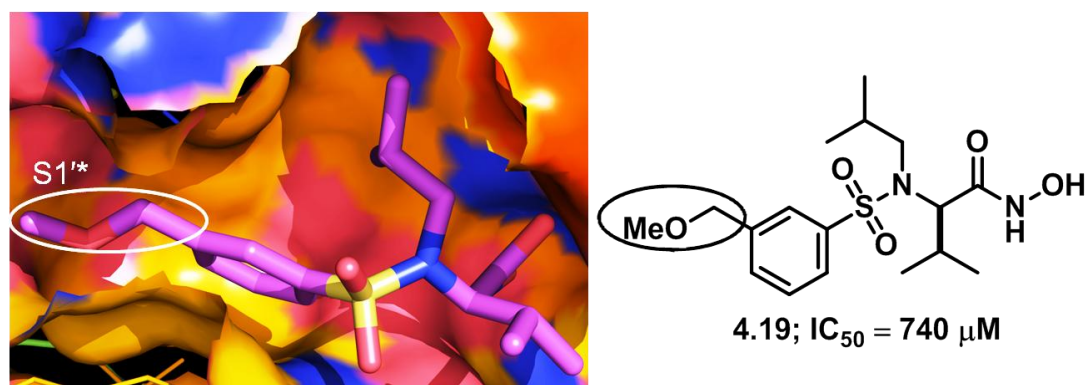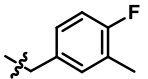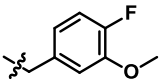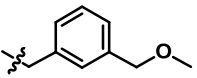
Table 4.1 presents FRET assay results for **4.17-4.19**. Unfortunately, the potency of **4.18** and **4.19** did not improve compared to **4.17**. This is because the 3-methoxy group of **4.18** and the 3-methoxymethyl group of **4.19** did not engage in H-bond interactions with residues in the S1'* tunnel. Nevertheless, this study shows that the S1' subsite is more flexible than previously thought and capable of accommodating relatively

102

large hydrophobic modifications. This ligand-induced conformational change in the S1′ subsite of LF had not been demonstrated before this study. The virtual screening of ~11,000,000 compounds against 4XM6 will be described in the following chapter.

Table 4.1. LF FRET assay results for **4.17-4.19**.

GPHR #[a]	PDB ID	Cpd #	R	LF IC$_{50}$ (μM)[b]
00223405	4XM6	**4.17**		90
00227098	4XM7	**4.18**		600
00278888	4XM8	**4.19**		740

[a]GPHR # is a compound number in our in-house collection of compounds
[b]IC$_{50}$ is a half maximal inhibitory concentration

4.5 CONCLUDING REMARKS

In this study, we identified a ligand-induced binding pocket in the S1′ site of LF designated S1′*, which is created by the movement of Lys656 and Leu677. This conformational change results in a solvent exposed tunnel that we expect may be used to develop potent and novel LF inhibitors in future. To note, this was the first time this ligand-induced conformational change was reported with LF inhibitors. In order to take advantage of the S1′* site, we have synthesized two analogs of **4.17** bearing polar moieties at the 3-phenyl position. Co-crystal structures of these compounds were obtained and show that the modifications in both **4.18** and **4.19** occupy the newly discovered S1′* solvent exposed tunnel. The potency of these analogs did not increase compared to **4.17** due to the inability of the 3-methoxy group of **4.18** and the 3-methoxymethyl group of **4.19** to engage in H-bond interactions with the residues in the

103

S1'* tunnel. Although we were unable to obtain more potent compounds, these studies show that the S1'* site may be used in future design to develop potent LF inhibitors with a novel binding mode.

4.6 EXPERIMENTAL

General Synthesis Information. Chemical reagents were purchased from commercial sources and used without additional purification. Bulk solvents were purchased from Fisher Scientific and anhydrous *N,N'*-dimethylformamide (DMF) was purchased from EMD Chemicals. Reactions were performed under an atmosphere of dry N_2 unless otherwise noted. Silica gel chromatography was performed on self-packed columns with SiliaFlash 60Å silica gel (SiliCycle). Preparatory thin layer chromatography (TLC) was performed on plates with glass backed SiliaPlate 60Å silica gel (SiliCycle). HPLC analyses were performed on an Agilent 1100 series instrument equipped with a diode array detector and a Zorbax SB-C18 column (0.5 x 150 mm, 5 µm, Agilent Technologies). LC-MS analyses were performed on an Agilent 1100 series instrument equipped with an Agilent MSD SL Ion Trap mass spectrometer (positive-ion mode) and a Zorbax SB-C18 column (0.5 x 150 mm, 5 µm, Agilent Technologies). The analysis method (15 µL/min flow rate) involved isocratic 10% MeCN (containing 0.1% TFA) in ddH_2O (containing 0.1% HCO_2H; 0 to 2 mins) followed by a linear gradient of 10% to 90% MeCN (containing 0.1% TFA) in ddH_2O (containing 0.1% HCO_2H; 2 to 24 mins), and isocratic 90% MeCN (containing 0.1% TFA) in ddH_2O (containing 0.1% HCO_2H; 24-26 mins). The column was heated to 40 °C. Wavelengths monitored = 254 nm and 215 nm. Nuclear magnetic resonance (NMR) spectra were recorded in $CDCl_3$, CD_3OD, or DMSO-d_6 on a Varian instrument operating at 400 MHz (for ^{1}H) and 100 MHz (for ^{13}C) at

ambient temperature. Chemical shifts are reported in parts per million (ppm) and normalized to internal solvent peaks or tetramethylsilane (0 ppm).

Representative Procedure for Synthesis of 4.5-4.7

(R)-tert-Butyl 2-(4-fluoro-3-methylphenylsulfonamido)-3-methylbutanoate (4.5)
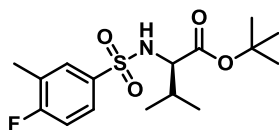
D-Valine *tert*-butyl ester hydrochloride (**4.1**, 0.6 g, 2.9 mmol) was added to a solution of K_2CO_3 (0.8 g, 6.0 mmol) in dioxane/water (6 mL, 1:1, v/v). To this, a solution of 4-fluoro-3-methylphenyl-sulfonylchloride (**7**, 0.5 g, 2.4 mmol) in dioxane (1 mL) was added immediately after with vigorous stirring. The mixture was stirred at rt overnight. Upon consumption of the starting material as determined by TLC, the solvent was reduced to one-third the reaction volume under reduced pressure. The aq. layer was extracted with EtOAc (3 × 20 mL). The combined organic layers were washed with sat. aq. NH_4Cl (1 × 20 mL) and brine (1 × 20 mL), dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude product was purified over SiO_2 using an eluent of EtOAc/hexane (1/5, v/v) to yield **4.5** as a white solid (0.8 g, 94%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77-7.72 (m, 2H), 7.10 (t, J = 8.4 Hz, 1H), 5.77 (d, J = 10.0 Hz, 1H, NH), 3.66 (dd, J = 10.0 Hz, J = 4.4 Hz, 1H), 2.30 (s, 3H), 2.11-2.0 (m, 1H), 1.26 (s, 9H), 1.0 (d, J = 6.8 Hz, 3H), 0.88 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.12, 163.46 (d, J = 251.2 Hz), 135.72 (d, J = 3.8 Hz), 130.91 (d, J = 6.8 Hz), 127.26 (d, J = 9.9 Hz), 125.91 (d, J = 18.2 Hz), 115.43 (d, J = 23.5 Hz), 82.0, 61.31, 31.49, 27.46, 18.93, 17.02, 14.24 (d, J = 3.8 Hz).

(R)-tert-Butyl 2-(4-fluoro-3-methoxyphenylsulfonamido)-3-methylbutanoate (4.6)
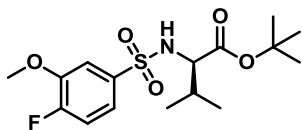
White solid (202 mg, 29%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.50-7.44 (m, 2H), 7.16 (t, J = 8.8 Hz, 1H), 5.57 (d, J = 10.4 Hz, 1H, NH), 3.11 (s, 3H), 3.64 (dd, J = 10.0 Hz, J = 4.4 Hz, 1H), 2.10-2.02 (m, 1H), 1.26 (s, 9H), 1.0 (d, J = 6.8 Hz, 3H), 0.87 (d, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 170.17, 154.73 (d, J = 252.8 Hz), 147.97 (d, J = 11.4 Hz), 136.04 (d, J = 3.8 Hz), 120.61 (d, J = 8.4 Hz), 116.14 (d, J = 19.8 Hz), 112.41 (d, J = 3.0 Hz), 82.28, 61.36, 56.37, 31.55, 27.58, 18.98, 17.02.

(R)-tert-Butyl 2-(3-(methoxymethyl)phenylsulfonamido)-3-methylbutanoate (4.7)
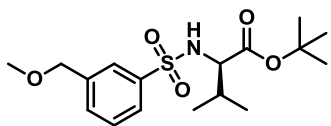
White solid (250 mg, 79%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.84 (s, 1H), 7.78 (d, J = 7.6 Hz, 1H), 7.53 (d, J = 8.0 Hz, 1H), 7.46 (t, J = 8.0 Hz, 1H), 5.54 (d, J = 10.0 Hz, 1H, NH), 4.49 (s, 2H), 3.65 (dd, J = 9.6 Hz, J = 4.4 Hz, 1H), 3.39 (s, 3H), 2.08-2.01 (m, 1H), 1.22 (s, 9H), 0.99 (d, J = 6.8 Hz, 3H), 0.86 (d, J = 6.4 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 170.12, 140.03, 139.77, 131.39, 128.99, 126.30, 126.02, 82.12, 73.47, 61.26, 58.22, 31.58, 27.57, 18.99, 17.05.

Representative Procedure for N-Alkylation (4.8-4.10)

(R)-tert-Butyl 2-(4-fluoro-N-isobutyl-3-methylphenylsulfonamido)-3-methylbutanoate (4.8)
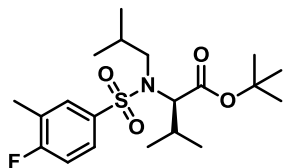
To a solution of **4.5** (0.4 g, 1.0 mmol) and 1-iodo-2-methylpropane (0.3 mL, 3.0 mmol) in anhydrous DMF (2.0 mL) was added K_2CO_3 (0.7 g, 5.0 mmol). The reaction mixture was stirred at room temperature. After 48 h, the solvent was removed under reduced pressure and the resulting residue was taken up in H_2O, extracted with EtOAc (3 × 15 mL), dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude product was purified over SiO_2 using an eluent of EtOAc/hexane (1/10, v/v) to yield **4.8** as a colorless oil (240 mg, 59%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72-7.68 (m, 2H), 7.09 (t, J = 8.4 Hz, 1H), 4.0 (d, J = 10.4 Hz, 1H), 3.27_{CHA} (dd, J = 14.8 Hz, J = 6.0 Hz, 1H), 3.05_{CHB} (dd, J = 14.8 Hz, J = 8.8 Hz, 1H), 2.33 (s, 3H), 2.14-2.05 (m, 2H), 1.32 (s, 9H), 1.06 (d, J = 6.8 Hz, 3H), 0.93 (d, J = 6.8 Hz, 3H), 0.85 (t, J = 6.8 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.27, 163.39 (d, J = 251.3 Hz), 136.58 (d, J = 3.8 Hz), 131.13 (d, J = 6.0 Hz), 127.46 (d, J = 9.1 Hz), 125.73 (d, J = 18.2 Hz), 115.38 (d, J = 23.5 Hz), 81.72, 67.15, 53.20, 29.1, 27.87, 27.74, 20.57, 20.26, 19.86, 19.40, 14.53 (d, J = 3.8 Hz).

(R)-tert-Butyl 2-(4-fluoro-N-isobutyl-3-methoxyphenylsulfonamido)-3-methylbutanoate (4.9)
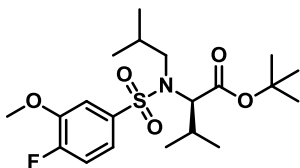
Colorless oil (153 mg, 66%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.49-7.43 (m, 2H), 7.16 (t, J = 8.4 Hz, 1H), 3.99-3.96 (m, 4H), 3.26 $_{CHA}$ (dd, J = 14.4 Hz, J = 6.0 Hz, 1H), 3.05 $_{CHB}$ (dd, J = 14.4 Hz, J = 8.8 Hz, 1H), 2.13-2.04 (m, 2H), 1.30 (s, 9H), 1.06 (d, J = 6.8 Hz, 3H), 0.94 (d, J = 6.8 Hz, 3H), 0.87 (d, J = 7.2 Hz, 6H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 169.17, 154.64 (d, J = 252.8 Hz), 147.68 (d, J = 11.4 Hz), 136.92 (d, J = 3.8 Hz), 120.85 (d, J = 7.6 Hz), 116.14 (d, J = 19.0 Hz), 113.13 (d, J = 3.0 Hz), 81.9, 67.46, 56.53, 53.34, 29.16, 28.0, 27.76, 20.6, 20.37, 19.93, 19.43.

(R)-tert-Butyl 2-(N-isobutyl-3-(methoxymethyl)phenylsulfonamido)-3-methylbutanoate (4.10)
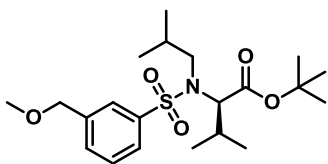
Colorless oil (163 mg, 56%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.82-7.78 (m, 2H), 7.53 (d, J = 8.0 Hz, 1H), 7.47 (t, J = 7.6 Hz, 1H), 4.50 (s, 2H), 4.03 (d, J = 10.8 Hz, 1H), 3.39 (s, 3H), 3.26 $_{CHA}$ (dd, J = 14.8 Hz, J = 6.4 Hz, 1H), 3.07 $_{CHB}$ (dd, J = 14.0 Hz, J = 8.8 Hz, 1H), 2.14-2.04 (m, 2H), 1.29 (s, 9H), 1.04 (d, J = 6.4 Hz, 3H), 0.93 (d, J = 6.8 Hz, 3H), 0.85 (d, J = 6.4 Hz, 3H), 0.82 (d, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 169.32, 141.08, 139.45, 131.16, 128.82, 126.66, 126.30, 81.70, 73.62, 67.12, 58.23, 53.14, 29.03, 27.79, 27.74, 20.59, 20.24, 19.84, 19.42.

Representative Procedure for *tert*-Butyl Deprotection (4.11-4.13)

To an ice-chilled solution of **4.8** (0.2 g, 0.6 mmol) in DCM (2.0 mL) was added TFA (1.0 mL) dropwise. The reaction mixture was stirred for 6 hours at rt. Upon consumption of the starting material as determined by TLC, the solvent was removed under reduced pressure, and the product was submitted to the next reaction without further purification.

Representative Procedure for the Synthesis of 4.14-4.16

(2R)-2-(4-Fluoro-N-isobutyl-3-methylphenylsulfonamido)-3-methyl-N-((tetrahydro-2H-pyran-2-yl)oxy)butanamide (4.14)
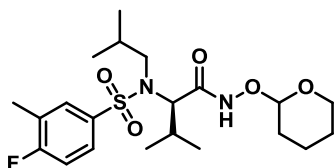
To a solution of **4.11** (130 mg, 0.38 mmol) in DMF (2 mL), HOBt (62 mg, 0.46 mmol), NMM (125 µL, 1.14 mmol), THPONH$_2$ (137 mg, 1.17 mmol) and EDC (102 mg, 0.53 mmol) were added. The mixture was stirred at rt overnight. Upon consumption of the starting material as determined by TLC, H$_2$O (10 mL) was added. The aq. layer was extracted with EtOAc (3 × 20 mL).The combined organic layers were washed with saturated solution of NaHCO$_3$ (1 × 20 mL), brine (1 × 20 mL), dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure. The crude product was purified over SiO$_2$ using an eluent of EtOAc/hexane (1/4, v/v) to yield a colorless oil (50 mg, 30%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.43 (s, 1H), 7.72-7.68 (m, 2H), 7.09 (t, *J* = 8.4 Hz, 1H), 4.95 (s, 1H), 4.0-3.90 (m, 1H), 3.69-3.54 (m, 2H), 3.30-3.21 (m, 1H), 3.01-2.96 (m, 1H), 2.33 (s, 3H), 2.25-2.05 (m, 2H), 1.92-1.6 (m, 6H), 0.93-0.81 (m, 9H), 0.5-0.39 (m, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.10, 163.39 (d, *J* = 251.3 Hz), 136.58 (d, *J* = 3.8 Hz),

131.13 (d, J = 6.0 Hz), 127.46 (d, J = 9.1 Hz), 125.73 (d, J = 18.2 Hz), 115.38 (d, J = 23.5 Hz), 102.1, 63.67, 62.13, 52.74, 27.83, 27.35, 26.36, 24.99, 20.38, 20.15, 19.93, 19.25, 18.28, 14.53 (d, J = 3.8 Hz).

(2R)-2-(4-Fluoro-N-isobutyl-3-methoxyphenylsulfonamido)-3-methyl-N-((tetrahydro-2H-pyran-2-yl)oxy)butanamide (4.15)
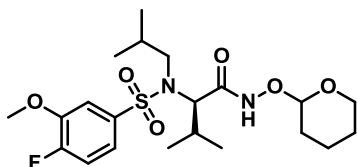
Colorless oil (160 mg, 90%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.43 (s, 1H), 7.72-7.68 (m, 2H), 7.09 (t, J = 8.4 Hz, 1H), 4.95 (s, 1H), 4.01-3.94 (m, 4H), 3.72-3.62 (m, 2H), 3.33-3.27 (m, 1H), 3.03-2.98 (m, 1H), 2.24-2.05 (m, 2H), 1.92-1.6 (m, 6H), 0.93-0.81 (m, 9H), 0.62-0.52 (m, 3H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.10, 163.39 (d, J = 251.3 Hz), 136.58 (d, J = 3.8 Hz), 131.13 (d, J = 6.0 Hz), 127.46 (d, J = 9.1 Hz), 125.73 (d, J = 18.2 Hz), 115.38 (d, J = 23.5 Hz), 102.1, 63.67, 62.13, 56.54, 52.74, 27.83, 27.35, 26.36, 24.99, 20.38, 20.15, 19.93, 19.25, 18.28.

((2R)-2-(N-Isobutyl-3-(methoxymethyl)phenylsulfonamido)-3-methyl-N-((tetrahydro-2H-pyran-2-yl)oxy)butanamide (4.16)
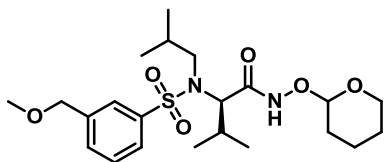
Colorless oil (173 mg, 97%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.13 (s, 1H), 7.43-7.37 (m, 2H), 7.16-7.11 (m, 2H), 4.57 (s, 1H), 4.12 (s, 2H), 3.71-3.52 (m, 1H), 3.35-3.23 (m, 2H), 3.02 (s, 3H), 2.93-2.89 (m, 1H), 2.65-2.50 (m, 1H), 1.84-1.67 (m, 2H), 1.44-1.21 (m, 6H),

0.55-0.42 (m, 9H), 0.13-0.01 (3, H). ^{13}C NMR (100 MHz, CDCl$_3$) δ 167.06, 140.45, 140.02, 131.63, 129.11, 126.33, 125.96, 101.98, 63.56, 62.0, 60.29, 58.24, 52.70, 27.83, 27.35, 26.36, 24.99, 20.38, 20.15, 19.93, 19.25, 18.28.

Representative Procedure for THP Deprotection (4.17-4.19)

(R)-2-(4-Fluoro-N-isobutyl-3-methylphenylsulfonamido)-N-hydroxy-3-methylbutanamide (4.17)
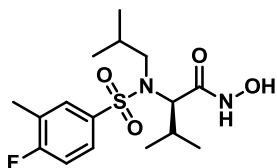
To an ice-chilled solution of **4.14** (0.05 g, 0.11 mmol) in DCM (2.0 mL) was added TFA (1.0 mL) dropwise. The reaction mixture was stirred overnight at rt. Upon consumption of the starting material as determined by TLC, the solvent was removed under reduced pressure and the resulting residue was purified by preparatory TLC using DCM/MeOH (19/1, v/v) to yield a white solid (0.02 g, 50%). The HPLC retention time is 18.7 min. Compound purity is >99% as determined by two-wavelength HPLC analysis (254 nm and 215 nm). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.77-7.68 (m, 2H), 7.20 (t, J = 8.8 Hz, 1H), 3.74 (d, J = 11.2 Hz, 1H), 3.45 $_{CHA}$ (dd, J = 14.4 Hz, J = 6.8 Hz, 1H), 3.0 $_{CHB}$ (dd, J = 14.8 Hz, J = 8.0 Hz, 1H), 2.34 (s, 3H), 2.16-2.08 (m, 2H), 0.89-0.82 (m, 12H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 167.06, 163.56 (d, J = 250.4 Hz), 136.1 (d, J = 3.8 Hz), 130.79 (d, J = 6.8 Hz), 127.24 (d, J = 9.9 Hz), 126.07 (d, J = 18.9 Hz), 115.18 (d, J = 23.5 Hz), 64.10, 52.70, 27.96, 27.67, 19.49, 19.2, 18.83, 18.30, 12.99 (d, J = 3.8 Hz). MS (ESI) 361.20 [M + H]$^+$.

111

(R)-2-(4-Fluoro-N-isobutyl-3-methoxyphenylsulfonamido)-N-hydroxy-3-methylbutanamide (4.18)
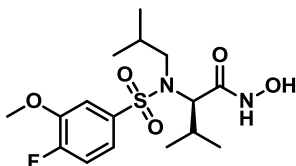
White solid (40 mg, 31%). The HPLC retention time is 18.1 min. Compound purity is 94% by 254 nm and 97% by 215 nm as determined by HPLC analysis. ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.50-7.43 (m, 2H), 7.25 (t, J = 8.4 Hz, 1H), 3.95 (s, 3H), 3.8 (d, J = 10.8 Hz, 1H), 3.46 $_{CHA}$ (dd, J = 14.8 Hz, J = 6.8 Hz, 1H), 3.0 $_{CHB}$ (dd, J = 14.4 Hz, J = 8.0 Hz, 1H), 2.18-2.1 (m, 2H), 0.91-0.80 (m, 12H). ^{13}C NMR (100 MHz, CD$_3$OD) δ 167.16, 154.73 (d, J = 252.7 Hz), 148.13 (d, J = 10.6 Hz), 136.68 (d, J = 3.8 Hz), 120.71 (d, J = 8.3 Hz), 115.78 (d, J = 19.8 Hz), 112.14 (d, J = 3.1 Hz), 64.19, 55.68, 52.75, 27.91, 27.59, 19.54, 19.23, 18.83, 18.32. MS (ESI) 377.10 [M + H]$^+$.

(R)-N-Hydroxy-2-(N-isobutyl-3-(methoxymethyl)phenylsulfonamido)-3-methylbutanamide (4.19)
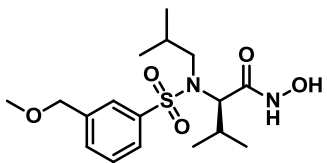
White solid (80 mg, 57%). The HPLC retention time is 15.5 min. Compound purity is 97% by 254 nm and 86% by 215 nm as determined by HPLC analysis. ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.83 (s, 1H), 7.78 (d, J = 7.2 Hz, 1H), 7.60 (d, J = 7.6 Hz, 1H), 7.52 (t, J = 7.6 Hz, 1H), 4.53 (s, 2H), 3.78 (d, J = 10.4 Hz, 1H), 3.47-3.38 (m, 4H), 3.02 $_{CHB}$ (dd, J = 14.4 Hz, J = 7.6 Hz, 1H), 2.16-2.10 (m, 2H), 0.91-0.80 (m, 12H). ^{13}C NMR (100 MHz,

CD$_3$OD) δ 167.14, 140.51, 139.79, 131.50, 128.81, 126.29, 125.99, 73.21, 64.06, 57.28, 52.69, 27.88, 27.61, 19.59, 19.27, 18.85, 18.39. MS (ESI) 373.10 [M + H]$^+$.

Chapter 5

LARGE-SCALE VIRTUAL SCREENING TO IDENTIFY COMPOUNDS THAT TARGET THE NEWLY DISCOVERED S1'* SUBSITE OF ANTHRAX TOXIN LETHAL FACTOR

5.1 ACKNOWLEDGEMENTS

5.1 INTRODUCTION

High-throughput screening (HTS) is defined as the rapid experimental evaluation of large chemical libraries (10,000-100,000 compounds tested per day) to identify modulators of a validated drug target. Many pharmaceutical and biotechnology companies use this approach to fuel their drug discovery programs.[152] As a result, fourteen drugs that were approved between 1991 and 2008 were originally discovered through HTS.[153] Nonetheless, there are several limitations associated with HTS, including the high cost of infrastructure, long screening times, high rates of false positives, low rates for actual hits, and the limited chemical space represented by commercially-available HTS libraries.[154,155] Virtual screening, on the other hand, is devoid of such limitations (although it certainly demonstrates others) and can be used to screen large chemical libraries very efficiently. When a crystal structure is available for a protein of interest, the most common virtual screening strategy is docking and scoring.[156,157] A comprehensive review by Kubinyi describes many successful examples of docking and scoring-based virtual screens that resulted in drug candidates and approved drugs.[158]

Overview of Glide: In this study, we used Glide 5.9,[114–117] available through the Schrödinger Maestro Discovery Suite 9.4 (Schrödinger, Inc.),[118] to screen large

compound libraries with the intent of identifying novel inhibitors of anthrax toxin lethal factor. At the preprocessing stage, Glide computes and generates a receptor grid that represents the shape and properties of the protein target's active site. In the next step, Glide generates exhaustive conformations of each ligand and prescreens the small molecules for promising ligand poses, eliminating conformations with high-energies not suitable for binding to a receptor. These initial ligand poses (100-400) are energy minimized in the active site of the protein using the OPLS-AA force field.[121] Then, three to six of the lowest-energy poses for each ligand are subjected to the Monte Carlo procedure for nearby torsional minima exploration. Once minimized conformations for each molecule are obtained, affinity scoring functions are applied that are designed to predict biological activity through the evaluation of protein-ligand interactions. Binding affinities for the minimized ligand poses are calculated and rank-ordered using Schrödinger's proprietary GlideScore scoring function. GlideScore estimates the free energy of binding (ΔG) of a ligand to a target, which is related to its binding affinity by $\Delta G = -RT \ln K_A$, where binding affinity is represented by $K_A = 1/K_d = \frac{[EI]}{[E][I]}$, and EI is an enzyme-inhibitor complex. The GlideScore function incorporates van der Waals energy, Coulomb energy, a lipophilic term, a hydrogen-bonding term, a metal-binding term, a polar-hydrophobic term, a penalty for freezing rotatable bonds, and solvation terms. The best pose for each ligand, however, is selected using an Emodel score, which gives more weighting to electrostatic and van der Waals energies. In other words, the Emodel score is used to identify the best pose of a ligand, whereas GlideScore ranks the best poses among ligands.[115]

Ligands can be docked to protein targets using three modes: high - throughput virtual screening (HTVS) mode, standard precision (SP) mode, and extra-precision (XP)

mode. HTVS and SP modes use the same scoring function; however, SP mode samples ligand conformations more thoroughly than HTVS mode. As a result, SP mode is more accurate, but slower than HTVS mode. Primarily, HTVS mode is used in the docking and scoring of millions of compounds, whereas SP mode is used in the docking and scoring of thousands of compounds. The slowest and most accurate mode is XP mode.[116] It uses even more extensive conformational ligand sampling than SP mode. XP mode also uses a more elaborate scoring function than SP GlideScore. XP GlideScore includes van der Waals energy, Coulomb energy, a lipophilic term, a hydrogen-bonding term, rewards for π-π stacking and π-cation interactions, for Cl or Br in a hydrophobic environment that pack against Asp or Glu, and for hydrophobic enclosure. The XP GlideScore also penalizes polar atom burial, desolvation, intra-ligand contacts, ligands with large hydrophobic contacts but low H-bond scores, exposed hydrophobic ligand groups, and freezing rotatable bonds.[114,116] Primarily, XP mode is used for docking hundreds of compounds, especially during the hit-to-lead optimization stage of a drug discovery program.

In our previous work, we synthesized and co-crystallized **4.17**, an analog of **MK-31**, which shows a different binding mode than our other N-alkylated analogs. This crystal structure will be published in due course as PDB ID 4XM6. Instead of occupying the S2' subsite, the isobutyl substituent causes a conformational change in the S1' subsite, resulting in a solvent exposed tunnel (**Figure 4.1**). Using 4XM6**,** we docked and scored ~11 million drug-like compounds from the ZINC[159] database in order to discover novel binders of the S1′* subsite. Compounds predicted to have high inhibitory potency while targeting the modified S1′ subsite were subsequently purchased and evaluated for LF inhibition.

117

5.3 MATERIALS AND METHODS

5.3.1 Protein and Ligand Preparation
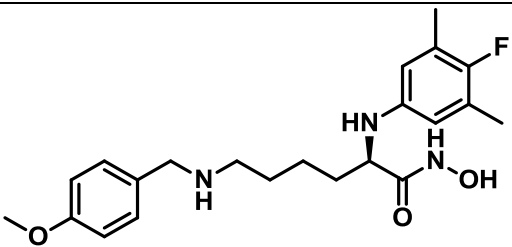
4XM6 was prepared for docking studies using the protein preparation wizard in Schrödinger's Maestro Discovery Suite 9.4 (Schrödinger, Inc.) as previously discussed. We selected the ZINC database compiled by Shoichet *et al.*[159] for virtual screening due to the diversity and drug-like properties of these compounds, which were also pre-filtered to pass Lipinski's[160] rules. Compounds with PSA > 150 and > 7 rotational bonds were also excluded. Compounds that passed these filtering criteria were energy minimized and protonated at pH 7.0 by Shoichet group.[161]
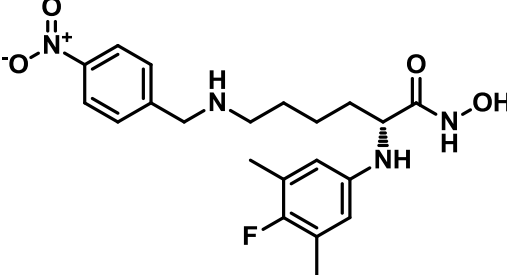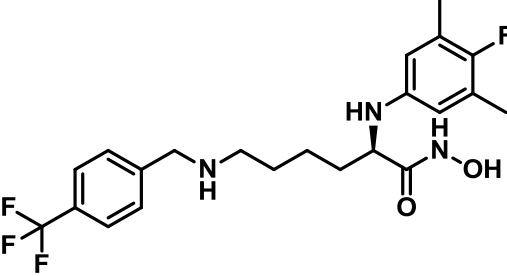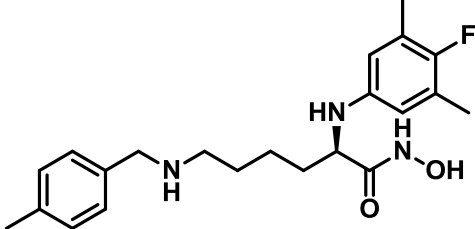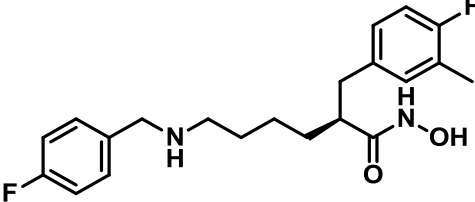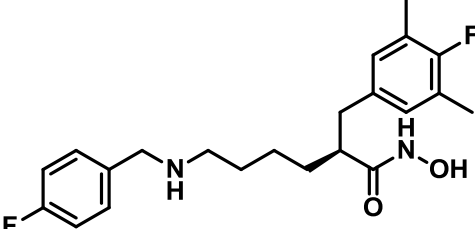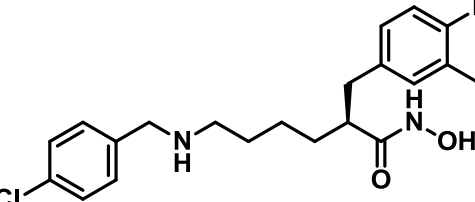
5.3.2 Docking Method Development
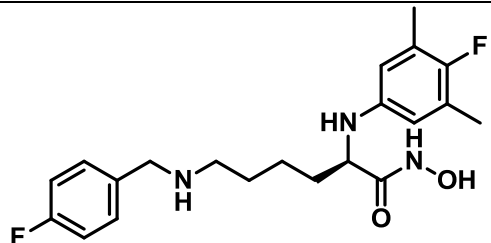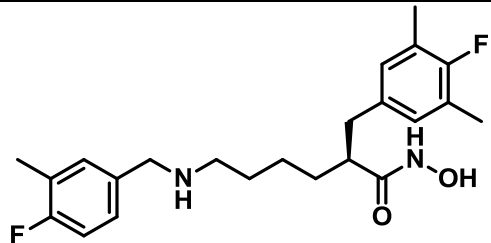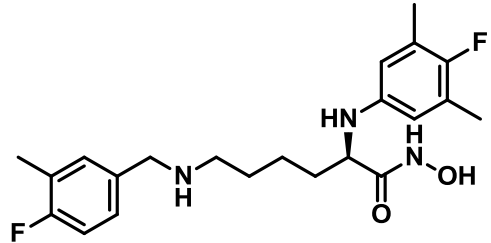
To maximize accuracy while minimizing computational cost, we first docked all the compounds (~11 million) using Glide in HTVS mode and refined the top scoring compounds using the SP and XP modes. We evaluated two specific docking methodologies as outlined below using the 1YQY crystal structure.[110]

For this study, we created database 1 (**DB1**) which included ten of the most potent LF inhibitors[92,162] from published literature (**Table 5.1**) and 10,000 decoy compounds randomly picked from the ZINC[159] database.

Table 5.1. Ten known LF inhibitors for initial docking and scoring studies[92,162]

Compound ID	Structure	IC$_{50}$ (nM)
5.1		0.44

118

5.2		0.3
5.3		0.47
5.4		0.41
5.5		0.25
5.6		0.13
5.7		0.36

5.8		0.24
5.9		0.39
5.10		0.05

In methodology I, we docked all 10010 compounds in **DB1** to 1YQY in HTVS mode, then took the top 10% (1000) of the docked compounds and redocked them in SP mode. The top 1% (10) of the compounds were then docked in XP mode. In methodology II, we docked all 10010 compounds in HTVS mode, then took the top 10% (1000) of the compounds and docked them in SP mode. The top 10% (100) of those docked compounds were then docked in XP mode. The only difference between these two methods was the number of compounds evaluated in XP docking mode. To compare these two strategies, we calculated the enrichment factors[163] (EF) for each methodology to compare docking accuracy. EF values are computed as follows[163]:

$$EF = \frac{\text{Hits}_{\text{sampled}}/N_{\text{sampled}}}{\text{Hits}_{\text{total}}/N_{\text{total}}}$$

Methodology I was able to recover four active compounds with an EF of 400:

$$EF \; = \; \frac{4/10}{10/10010} = 400$$

Methodology II also returned four active compounds with an EF of 400:

$$EF \; = \; \frac{4/10}{10/10010} = 400$$

For comparison, the HTVS docking mode was only able to recover one active compound with an EF = 100:

$$EF \; = \; \frac{1/10}{10/10010} = 100$$
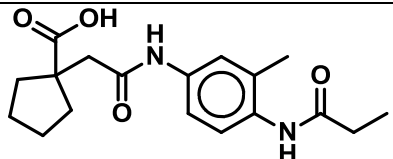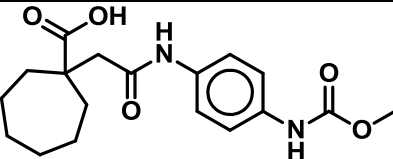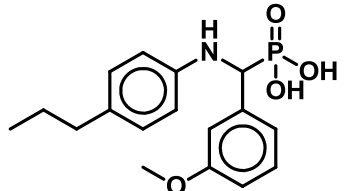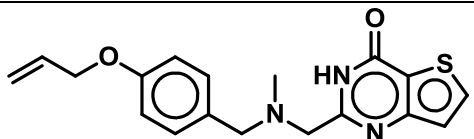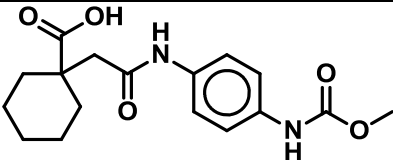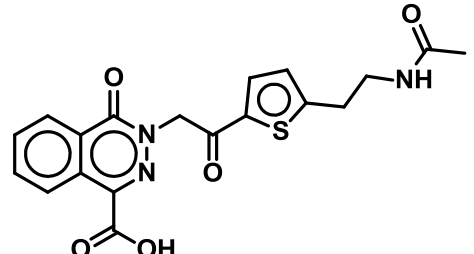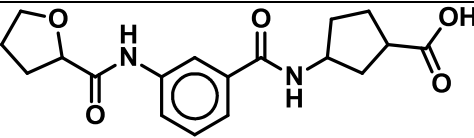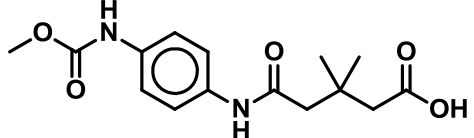
The maximum possible EF for these methodologies is 1000 (perfect accuracy). As there was no difference in docking accuracy between the two methodologies, we chose to employ methodology I, for future studies, due to its lower computational cost.
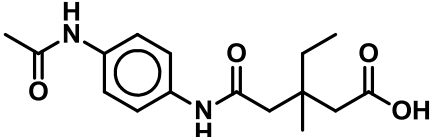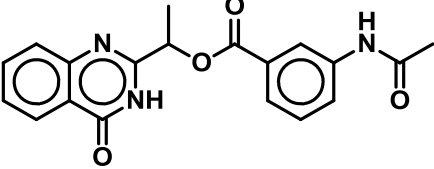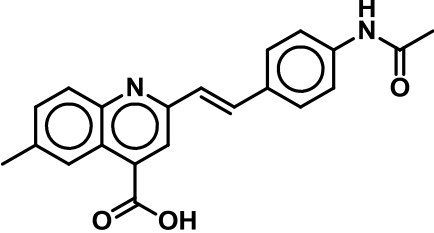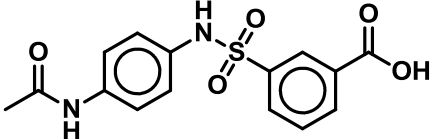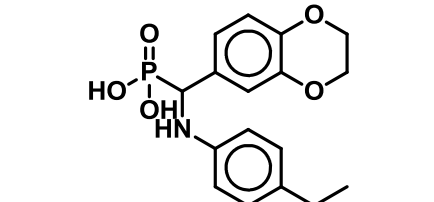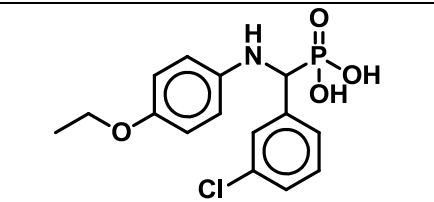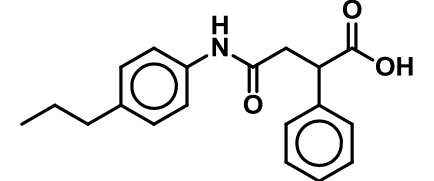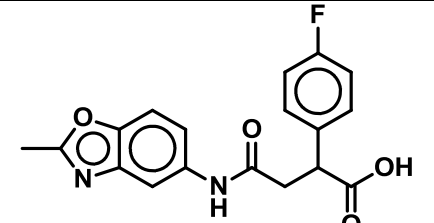
With methodology I, we docked and scored ~11 million compounds. Initially, the ~11 million compounds were docked with the HTVS method. The top 10% of compounds identified in HTVS mode (~1.1 million compounds) were advanced to SP docking. Finally, the top 1% of compounds docked in SP mode (~11,000 compounds) were selected for XP docking.
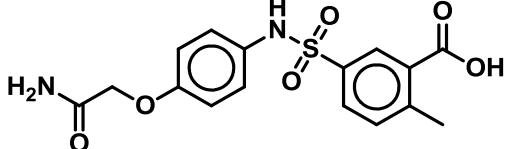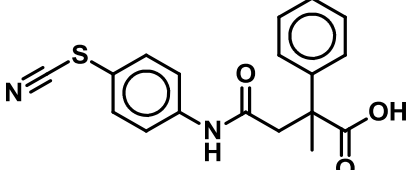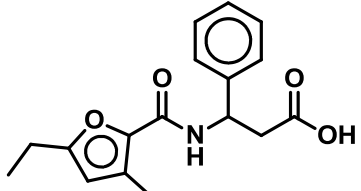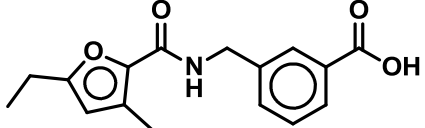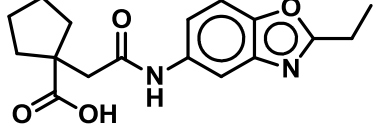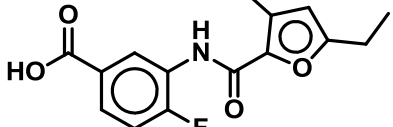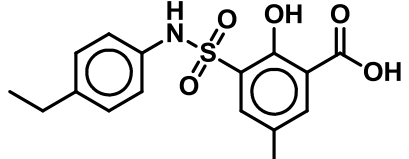
5.4 RESULTS AND DISCUSSION

We analyzed the resulting ~11,000 compounds for binding to the S1′* site and selected 190 compounds with high predicted inhibitory potency. Of these 190 compounds, only 65 were commercially available and of those, 23 were purchased. **Table 5.2** illustrates the 23 purchased compounds.

Table 5.2. Twenty-three purchased compounds from Enamine, Ltd predicted to inhibit LF, with Glide docking scores.

Structures	Compound ID	Docking score
	5.11	-10.5
	5.12	-10.4
	5.13	-10.4
	5.14	-10.3
	5.15	-10.2
	5.16	-10.1
	5.17	-10.1
	5.18	-10.0

Structure	Compound	Value
	5.19	-9.9
	5.20	-9.6
	5.21	-9.1
	5.22	-9.1
	5.23	-9.1
	5.24	-9.0
	5.25	-9.0
	5.26	-8.7

	5.27	-8.5
	5.28	-8.4
	5.29	-8.4
	5.30	-8.4
	5.31	-8.3
	5.32	-8.3
	5.33	-8.1

We studied the docking poses of the purchased compounds and analyzed important protein-ligand interactions within the active site of LF. Details are shown in **Table 5.3**.

Table 5.3. Frequency counts of residues engaged in protein-ligand interactions with the 23 purchased compounds.

	Asn679 Backbone H-Bond Acceptor	Gly674 Backbone H-Bond Donor	Glu687 Sidechain H-Bond Acceptor	Tyr728 Sidechain H-Bond Donor	Val675 Backbone H-Bond Donor	Lys656 Backbone H-Bond Donor	Gly657 Sidechain H-Bond Acceptor
23 Compounds	7	4	15	12	6	3	3

One example of a top-scoring compound interacting with the LF active site is shown in **Figures 5.1** and **5.2**. Compound **5.11** is predicted to engage in H-bond donating interactions with the backbone of Asn679, which is located at the entrance of the S1′* channel. The carboxylic acid moiety chelates zinc and engages in H-bond interactions with the sidechain of Tyr728. The peptidic NH is engaged in H-bond donating interactions with the sidechain of Glu687 (**Figures 5.1** and **5.2**).

Figure 5.1. 2D protein-ligand interaction map of **5.11**, which is predicted to be active against LF (PDB ID 4XM6). (Schrödinger Maestro Discovery Suite 9.4).
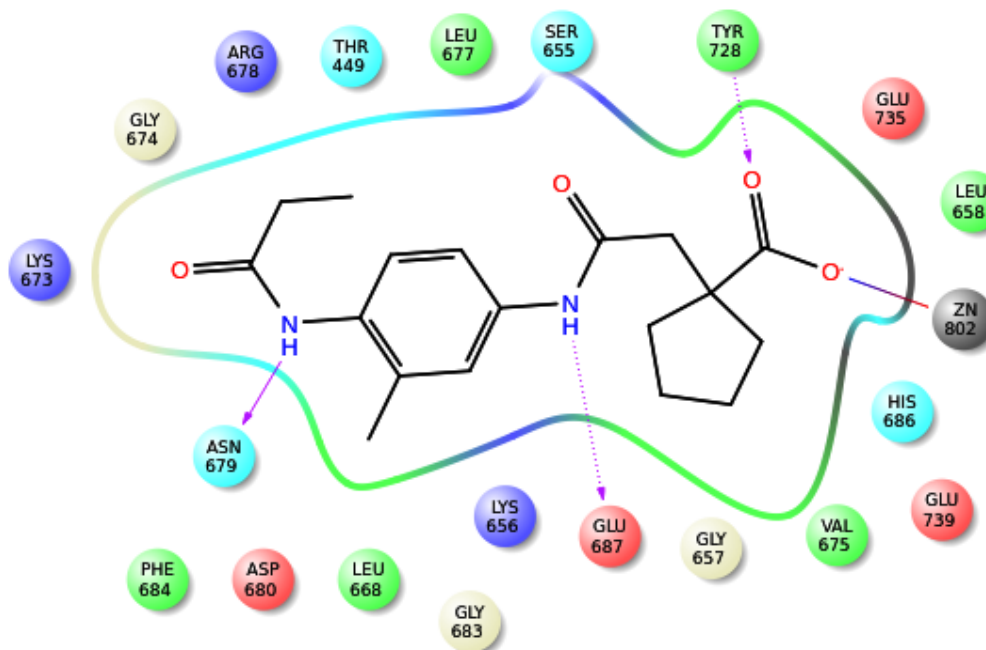
Figure 5.2. Three-dimensional (3D) protein-ligand interaction image of **5.11** (green), which is predicted to be active against LF (PDB ID 4XM6). (Schrödinger Maestro Discovery Suite 9.4).
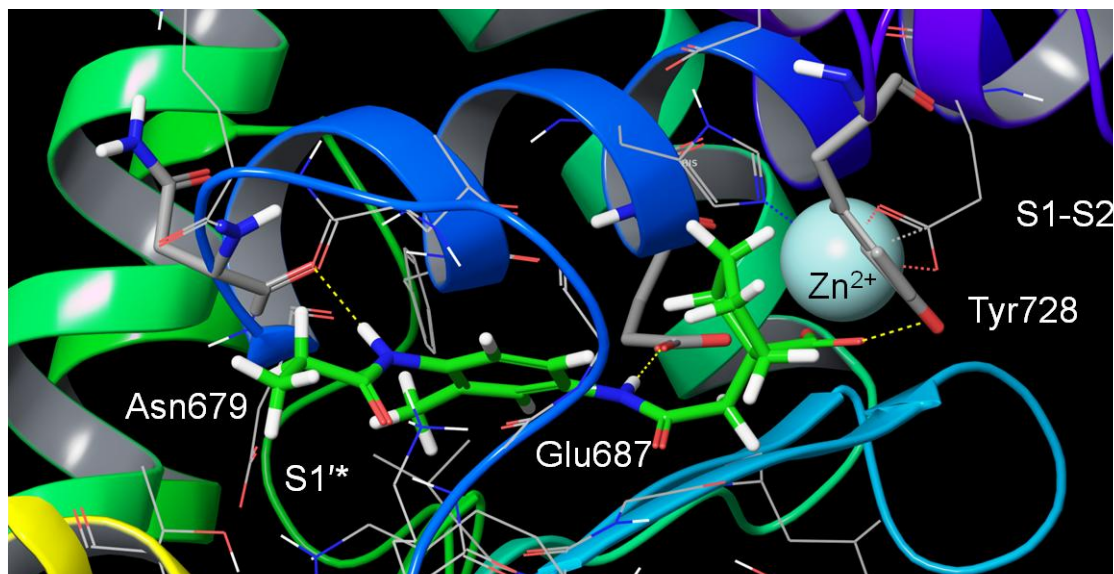
The 23 purchased compounds were experimentally evaluated for LF inhibition using our previously discussed *in vitro* LF FRET assay.[111] Only **5.33** showed LF inhibition, with an IC_{50} value of 126 μM and a ligand efficiency (LE, binding affinity measure as a function of size) of 0.24. LE is calculated using the following equation: LE = 1.4(-logIC_{50})/N, where N is the number of non-hydrogen atoms). The LE of **5.33** is superior to HTS hit **8.16**, which has a LE = 0.20 (See Chapter 8). **Figure 5.3** shows the predicted binding mode for **5.33** when docked to 4XM6. Compound **5.33** is predicted to chelate zinc through an *o*-hydroxybenzoic acid moiety, and the carboxylic acid is also engaged in H-bond interactions with the backbone of Tyr659. There are no predicted interactions with the sulfonamide moiety. The ethylphenyl moiety is occupying the S1′* subsite.

126

Figure 5.3. 3D protein-ligand interaction image of **5.33** (yellow) with LF (PDB ID 4XM6). (Schrödinger Maestro Discovery Suite 9.4).
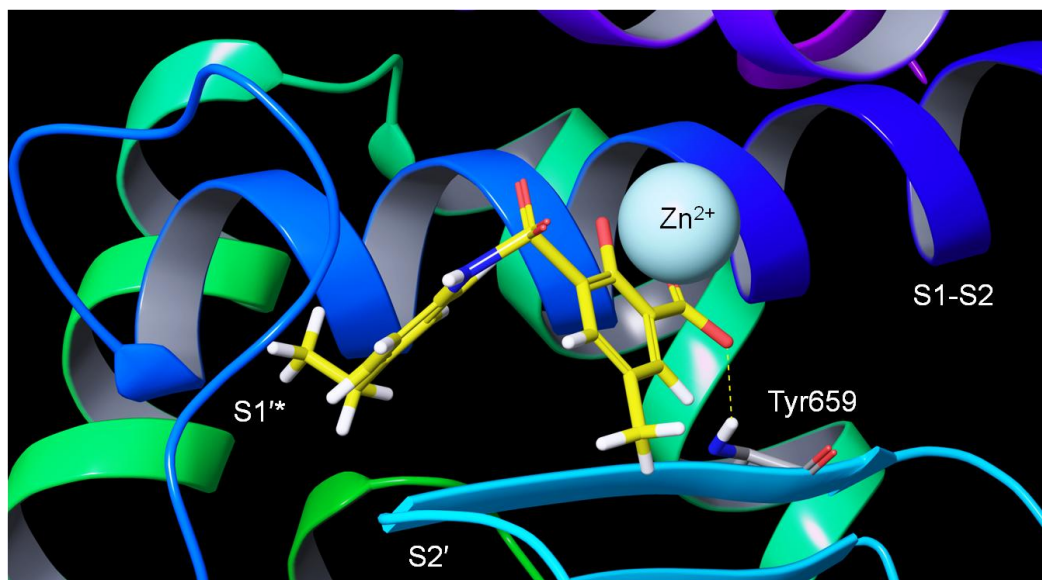
An alternative binding mode for **5.33** is shown in **Figure 5.4** when **5.33** was docked into 1YQY. In this binding mode, the o-hydroxybenzoic acid moiety coordinates zinc as in **Figure 5.3,** whereas the ethylphenyl moiety occupies the S1-S2 subsite rather than the S1′* subsite. The sulfonamide functionality engages in H-bond interactions with the backbone of Tyr659.

Interestingly, various *o*-hydroxybenzoic acid fragments were tested against LF by Cohen *et al*. and reported to be inactive.[164] In this case, **5.33** is hypothesized to be active because of the electron withdrawing sulfonamide functionality on the phenyl ring, which can decrease the pK_a of the phenol, making it more likely to bind zinc as O^-. Phenoxides coordinate zinc more strongly than phenols. X-ray studies are underway to determine the experimental binding mode of **5.33**. Once the binding mode is elucidated, structure-based drug design can be employed to improve the activity of **5.33**.

Figure 5.4. 3D protein-ligand interaction image of **5.33** (yellow) with LF (PDB ID 1YQY). (Schrödinger Maestro Discovery Suite 9.4).
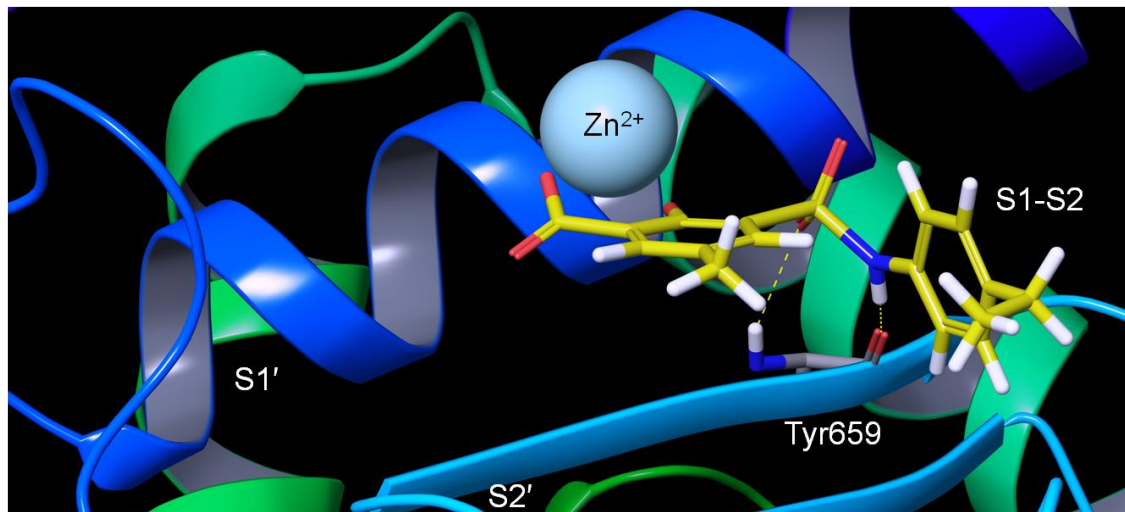
5.5 CONCLUSION

The goal of this study was to identify LF inhibitory compounds that bind to the newly identified S1′* subsite of LF. For that purpose, we developed and tested an accurate docking methodology that can be used to virtually screen large databases in a timely manner. Using this approach, ~11 million compounds from the ZINC database were docked and scored. Methodology I prioritized ~11,000 compounds out of ~11 million to be docked and scored in XP mode. We have analyzed these compounds for binding to the S1′* site and purchased 23 compounds for further interrogation. We identified **5.33** as a novel LF inhibitor with an IC_{50} = 126 µM. The hit rate of this virtual screen was 1/23 (4.3%) versus 1/250,000 (0.0004%) in our large-scale experimental HTS campaign (See Chapter 8). X-ray studies are underway to determine the binding mode of **5.33**.

Chapter 6

LARGE-SCALE VIRTUAL DATABASE SCREENING TOWARDS THE IDENTIFICATION OF NOVEL LF INHIBITOR SCAFFOLDS

6.1 INTRODUCTION

In the previous chapter, LF crystal structure 4XM6 was used to screen ~11,000,000 compounds from the ZINC[159] database and a novel LF inhibitory scaffold was discovered. Despite the success of this presented strategy, its general use for other enzymatic targets may be limited due to its high computational cost. Here, I present a new protocol that was designed to lower the computational cost of virtual screening for metalloprotein drug targets by selecting only drug-like compounds with zinc binding groups (ZBGs) for docking and scoring studies. In this study, we used Glide 5.9,[114–117] available through the Schrödinger Maestro Discovery Suite 9.4 (Schrödinger, Inc.),[118] to screen large compound libraries from eMolecules to identify novel potential inhibitors of LF. eMolecules is a search engine for chemicals that contains ~7,000,000 compounds from commercial suppliers like Enamine, ChemDiv, and ChemBridge to name a few.[165] The benefit of using the eMolecules database versus the ZINC database is that all compounds included in eMolecules are commercially available. One limitation of our previous study was that many compounds we identified using the ZINC database were not commercially available (out of 190 selected compounds, only 65 were commercially available).

6.2 COMPUTATIONAL METHODS

6.2.1 Selection of Appropriate LF Crystal Structure/s for Docking

Preparation of Crystal Structures For Docking and Scoring: Prior to screening, we selected specific X-ray structures, out of the 16 extant in the PDB,[166] that would yield the best screening results. By visual analysis, we removed 1JKY[50] and 1PWV[91] as they did not contain the active site zinc, 1J7N[50] as it did not contain a bound ligand, and 1PWW[91]

as it contained a mutated active site residue (E687C). The twelve remaining crystal structures were aligned using the 'ATLF' overlay method developed by Finzel *et al.*, which was described in Chapter 2.[167] The overlaid structures were prepared using the protein preparation wizard[122,123] in Maestro, and missing sidechains were added with Prime.[124] These final structures were used in receptor grid generation for the subsequent docking studies.

Initial Docking and Scoring: Before screening large compound databases, we performed a validation study using **DB1** (introduced in chapter 5) to select LF structures that best retrieve known active compounds. Prepared **DB1** was docked into the twelve aforementioned LF crystal structures using Glide 5.9[114–117] with standard precision (SP) in the Schrödinger Maestro Discovery Suite 9.4 (Schrödinger, Inc.).[118] Enrichment factors[163] (EF) for each crystal structure were calculated to compare the screening performance by counting the number of known actives among the top ten scoring compounds (*i.e.* 0.1% of total **DB1**). We computed EF values as follows[163]:

$$\mathrm{EF} = \frac{\mathrm{Hits}_{\mathrm{sampled}}/N_{\mathrm{sampled}}}{\mathrm{Hits}_{\mathrm{total}}/N_{\mathrm{total}}}$$

For example, 1YQY retrieves nine known actives among the top ten scoring compounds, hence EF for 1YQY is:

$$\mathrm{EF} = \frac{9/10}{10/10010} = 901$$

The maximum possible EF for these structures is 1000 (perfect accuracy). **Table 6.1** illustrates that 1YQY (bolded) best recovered the known actives from the decoys with an EF = 901, recovering 90% of actives in 0.1% of **DB1**. This test also shows that SP mode was sufficient to recover most of the actives when used in screening against

131

1YQY. Thus, we predicted that a strategy to virtually screen against 1YQY would translate to the identification of more experimentally active compounds than would be identified if virtual screening was performed against any of the other LF crystal structures.

Table 6.1. Enrichment factors for docking and scoring of **DB1** against twelve LF structures.

Crystal Structures	EF (0.1%)	% of Actives Recovered
1PWQ[91]	0	0
4DV8[86]	300	30
1YQY[110]	**901**	**90**
4PKR[125]	501	50
4PKS[125]	501	50
4PKT[125]	601	60
4PKU[125]	200	20
4PKV[125]	100	10
1PWU[91]	200	20
4WF6	400	40
1PWP[100]	0	0
1ZXV[81]	0	0

6.2.2 Development of Docking and Scoring Protocol

Described below is the protocol we developed to prepare, filter, and dock drug-like compounds into the crystal structure of 1YQY.

1. Download ~7 million commercial compounds from eMolecules database.[165]

2. Remove reactive molecules using Schrödinger's Canvas.[15–17] Canvas removes compounds with specific reactive functional groups that are known to be problematic and interfere in biological assays. The list of such groups was first reported by Baell *et al.*[112]

3. Use Ligprep and Epik to prepare compounds as previously described. Compounds were prepared using LigPrep (Schrödinger),[120] which generates 3D-minimized structures at pH 7.0. Because LF is a zinc metalloprotein, Epik (Schrödinger)[168] was used to generate "metal binding states" of the ligands, which are ligand ionization states that are more likely to bind active site metals. For example, hydroxamic acids are protonated at pH 7.0, as in –CONHOH. However, we know that hydroxamic acids bind to zinc as –CONHO⁻. Epik recognizes hydroxamic acids as metal binding groups and alternatively prepares the ligand as –CONHO⁻. Thus, Epik is an essential step of ligand preparation, enabling us to appropriately screen compounds with zinc-binding groups.

4. Filter compounds by drug-likeness:

a. Remove compounds with MW ≥ 500, ClogP ≥ 5, H-bond acceptors ≥ 10, and H-bond donors ≥ 5 (Lipinski's Rules).[160]

b. Use QikProp[169] to calculate polar surface area (PSA or FISA in QikProp) for each ligand, and remove compounds with FISA > 140 Å^2 and rotatable bonds > 10.[170]

5. Filter compounds to retain only zinc-binders.

6. Dock compounds that pass all the aforementioned filters using SP mode.

7. Select and re-dock the top 1% using XP mode.

Rationale for Enriching Zinc Binders: Because LF is a zinc-containing enzyme, we incorporated filtering step **5** to retain only compounds capable of binding zinc. The

133
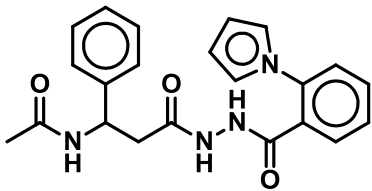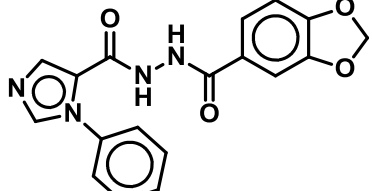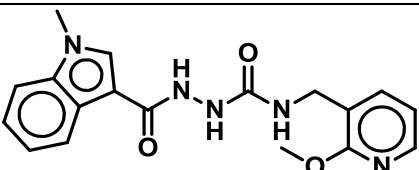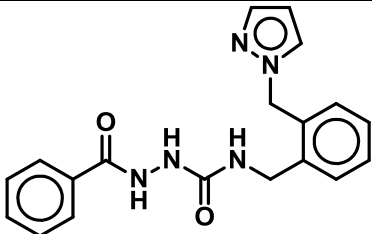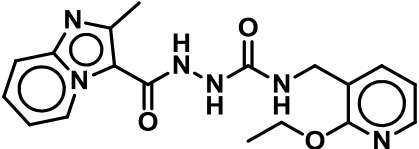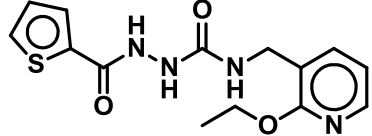
retained compounds can be characterized as having negative charges (-1 or -2), metal binding states that were generated by Epik, or carboxylic acids (1-2 groups). We hypothesized that docking only compounds that passed this filter would be as accurate as docking the entire compound library because zinc-binding ligands would be predicted as the highest scoring binders. Using this strategy, we were able to drastically improve our time-efficiency, which is crucial in settings with limited or shared computational resources. We tested this hypothesis by filtering a ~200,000 compound library from eMolecules using the proposed method. After selecting for zinc-binders, ~30,000 compounds remained. We then docked and scored both the filtered and unfiltered versions of this library with SP mode. The top 292 compounds identified from the two libraries were identical. Thus, by instituting one additional filter we were able to enrich the library for LF binders and reduce our computing time ~10-fold. This method offers a "smart and fast strategy" for virtually screening large libraries against zinc metalloproteins by reducing the library size to include only those with high predicted zinc-binding affinities. We believe that this protocol can also be used for other metal-containing active sites.
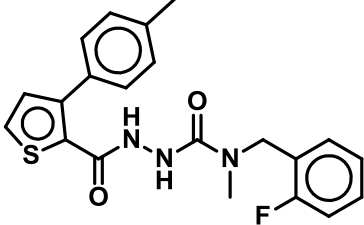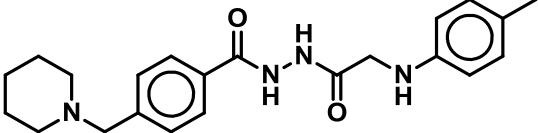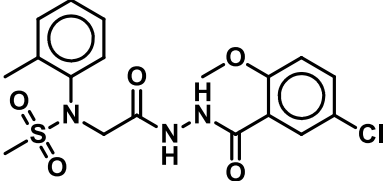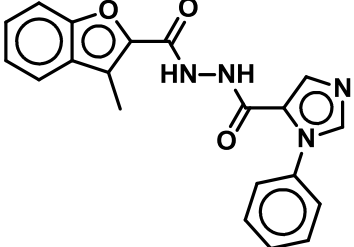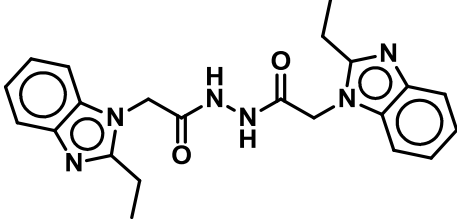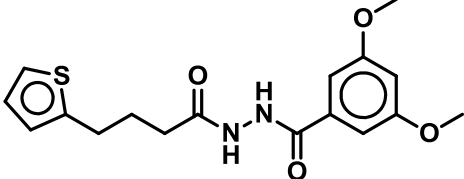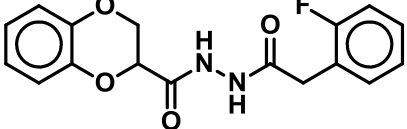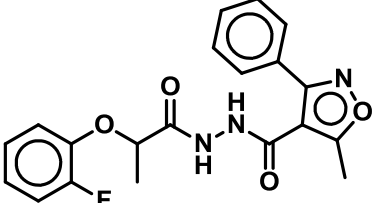
6.3 RESULTS AND DISCUSSION
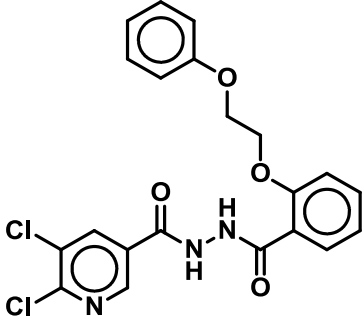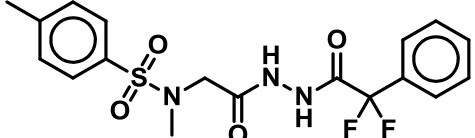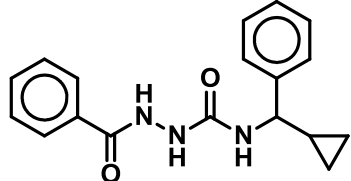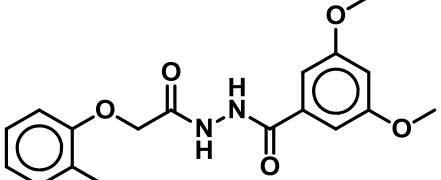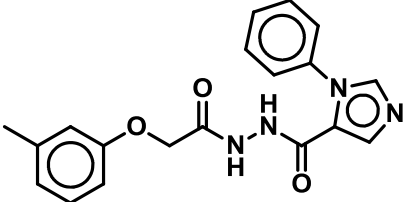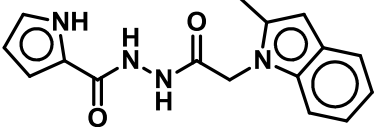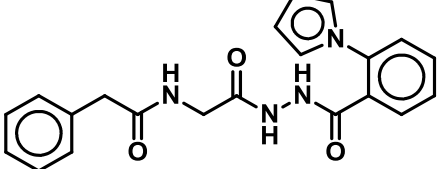
From the ~7,000,000 compounds downloaded from eMolecules, ~1,000,000 were removed by Canvas. Moreover, using drug-like filters (steps 4a/b), we removed ~2,000,000 compounds. Filtering to remove compounds that do not bind zinc reduced the library from ~4,000,000 compounds to ~540,000 compounds, all of which were docked and scored using SP mode. Only the top 4000 compounds (~1%) were prioritized and then re-docked using Glide's XP mode. A large representation of top-scoring compounds from the XP and SP docking modes were hydrazides, which are

134

bioisosteres of hydroxamic acids. Of the top-scoring hydrazide-containing compounds, we selected 21 to be purchased and tested for LF inhibition. Structures of the top-scoring, purchased compounds are presented in **Table 6.2**. The extreme closeness of docking scores should be noted, and range from -11.0 to -12.2 kcal/mol.

Table 6.2. Twenty-one purchased compounds from Enamine, Ltd predicted to inhibit LF with Glide docking scores

Structures	Compound ID	Docking Score (kcal/mol)
	6.1	-12.2
	6.2	-12.2
	6.3	-11.9
	6.4	-11.7
	6.5	-11.7
	6.6	-11.6

	6.7	-11.6
	6.8	-11.5
	6.9	-11.4
	6.10	-11.3
	6.11	-11.2
	6.12	-11.2
	6.13	-11.2
	6.14	-11.2

	6.15	-11.1
	6.16	-11.1
	6.17	-11.1
	6.18	-11.1
	6.19	-11.0
	6.20	-11.0
	6.21	-11.0

We have studied the docking poses of the purchased hydrazides and analyzed important protein-ligand interactions within the active site of LF. The details of the

frequency and types of these interactions for the 21 purchased compounds are shown in

Table 6.3.

Table 6.3. Frequency counts of residues engaged in protein-ligand interactions with the 21 purchased hydrazides.

Residues	Tyr659 H-Bond Donor	Gly657 H-Bond Donor	His690 π-π	His686 π-π	Val675 H-Bond Donor	Glu687 H-Bond Acceptor	Lys656 H-Bond Donor
21 Purchased Compounds	19	15	8	12	7	9	7

An example of a top-scoring hydrazide interacting with the LF active site is shown in **Figure 6.1**. The hydrazide moiety of these compounds chelates zinc similarly to a hydroxamic acid (as in **MK-702/LF1-B,** which was co-crystallized with LF (PDB ID 1YQY)) in a bidentate mode. However, compared to hydroxamic acids, hydrazide functional groups have an advantage. The hydrazide can be modified to engage in protein-ligand interactions with residues in the S1-S2 subsite, in addition to, the S1′ subsite. Essentially, the zinc binding hydrazide can act as a metal chelating bridge between two halves of the small molecule. For example, the carbonyl moiety of **6.1** is engaged in H-bond interactions with the backbone of Tyr659 in the S1-S2 subsite, while the phenyl group is occupying the hydrophobic S1′ subsite and is engaged in π-π interactions with His686 and π-cation interactions with active site zinc (**Figures 6.1, 6.2**). This mode of interaction is representative of this class of compounds (**Table 6.4**).

Figure 6.1. 2D protein-ligand interaction map of **6.1**, which is predicted to be active against LF (PDB ID 1YQY). (Schrödinger Maestro Discovery Suite 9.4).
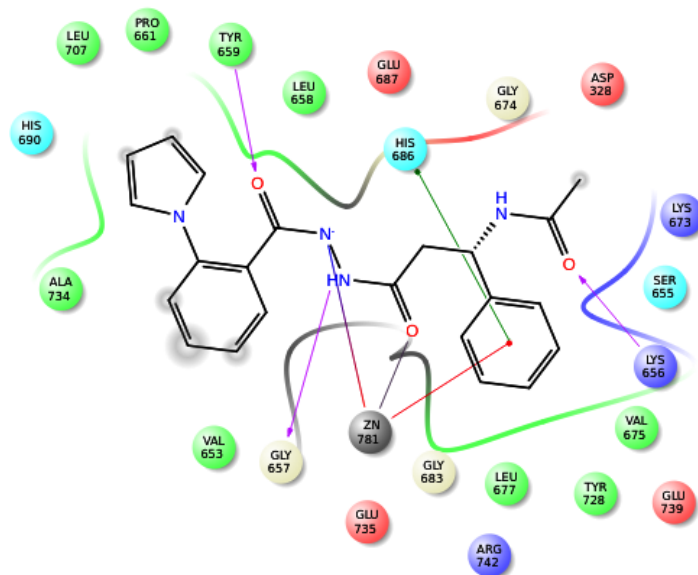

Some of these compounds are also engaged in key hydrogen bond interactions with Gly657, Lys656, and Val675, which are crucial for the binding affinity of the compounds (**Table 6.3**).

Biochemical evaluation of the 21 purchased compounds was performed using a previously published *in vitro* FRET assay.[111] Unfortunately, none of the purchased compounds showed LF inhibition when tested up to 1 mM concentrations. This result can be rationalized by the difficulty in accurately predicting the absolute protein-ligand binding free energies of proteins that contain active site zinc metals through docking and scoring programs. The complex coordination environment, polarization, and charge transfer effects associated with zinc make ligand binding free energy calculations very challenging.[171,172]

Figure 6.2. 3D protein-ligand interaction image of **6.1** (green) which is predicted to be active against LF (PDB ID 1YQY). (Schrödinger Maestro Discovery Suite 9.4).
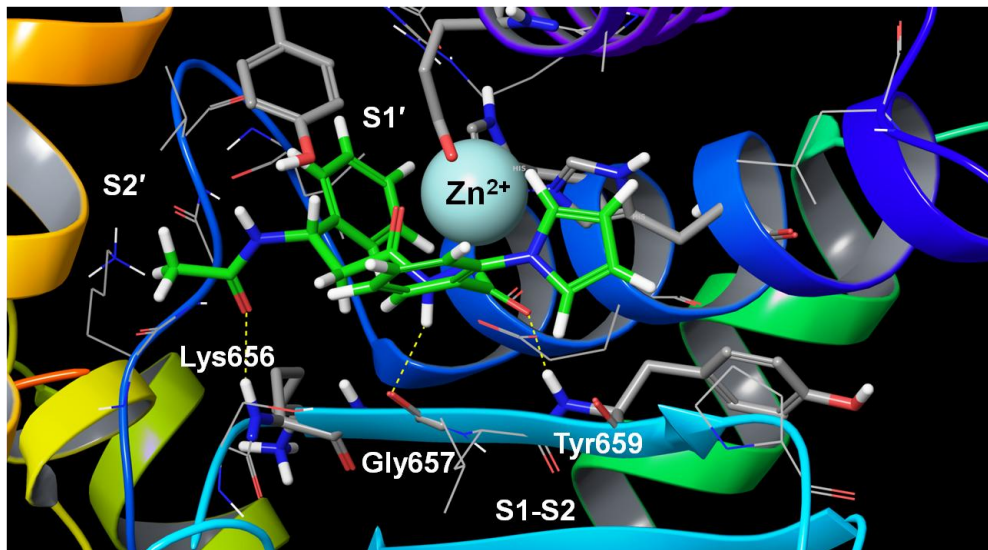
A number of methods have been reported to improve binding affinity calculations of ligands to zinc-containing metalloproteins. For example, it has been reported by Khandelwal *et al*. that QM/MM methods coupled with MD simulations can produce accurate estimations of binding affinities for 28 MMP-9 ligands.[173] These methods, however, are not amenable for the virtual screening of large compound libraries. Glide, which uses only MM methods, is not nearly as accurate in calculating metalloprotein-ligand binding free energies, but is amenable for the efficient screening of millions of compounds. The developers of Glide note that the errors in binding free energy calculations can be as high as 4 kcal/mol, and are likely much higher for zinc metalloproteins due to the aforementioned difficulties.[174] They attribute this error to several factors, including the use of rigid protein structures, imperfect scoring functions, gridded potentials, and the complex coordination environment, polarization, and charge transfer effects due to zinc. Moreover, Glide is primarily intended for database enrichment, rather than accurate calculation of absolute protein-ligand binding free

140

energies.[174] Knowing these limitations, it is not surprising, although disappointing, that none of the purchased hydrazides showed inhibitory activity.

6.4 CONCLUSION

Docking and scoring can be a very useful tool for screening large compound databases in an effort to discover novel inhibitors of protein targets. We have used this tool for efficient virtual screening of a large compound database against LF. Using multiple tests, we identified 1YQY as the crystal structure that best retrieves active compounds. Using this discovery, we employed 1YQY for the efficient virtual screening of a large compound library from eMolecules. Our research objectives were 1) to develop a virtual screening protocol that decreases computational cost without over-looking any top-scoring hits and 2) to identify potential LF inhibitors by applying the developed protocol to eMolecules. Using the developed protocol, we have shown that we can enrich compound libraries for small molecules that have zinc binding capabilities, without losing any potential top-scoring hits. As a result, we were able to reduce the size of the compound library from ~7,000,000 to ~540,000. This reduced our computational requirements ~10-fold, allowing us to finish this project within a reasonable time frame. We achieved our primary objective of developing an efficient protocol that decreases computational cost while retaining accuracy. By analyzing the top-scoring compounds, we prioritized hydrazide-containing ligands for further interrogation. In docking studies, these ligands displayed bidentate zinc chelation similar to that of hydroxamic acids, and engaged in key protein-ligand interactions with active site residues such as Tyr659, Gly657, Lys656, Val675, and Glu687. We selected 21 hydrazides and tested them against LF. The observation that none of the purchased compounds inhibited LF likely results from the challenges associated with accurately calculating metalloprotein-ligand

binding free energies via docking and scoring methods. Hence, this result is due to the docking and scoring method not the protocol that we developed. Therefore, we believe that this virtual screening protocol offers an efficient strategy for screening large libraries against zinc metalloproteins by reducing the library size to include only those with high predicted zinc-binding affinities. This method can be applied to drug discovery efforts for any zinc-containing enzyme, and potentially, other metalloproteins.

Chapter 7

**ELECTROSTATICALLY EMBEDDED MANY-BODY EXPANSION OF THE ENERGY
(EE-MB) AND THE CORRELATION ENERGY (EE-MB-CE) FOR ZN AND CD MODEL
SYSTEMS INCLUDING A MODEL OF THE CATALYTIC SITE OF THE ZINC-
BEARING ANTHRAX TOXIN LETHAL FACTOR**

Adapted with permission from:

Kurbanov, E. K.; Leverentz, H. R.; Truhlar, D. G.; Amin, E. A. Electrostatically
Embedded Many-Body Expansion for Neutral and Charged Metalloenzyme Model
Systems. *J. Chem. Theory Comput.* **2012**, *8* (1), 1–5

Kurbanov, E. K.; Leverentz, H. R.; Truhlar, D. G.; Amin, E. A. Analysis of the
Errors in the Electrostatically Embedded Many-Body Expansion of the Energy and the
Correlation Energy for Zn and Cd Coordination Complexes with Five and Six Ligands
and Use of the Analysis to Develop a Generally Successful Fragmentation Strategy. *J.
Chem. Theory Comput.* **2013**, *9* (6), 2617–2628.

7.1 ACKNOWLEDGEMENTS

7.2 INTRODUCTION

In order to address some of the challenges regarding the accurate modeling of zinc metalloenzymes discussed in chapter 6, we employed quantum mechanical EE-MB and EE-MB-CE methods to obtain accurate bond dissociation energies for Zn and Cd model systems. Our methods can be used in future efforts to parameterize inorganic reactive force fields for use in studying zinc metalloenzymes, or could be used directly in QM-based simulations to obtain more accurate results that would otherwise be impossible for large systems (for example, systems with large ligands) where full quantum mechanical calculations on the whole complex with a reliable method are computationally unaffordable.

Zinc is an essential transition metal required for the catalytic and structural activity of many enzymes,[175] and it participates in a number of key biological processes in living systems, including immune function,[176,177] protein synthesis,[176,178] wound healing,[179,180] DNA synthesis,[176,181] and cell division.[176,181] Zinc metalloenzymes carry out essential functions in a wide variety of biochemical pathways and have attracted much attention as drug design targets; examples include the anthrax toxin lethal factor,[111] insulin, phosphotriesterase, the matrix metalloproteinases, cytidine deaminase, histone

144

deacetylases, zinc-finger proteins, and human carbonic anhydrase. In these enzymes zinc may play structural and/or catalytic roles, with catalysis taking place in the first coordination shell.[182] *In silico* techniques have generally proven valuable for rational drug design and enzyme modeling; however, reliable representation of zinc and other transition metal centers in macromolecules is nontrivial due to the complexity of the coordination environment and charge distribution at the catalytic center. Hybrid density functional theory[183] and post-Hartree–Fock correlated wave function methods, such as second-order Møller-Plesset perturbation theory, MP2,[184] coupled cluster theory with single and double excitations, CCSD,[185,186] and CCSD with quasiperturbative connected triple excitations, CCSD(T)[187] are all able to calculate accurate energies for selected small and moderately sized systems, but they are often too computationally costly— either because the system (or model system) to be studied is too large, or because a large number of calculations must be performed to achieve adequate sampling in a simulation. We note that in general the "expense" of a calculation depends on the program, the algorithm, the degree of parallelization, and the computer, and one should take account of such factors as computer time, memory requirements, communication among processors, input–output, human time, and other factors, but for discussion purposes we simply consider the number of arithmetic operations, and for convenience we call that the "expense". It is well known[188] that if one keeps the average number of basis functions per atom fixed, the expenses of, for example, hybrid density functional theory, MP2, CCSD, and CCSD(T)—with conventional basis sets—nominally scale, in the large—N limit, as N^4, N^5, N^6, and N^7 respectively, where N is the number of atoms. Therefore, enabling accurate calculations on large systems including the active sites of Zn-containing enzymes with post-Hartree–Fock methods remains challenging due to the rapid scaling of the computational cost. To make the problem more tractable, localized

molecular orbitals[189–193] and fragmentation[194–212] methods have been designed. In our previous work, we developed and implemented the electrostatically embedded many-body method[202,213–219] (EE-MB) and the electrostatically embedded many-body expansion of the correlation energy[214,220] (EE-MB-CE), which are fragment-based approaches for calculating the energies of large systems. These methods, with pairwise additive (PA) or three-body (3B) truncation of the many-body expansion, nominally scale as N^3 or lower for EE-MB and as N^4 or lower for EE-MB-CE (where N is now the number of monomers, see below), even if the individual dimer or trimer calculations scale less favorably because the cost of the individual oligomer calculations does not increase with N.

As described in our previous work,[202,213–219] the EE-MB method tackles the challenge of system size by partitioning larger complexes into a series of fragments called monomers, and calculating the energies of monomers, dimers, and optionally trimers or higher oligomers by embedding them in a field of point charges representing the remaining N-1, N-2, N-3, ... monomers, and running calculations in parallel. (A monomer can be defined as a single molecule, a portion of a molecule, or a collection of molecules. For example, in the systems considered here, a monomer could be an ammonia molecule or Zn^{2+} with two ammonia ligands). There are two variations of the EE-MB method: the electrostatically embedded pairwise additive method (EE-PA), which is based on the energies of monomers and dimers, and the electrostatically embedded three-body method (EE-3B), which is based on the energies of monomers, dimers, and trimers. The EE-3B method is able to predict bond energies obtained by conventional full-system calculations done at the same level of theory to within 1.0 kcal/mol for cationic, neutral, and negatively charged Zn^{2+} complexes.[218,219]

146

In EE-MB-CE, one applies the many-body expansion only to the correlation energy, that is, to the post-Hartree–Fock part of the energy calculation. Here we apply the EE-MB-CE method to predict MP2 correlation energies for a variety of pentacoordinate and hexacoordinate Zn^{2+} and Cd^{2+} systems, and we compare its performance to the EE-MB method. Most importantly, we present a new, simple, and unambiguous fragmentation strategy that maximizes the accuracy and efficacy of both the EE-MB and the EE-MB-CE calculations for Zn^{2+} and Cd^{2+} complexes studied in this paper. The number of computational operations in the EE-PA method scales as N^2, the number for the EE-3B method scales as N^3, and the numbers for the EE-PA-CE and EE-3B-CE methods scale as N^4, where N is the number of monomers.

7.3 THEORY

7.3.1 EE-MB

For any level of theory (e.g., MP2 or CCSD(T) with a given basis) we can either perform full (i.e., conventional) calculations of the potential energy, E, or many-body expansions. In the EE-MB method[202,213–218,220] the fragments into which a system is partitioned are called monomers. In the present study, we examine two variants of this method: the electrostatically embedded pairwise additive (EE-PA) approximation and the electrostatically embedded three-body (EE-3B) approximation. The EE-MB method approximates the energy of systems composed of monomers i, j, k, ... as

$$E = E^{(1)} + \Delta E^{(2)} + \Delta E^{(3)} + ... + \Delta E^{(N)} \tag{1}$$

where

147

$$E^{(1)} = \sum_i E_i \tag{2}$$

$$\Delta E^{(2)} = \sum_{i<j} (E_{ij} - E_i - E_j) \tag{3}$$

$$\Delta E^{(3)} = \sum_{i<j<k} [(E_{ijk} - E_i - E_j - E_k) - (E_{ij} - E_i - E_j) - (E_{ik} - E_i - E_k) - (E_{jk} - E_j - E_k)] \tag{4}$$

where E_i, E_{ij}, E_{ijk} are the energies of a monomer, dimer, and trimer, respectively, embedded in fields of point charges representing the other monomers; and the individual energies are obtained using any type of electronic structure theory. Then the EE-PA approximation is

$$E^{(\mathrm{PA})} = E^{(1)} + \Delta E^{(2)} \tag{5}$$

and the EE-3B approximation is

$$E^{(3\mathrm{B})} = E^{(\mathrm{PA})} + \Delta E^{(3)} \tag{6}$$

7.3.2 EE-MB-CE

Many-body expansion methods examined in the present study are the EE-PA approximation, the EE-3B approximation, the electrostatically embedded pairwise additive approximation of the correlation energy (EE-PA-CE), and the electrostatically embedded three-body approximation of the correlation energy (EE-3B-CE) methods. The theory of the EE-MB method was previously discussed. The electronic energy at the MP2 level of theory can be written as

$$E_{\text{MP2}} = E_{\text{HF}}^{\text{total}} + \Delta E_{\text{corr,MP2}} \tag{7}$$

where $E_{\text{HF}}^{\text{total}}$ is the Hartree–Fock energy of the system, and $\Delta E_{\text{corr,MP2}}$ is the MP2 correlation energy which can be rewritten as the many-body expansion of the correlation energy:

$$\Delta E_{\text{corr,MP2}} = \Delta E_{\text{corr,MP2}}^{(1)} + \Delta E_{\text{corr,MP2}}^{(2)} + \Delta E_{\text{corr,MP2}}^{(3)} + ... + \Delta E_{\text{corr,MP2}}^{(N)} \tag{8}$$

$$\Delta E_{\text{corr,MP2}}^{(1)} = \sum_i \Delta E_i^{\text{corr}} \tag{9}$$

$$\Delta E_i^{\text{corr}} = E_i^{\text{MP2}} - E_i^{\text{HF}} \tag{10}$$

where E_i^{HF} and E_i^{MP2} are respectively the Hartree–Fock and MP2 energies of a monomer i, and ΔE_i^{corr} is the MP2 correlation energy of a monomer i,

$$\Delta E_{\text{corr,MP2}}^{(2)} = \sum_{i<j} (\Delta E_{ij}^{\text{corr}} - \Delta E_i^{\text{corr}} - \Delta E_j^{\text{corr}}) \tag{11}$$

$$\Delta E_{ij}^{\text{corr}} = E_{ij}^{\text{MP2}} - E_{ij}^{\text{HF}} \tag{12}$$

where E_{ij}^{HF} and E_{ij}^{MP2} are the respectively the Hartree–Fock and MP2 energies of a dimer ij, and $\Delta E_{ij}^{\text{corr}}$ is the MP2 correlation energy for a dimer ij; therefore EE-PA-CE energy can be defined as

$$E^{\text{PA-CE}} = E_{\text{HF}}^{\text{total}} + \Delta E_{\text{corr,MP2}}^{(1)} + \Delta E_{\text{corr,MP2}}^{(2)} \tag{13}$$

The EE-3B-CE energy can be written as

$$E^{\text{3B-CE}} = E^{\text{PA-CE}} + \Delta E_{\text{corr,MP2}}^{(3)} \tag{14}$$

$$\Delta E_{\text{corr,MP2}}^{(3)} = \sum_{i<j<k}[(\Delta E_{ijk}^{\text{corr}} - \Delta E_i^{\text{corr}} - \Delta E_j^{\text{corr}} - \Delta E_k^{\text{corr}}) - (\Delta E_{ij}^{\text{corr}} - \Delta E_i^{\text{corr}} - \Delta E_j^{\text{corr}}) -$$

$$(\Delta E_{ik}^{\text{corr}} - \Delta E_i^{\text{corr}} - \Delta E_k^{\text{corr}}) - (\Delta E_{jk}^{\text{corr}} - \Delta E_j^{\text{corr}} - \Delta E_k^{\text{corr}})] \tag{15}$$

$$\Delta E_{ijk}^{\text{corr}} = E_{ijk}^{\text{MP2}} - E_{ijk}^{\text{HF}} \tag{16}$$

where E_{ijk}^{HF} and E_{ijk}^{MP2} are respectively the Hartree–Fock and MP2 energies of a trimer ijk, and $\Delta E_{ijk}^{\text{corr}}$ is the MP2 correlation energy for a trimer ijk. The individual energies are obtained using the MP2 level of theory. One can obtain the EE-MB-CE approximation of the CCSD(T) energy by replacing "MP2" with "CCSD(T)" in equations 7–16.

In the EE-PA and EE-3B calculations, one only performs MP2 or CCSD(T) calculations on the monomers, dimers, and (for 3B) the trimers. In the EE-PA-CE and EE-3B-CE approximations, one also carries out a Hartree–Fock calculation on the entire system. The electronic energy at the MP2 or CCSD(T) level of theory can be written as the Hartree–Fock energy plus the correlation energy. In the EE-PA-CE and EE-3B-CE approximations, one uses expansions only on the correlation energy, and adds the correlation energy to the directly calculated Hartree–Fock calculation of the entire system (eq. 13 and 14). Since Hartree–Fock theory formally scales as N^4, where N is the number of atoms, it is less computationally demanding to carry out a Hartree–Fock calculation on the entire system than to carry out an MP2 or other correlated wave function calculation on the entire system, and the goal here is to obtain accuracy equivalent to a full MP2 or CCSD(T) calculation without the cost of the latter.

In past work,[214,220] the EE-MB-CE method has been applied to large water clusters, ranging in size from 5 to 20 water molecules and water hexamers. Here, we examine the application of the EE-MB-CE method to metal-ligand bonding in various pentacoordinate and hexacoordinate Zn and Cd complexes and present a fragmentation

scheme that can be applied to all Zn and Cd systems studied in the paper. These metal-ligand systems are more challenging than noncovalently bonded clusters because the electrostatic and induction effects are much larger, and there is some covalent character in the metal-ligand coordination bonds.

We also present an application where MP2 is replaced by CCSD(T).

7.4 EE-MB

7.4.1 METHODS

We have already shown that the EE-MB method can be used to calculate usefully accurate bond dissociation energies at low computational cost for positively charged Zn^{2+} systems; in particular the EE-3B method predicts bond energies obtained by conventional full-system calculations done at the same level of theory to within 1.0 kcal/mol for those cationic Zn^{2+} complexes.[218] In the present work, we recommend a set of specific fragmentation strategies to enhance the accuracy of EE-MB for coordination chemistry, and we assess the suitability of the EE-3B method for the more challenging neutral and negatively charged penta- and hexacoordinate Zn systems of biological importance; we also present EE-PA results for comparison.

Charges are calculated for each fragment at the geometry of that monomer in the overall system. For example, if we are calculating the energy of ZnABCDEF, where A, B, C, D, E, and F are ligands, and if one of the fragments is ZnBC, we calculate the partial atomic charges of ZnBC by removing A, D, E, and F from the system. Here we calculate charges using Merz-Kollman (MK) electrostatic fitting,[221] as in previous work on Zn compounds.[218]

151

All calculations were done with the M05-2X density functional[222] and the B2 basis set,[182] which is a polarized valence-triple-zeta basis set optimized and validated for use with Zn-containing complexes including biozinc coordination systems. Our earlier published work on a variety of Zn-ligand systems of importance in biology, nanotechnology and drug design[182,223] showed that incorporating relativistic effects on core electrons significantly increased the accuracy of geometric and energetic calculations for Zn coordination complexes; in the current study we therefore replaced the innermost ten electrons of Zn with the (MEFIT, *R*) relativistic effective core potential.[224,225] The M05-2X/B2 density functional/basis set combination was chosen because of previous evaluations[182,223] that yielded very accurate results for zinc complexes. We note explicitly, however, that the main objective of using DFT in this study is to assess whether the EE-MB approximation can reproduce full (unfragmented) calculations. If so, one could, for example, use the EE-MB approximation with coupled cluster calculations on the fragments to approximate full coupled cluster calculations that are currently unaffordable.

All unfragmented calculations were performed using *Gaussian 09*.[226] All EE-MB calculations were carried out using MBPAC 2011–2,[227] an in-house software package that allows the user to define a particular fragmentation scheme and then accesses a locally modified version of *Gaussian 09* to perform the necessary monomer, dimer, and trimer calculations.

In the current work, we consider four pentacoordinate and two hexacoordinate Zn systems (**Table 7.1**). **Table 7.1** shows the possible ligands of Zn^{2+} in a fragment for each studied system.

152

Table 7.1. Systems considered in this work and the largest fragment in each.[a]

Full System	The Largest Fragment
$[Zn(NH_3)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnBC)
$[Zn(NH_3)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnAB)
$[Zn(NH_3)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnBC, ZnAB)
$[Zn(NH_3)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnBC, ZnAB)
$[Zn(NH_3)_3(OH)_2]$	$Zn(OH)_2$ (ZnAB)
$[Zn(NH_3)_3(OH)_2]$	$Zn(OH)_2$ (ZnAB)
$[Zn(NH_3)_3(OH)_2]$	$Zn(OH)_2$ (ZnAB)
$[Zn(Imd)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnBC)
$[Zn(Imd)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnAB)
$[Zn(Imd)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnBC, ZnAB)
$[Zn(Imd)_2(OH)_3]^-$	$Zn(OH)_2$ (ZnBC, ZnAB)
$[Zn(Imd)_3(OH)_2]$	$Zn(OH)_2$ (ZnAB)
$[Zn(Imd)_3(OH)_2]$	$Zn(OH)_2$ (ZnAB)
$[Zn(Imd)_3(OH)_2]$	$Zn(OH)_2$ (ZnAB)
fac isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnBC)
fac isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAC)
fac isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB)
fac isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB, ZnBC, ZnAC)
fac isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB, ZnBC, ZnAC)
fac isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB, ZnBC, ZnAC)
mer isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnBC)
mer isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB)
mer isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB, ZnBC)
mer isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB, ZnBC)
mer isomer of $[Zn(NH_3)_3(OH)_3]^-$	$Zn(OH)_2$ (ZnAB, ZnBC)

[a] When there is more than one row for a given system, it is because the largest fragment is not the same in all calculations on that system.

The pentacoordinate complexes are model compounds based on experimental X-ray structures of two Zn metalloenzyme active sites relevant to biology and to the drug design process: the anthrax toxin lethal factor (LF) (PDB ID 1PWU)[91] and the matrix metalloproteinase-3 (MMP-3) catalytic site (PDB ID 1SLN)[228]. In LF, the catalytic Zn is coordinated by two histidines and one glutamic acid, and in 1PWU, the zinc is also ligated by two oxygens in the hydroxamate zinc-binding group (ZBG) of the cocrystallized inhibitor, forming the complete pentacoordinate system. In MMP-3, the catalytic zinc is similarly coordinated by three histidine residues, and in 1SLN, the two remaining coordination sites are occupied by the carboxylate ZBG of the cocrystallized inhibitor. We specifically chose pentacoordinate systems that include ligands from potential drug scaffold ZBGs, in order to test the ability of EE-MB to reproduce bond dissociation energies that would parallel the interactions of small molecules with drug-target catalytic centers.

We created two simple and two extended models of each biocenter, where the simple models **7.1** and **7.2** (**Figure 7.1**) represent His residues by ammonias and Glu sidechains and ZBG oxygens by hydroxyls, yielding $[Zn(NH_3)_2(OH)_3]^-$ as a model for the anthrax toxin lethal factor and $[Zn(NH_3)_3(OH)_2]$ as a model for MMP-3.

Figure 7.1. Structures of truncated model Zn biocenter complexes: (**7.1**) the anthrax toxin lethal factor active site (LF) (1PWU.pdb), $[Zn(NH_3)_2(OH)_3]^-$, and (**7.2**) matrix metalloproteinase-3 (MMP-3, stromelysin-1) (1SLN.pdb), $[Zn(NH_3)_3(OH)_2]$.
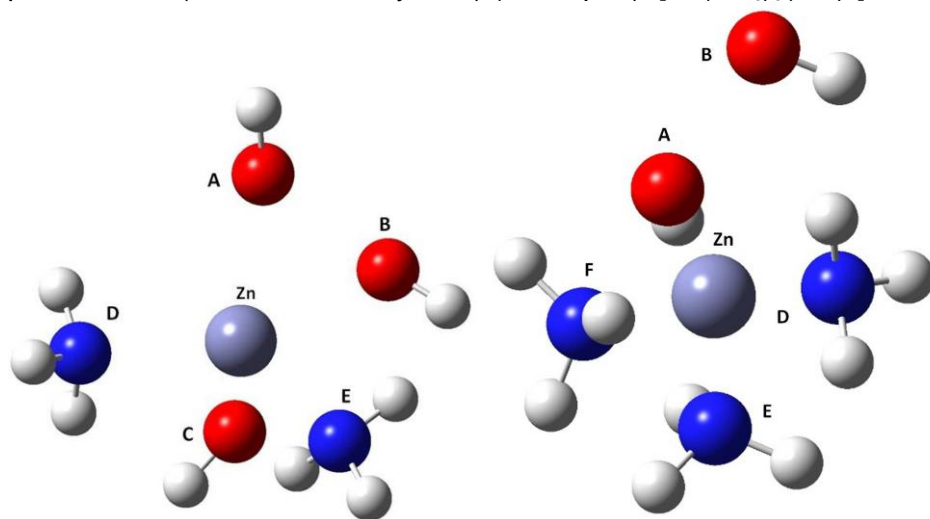
In the extended models **7.3** and **7.4** (**Figure 7.2**), the ammonias are replaced by full imidazole moieties while the hydroxyls are retained.

Figure 7.2. Structures of extended Zn biocenter complexes: (**7.3**) the anthrax toxin lethal factor active site (LF) (1PWU.pdb),[26] $[Zn(Imd)_2(OH)_3]^-$, and (**7.4**) matrix metalloproteinase-3 (MMP-3, stromelysin-1) (1SLN.pdb),[27] $[Zn(Imd)_3(OH)_2]$.
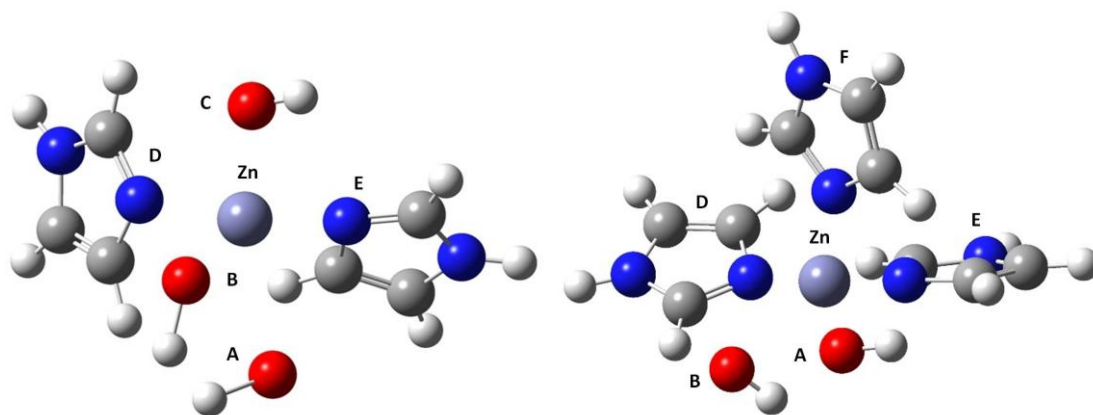
155

The hexacoordinate complexes examined here are the *fac* and *mer* isomers of [Zn(NH$_3$)$_3$(OH)$_3$]$^-$ (systems **7.5** and **7.6**, respectively, **Figure 7.3**). In total, these systems comprise four negatively charged and two neutral complexes. For systems **7.1–7.4**, all Zn–ligand distances were fixed at their experimental X-ray values. The hydrogen atoms on all NH$_3$ and OH ligands, and all Zn–ligand distances in systems **7.5** and **7.6,** were placed at standard distances and default orientations by the *GaussView*[229] program. The default N-H bond length in NH$_3$ is 1.00 Å, the default O-H distance is 0.96 Å, and for systems **7.5** and **7.6**, the default Zn–NH$_3$ bond length is 1.95 Å and the default Zn–OH distance is 1.91 Å. Default bond angles for ligand geometries in *GaussView* are obtained by AM1 optimizations. All structures are provided in Supporting Information.

Figure 7.3. Structures of two octahedral, hexacoordinate Zn complexes ([Zn(NH$_3$)$_3$(OH)$_3$]$^-$): (**7.5**) *fac* isomer, and (**7.6**) *mer* isomer.
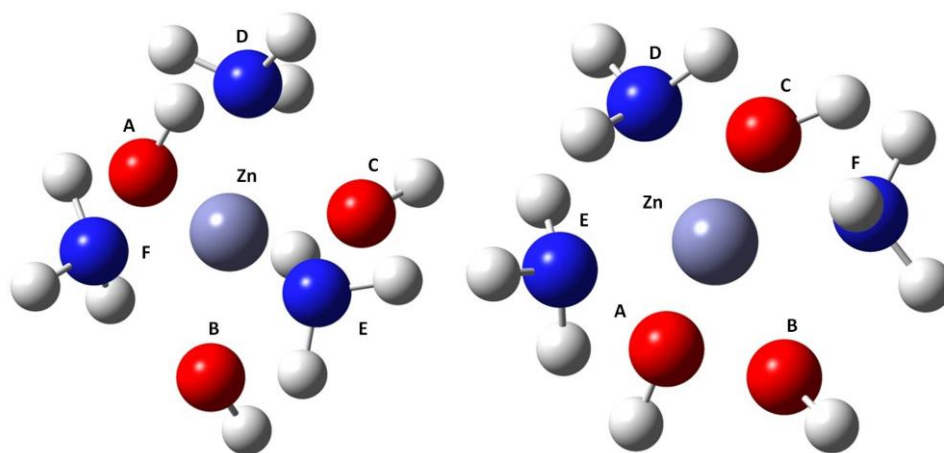
The quantity we calculate is a relative bond dissociation energy, which is defined as the energy to remove one of the ligands from the coordination system. As discussed in previous work,[218] this quantity is the sum of the energies of the two products (separated frozen fragments) minus the energy of the reactant, without including

vibrational energy (thus it is D_e, not D_0). When calculating the energies of a given dissociation product, the embedding charges of the other product are not included because the other product is considered to be infinitely separated.

After performing extensive calculations with various fragmentation schemes on systems **7.1** and **7.2**, we established four key fragmentation guidelines that, when applied, yielded the best results for all six systems in the current work. Next we present these four guidelines.

7.4.2 RESULTS AND DISCUSSION

First, our calculations on neutral and negatively charged Zn systems demonstrate, consistently with our previous findings,[218] that one must choose a fragmentation scheme where one of the monomers is Zn^{2+} coordinated to at least two ligands. We rationalize this rule in terms of partial atomic charges. In particular, the charge on unligated or monoligated Zn and even on biligated Zn is much larger than the charge on polyligated Zn; thus fragments consisting of unligated, monoligated, or—to a lesser extent—biligated Zn would not be representative of a portion of a larger system. But if each fragment already has two ligands on Zn, then even in dimers there are three ligands on Zn.

Second, as a corollary to rule 1, we do not dissociate bonds within fragments, as that would result in a product with Zn connected to a single ligand.

Third, at most one fragment can be charged. We rationalize rule 3 as eliminating the longest-range electrostatic effects.

Finally, our fourth guideline allows no trans coordination, i.e., Zn^{2+} cannot be coordinated within a fragment with two ligands that are *trans* to each other. This rule can

157

be understood as requiring links to be compact, although its origin is purely empirical at present.

We use the labeling scheme defined by **Figures 7.1–7.3**, in which A, B, and C (when present) are negatively charged hydroxyl ligands, and D, E, and F (when present) are neutral ligands. A consequence of rule 3 for the present study is that Zn^{2+} coupled with two hydroxyl groups must be part of the fragmentation scheme in all six complexes.

Rules 3 and 4, taken together, forbid applying EE-MB to dissociation of monomer B in **7.1** or monomer B in **7.6** because rule 3 would then require Zn to be coordinated within a fragment to ligands A and C in **7.1** and to ligands A and C in **7.6,** which in both cases would violate rule 4. After eliminating these processes that cannot be treated by the guidelines, we consider all the remaining processes, which may be classified as follows:

$$ZnABCDE^- \rightarrow ZnBCDE^- + A \qquad (R1)$$

$$ZnABCDE^- \rightarrow ZnABDE^- + C \qquad (R2)$$

$$ZnABCDE^- \rightarrow ZnABCE^- + D \qquad (R3)$$

$$ZnABCDE^- \rightarrow ZnABCD^- + E \qquad (R4)$$

$$ZnABDEF \rightarrow ZnABEF + D \qquad (R5)$$

$$ZnABDEF \rightarrow ZnABDF + E \qquad (R6)$$

$$ZnABDEF \rightarrow ZnABDE + F \qquad (R7)$$

$$ZnABCDEF^- \rightarrow ZnBCDEF^- + A \qquad (R8)$$

$$ZnABCDEF^- \rightarrow ZnACDEF^- + B \qquad (R9)$$

158

$$ZnABCDEF \rightarrow ZnABDEF + C \qquad (R10)$$

$$ZnABCDEF \rightarrow ZnABCEF + D \qquad (R11)$$

$$ZnABCDEF \rightarrow ZnABCDF + E \qquad (R12)$$

$$ZnABCDEF \rightarrow ZnABCDE + F \qquad (R13)$$
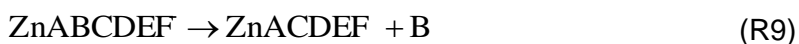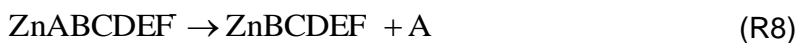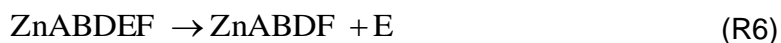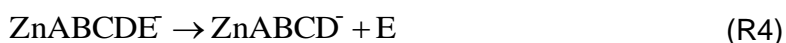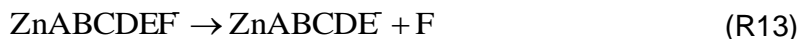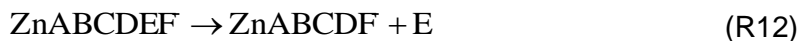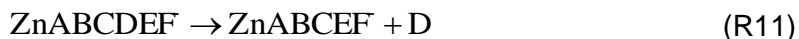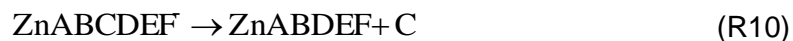
Keeping the four guidelines in mind, we considered dissociation processes (R1)-(R4) for systems **7.1** and **7.3**, we considered processes (R5)-(R7) for systems **7.2** and **7.4**, and we considered processes (R8)-(R13) for systems **7.5** and **7.6**, except for system **7.6,** where process (R9) was not considered because it would result in a monomer with ligands A and C positioned *trans* to each other.

Benchmark values for bond dissociation energies were obtained by full single-point calculations, i.e., without using the many-body approximation (see **Table 7.2**). Note that both the benchmark and the many-body calculations employ the same M05-2X/B2/MEFIT,*R* method. We measure "errors" as the deviation of the EE-MB results from the full calculations with the same method. If the error is small, then we assume that the method could be used with confidence for systems where full calculations on the entire system are impractically expensive or undoable, either due to system size (larger ligands, entire metalloenzymes) or due to using a higher level of electronic structure theory, for example coupled cluster theory.

Table 7.2. Benchmark bond energies (kcal/mol).

Reaction	System	Dissociated Bond	Bond Energy	Largest Zn Fragment(s) in Rxn
R1	**7.1**	Zn–A	35.12	ZnBC
R2	**7.1**	Zn–C	70.22	ZnAB

R3	**7.1**	Zn–D	19.32	ZnBC, ZnAB
R4	**7.1**	Zn–E	15.89	ZnBC, ZnAB
R5	**7.2**	Zn–D	-5.28	ZnAB
R6	**7.2**	Zn–E	-13.49	ZnAB
R7	**7.2**	Zn–F	6.19	ZnAB
R1	**7.3**	Zn–A	12.03	ZnBC
R2	**7.3**	Zn–C	57.01	ZnAB
R3	**7.3**	Zn–D	17.51	ZnBC, ZnAB
R4	**7.3**	Zn–E	20.68	ZnBC, ZnAB
R5	**7.4**	Zn–D	20.3	ZnAB
R6	**7.4**	Zn–E	-7.26	ZnAB
R7	**7.4**	Zn–F	9.53	ZnAB
R8	**7.5**	Zn–A	10.47	ZnBC
R9	**7.5**	Zn–B	9.17	ZnAC
R10	**7.5**	Zn–C	10.47	ZnAB
R11	**7.5**	Zn–D	-15.25	ZnAB, ZnBC, ZnAC
R12	**7.5**	Zn–E	-22.28	ZnAB, ZnBC, ZnAC
R13	**7.5**	Zn–F	-20.74	ZnAB, ZnBC, ZnAC
R8	**7.6**	Zn–A	26.98	ZnBC
R10	**7.6**	Zn–C	34.78	ZnAB
R11	**7.6**	Zn–D	-15.65	ZnBC, ZnAB
R12	**7.6**	Zn–E	-17.64	ZnBC, ZnAB
R13	**7.6**	Zn–F	-10.05	ZnBC, ZnAB

Tables 7.3-7.5 show the EE-MB bond-breaking energies and mean unsigned errors for all six systems. The systems are quite different, but the performance of the EE-3B method is uniformly good. For example, for **7.1**, the bond dissociation energies range from 16 to 70 kcal/mol, but the error of the EE-3B method is in the range 0.78–

0.85 kcal/mol for all four cases. The EE-3B method has a mean unsigned error (MUE) in bond dissociation energy of 0.82 kcal/mol for system **7.1**, 1.09 kcal/mol for system **7.2**, 1.03 kcal/mol for system **7.3**, and 0.83 kcal/mol for system **7.4**. It is encouraging that the EE-3B method performs very well for both the "truncated" model systems **7.1** and **7.2** and the "extended" model systems **7.3** and **7.4.** The MUEs in bond dissociation energies for the hexacoordinate systems **7.5** and **7.6** are comparable to those for the pentacoordinate systems, at 0.90 kcal/mol and 1.21 kcal/mol, respectively. As expected, the EE-PA method is less accurate, resulting in MUEs in bond dissociation energies ranging from 3.23 to 6.68 kcal/mol for the systems studied here. Altogether there are 25 cases in **Tables 7.3**, **7.4**, and **7.5**, and averaging the unsigned errors over all 25 gives an overall mean unsigned error of 5.10 kcal/mol for the EE-PA method but only 0.98 kcal/mol for the EE-3B method.

Table 7.3. Unsigned errors in bond energies (kcal/mol) for systems **7.1** and **7.3**

	EE-PA		EE-3B	
	7.1	**7.3**	**7.1**	**7.3**
R1	6.47	4.60	0.85	0.85
R2	6.81	8.71	0.78	1.24
R3	4.62	1.67	0.82	1.01
R4	1.38	4.01	0.81	1.04
mean	4.82	4.75	0.82	1.03

Table 7.4. Unsigned errors in bond energies (kcal/mol) for systems **7.2** and **7.4**

	EE-PA		EE-3B	
	7.2	**7.4**	**7.2**	**7.4**
R5	2.37	5.84	1.10	0.81
R6	5.41	5.57	1.09	0.85
R7	1.91	5.91	1.08	0.83
mean	3.23	5.77	1.09	0.83

Table 7.5. Unsigned errors in bond energies (kcal/mol) for systems **7.5** and **7.6**

	EE-PA		EE-3B	
	7.5	**7.6**	**7.5**	**7.6**
R8	8.80	6.57	0.59	1.38
R9	9.26		0.37	
R10	8.73	7.54	0.05	2.20
R11	4.51	3.06	1.44	0.16
R12	4.49	2.40	1.40	1.00
R13	4.32	2.53	1.54	1.30
mean[a]	6.68	4.42	0.90	1.21

[a] mean unsigned error for the five or six cases in the given column

7.4.3 CONCLUSION

The EE-3B method, when applied using our fragmentation guidelines, reliably yields bond dissociation energies within 1.21 kcal/mol of full-calculation DFT benchmark values, further demonstrating its utility and accuracy for neutral and negatively charged bio-inorganic structures, in addition to the positively charged systems evaluated in our previous work. Moreover, EE-MB exhibits high accuracy for "extended" active site models with His residues represented by full imidazole rings rather than ammonias, and for hexacoordinate Zn complexes, indicating its particular usefulness for larger metalloprotein active site systems for which full, high-level electronic structure calculations might be intractable or may incur a high computational cost. Finally, EE-MB is likely to find use in the drug discovery process; it performs very well for pentacoordinate systems representing a small-molecule drug lead coordinated to a catalytic metal center (which are otherwise quite challenging to model), and it can also be used to obtain key parameters such as bond dissociation energies that can be imported into molecular mechanics force fields to increase the accuracy of simpler and less costly calculations on macromolecular drug targets.

7.5 EE-MB-CE

7.5.1 METHODS

Most of the calculations were done using the MP2 level of theory. The B2[182] basis set was used for the Zn calculations, and the def2-TZVP[230] basis set was used for the Cd calculations. Both B2 and def2-TZVP are polarized valence-triple-zeta basis sets for Zn and Cd, respectively. The MP2 level was chosen because it is the least expensive of the post-Hartree–Fock methods, allowing for direct comparison of the EE-MB and the EE-MB-CE energies to the full MP2 energies. Such direct comparison for all clusters and fragmentation schemes studied in this work would not be practical at a reasonable cost with more expensive post-Hartree–Fock methods such as CCSD(T). Nevertheless, we present a CCSD(T) calculation on the system **7.1** to test the performance of the methods with other correlated methods. Our earlier published work[182,223] showed that incorporating relativistic effects on core electrons significantly increased the accuracy of geometric and energetic calculations for Zn coordination complexes; in the current study we therefore replaced the innermost ten electrons of Zn and 28 core electrons of Cd with the relativistic effective core potential (RECP).[224,225,231]

Charges are calculated for each fragment at the geometry of that monomer in the overall system. For example, if we are calculating the energy of ZnABCDEF, where A, B, C, D, E, and F are ligands, and if one of the fragments is ZnBC, we calculate the partial atomic charges of ZnBC by removing A, D, E, and F from the system. Here we calculate charges using Merz-Kollman (MK) electrostatic-potential fitting,[221] as in previous work on Zn compounds.[218,219] Charges for CCSD(T) calculations on system **7.1** were obtained using M05-2X level theory.[222]

163

All benchmark calculations were performed using *Gaussian 09*.[226] All EE-MB and EE-MB-CE calculations were carried out using MBPAC 2011–5,[232] a freely available software package that allows the user to define a particular fragmentation scheme and then accesses *Gaussian 09* to perform the necessary monomer, dimer, and trimer calculations.

In the current work, we consider seven pentacoordinate and three hexacoordinate Zn systems (see **Table 7.6**). Two pentacoordinate complexes are model compounds based on experimental X-ray structures of Zn metalloenzyme active sites relevant to the drug design process: the anthrax toxin lethal factor (LF) (PDB ID 1PWU)[91] active site, and the matrix metalloproteinase-3 (MMP-3) catalytic site (PDB ID 1SLN)[228] (systems **7.1** and **7.2**, respectively, **Figure 7.1**). These model structures are the same as those reported in our previously published work.[219,223] System **7.7** (**Figure 7.4**) is a model compound based on the X-ray structure of the matrix metalloproteinase-7 (MMP-7, also known as matrilysin) active site, co-crystallized with sulfodiimine (PDB ID: 1MMR).[233] In MMP-7, the catalytic Zn is coordinated by three histidines, and in 1MMR, the two remaining coordination sites are occupied by imine nitrogens in the co-crystallized inhibitor.

Figure 7.4. Structure of truncated model Zn biocenter complex: (**7.7**) matrix metalloproteinase-7) (MMP-7, matrilysin) (1MMR.pbs), ([Zn(NH$_3$)$_5$]$^{2+}$).
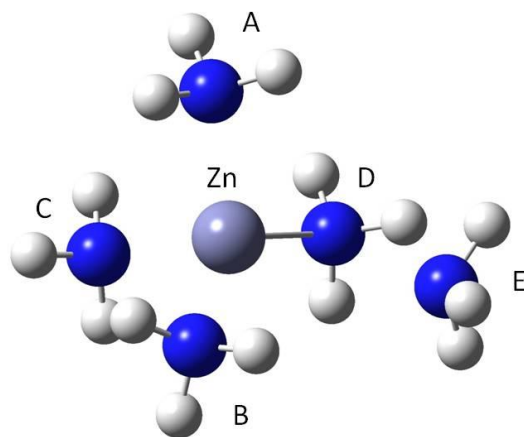
We use these systems to test the ability of the EE-MB-CE method to reproduce Zn-ligand bond dissociation energies of biological Zn-containing active sites. System **7.7** is also used to demonstrate the advantage of our newly developed fragmentation strategy over previously published fragmentation guidelines.[219] We also consider three other pentacoordinate systems: Zn(NH$_3$)$_5$$^{2+}$, Zn(H$_2$O)$_5$$^{2+}$, and Zn(H$_2$O)$_4(OH)^+$ (systems **7.8**, **7.9**, and **7.10**, respectively, **Figure 7.5**). These are used to test the applicability of the EE-MB-CE method to positively charged systems.

Figure 7.5. (**7.8**, top left) Structure of ($[Zn(NH_3)_5]^{2+}$) (**7.9**, top right) structure of ($[Zn(H_2O)_5]^{2+}$), and (**7.10**, bottom) structure of ($[Zn(H_2O)_4(OH)]^+$).
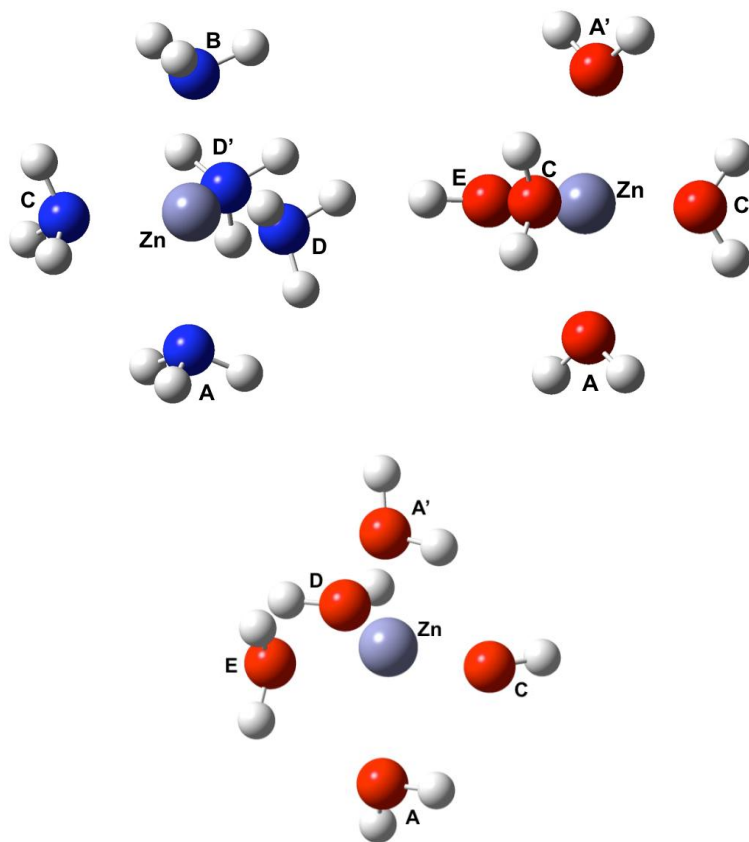
Calculations on a Cd complex (system **7.11, Figure 7.6**) were performed to test the applicability of our methods to a transition metal other than Zn. System **7.11** (**Figure 7.6**) is a model compound based on the X-ray structure of the cadmium carbonic anhydrase active site (CDCA1-R2) (PDB ID 3BOB).[234] In CDCA1-R2, the catalytic Cd is coordinated by two cysteine residues, a histidine residue, and two water molecules.

Figure 7.6. Structure of truncated model Cd biocenter complex: (**7.11**) cadmium carbonic anhydrase (3BOB), $[Cd(H_2O)_2(SH)_2(NH_3)]$.
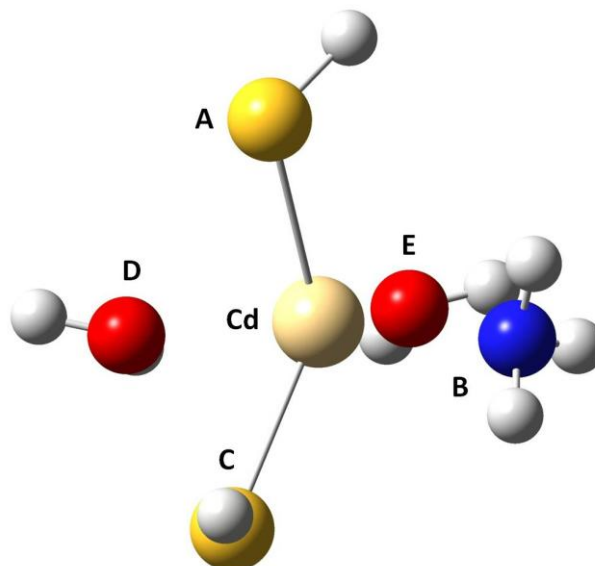
Two of the three hexacoordinate complexes examined here are the *fac* and *mer* isomers of $[Zn(NH_3)_3(OH)_3]^-$ (systems **7.5** and **7.6**, respectively, **Figure 7.3**), used here to evaluate the performance of the EE-MB-CE method with larger, hexacoordinate systems. We also use hexacoordinate $[Zn(OH)_6]^{4-}$ (system **7.12**, **Figure 7.7**) to test the applicability of the methods to symmetric Zn complexes. Zn – ligand distances in system **7.12** were placed at 2.10 Å. In order to make system **7.12** completely symmetric angles Zn – O – H were made equal to 180 degrees.

Figure 7.7. (**7.12**) Structure of ([Zn(OH)$_6$]$^{4-}$).
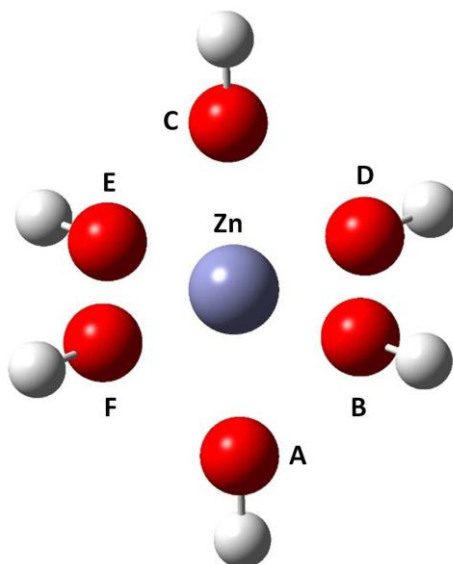
The labeling schemes of all systems are consistent with previously published work.[218,219] In total, these systems comprise two neutral, four positively charged, and four negatively charged complexes. All structures are provided in Supporting Information.

We use the labeling scheme defined in **Figures 7.4–7.7**, in which A, B, and C (when present) are negatively charged hydroxyl ligands and D, E, and F (when present) are neutral ligands. Each of the coordination complexes in Figure 3 has the structure of an irregular trigonal bipyramid, with axial ligands A and B and equatorial ligands C, D, and E. If r_X denotes the distance from the non-hydrogen atom of a ligand to Zn, we label the atoms so that $r_A \le r_B$ and $r_C \le r_D \le r_E$. However, if $r_B = r_A$, then B is also called A´; if $r_D = r_C$, then D is also called C´, and if $r_E = r_D$, then E is also called D´.

The quantity we calculate is an instantaneous bond dissociation energy, which is defined as the energy to remove one of the ligands from the coordination system at a given predefined geometry (if this were the equilibrium geometry, and if the separated

subsystems were re-optimized after removal of the ligand whose bond is being broken, then the instantaneous bond dissociation energy would be the equilibrium bond dissociation energy). As discussed previously,[218,219] the instantaneous bond dissociation energy is the sum of the energies of the two products (separated frozen fragments) minus the energy of the reactant, without reoptimization and without including vibrational energy. When calculating the energies of a given dissociation product, the embedding charges of the other product are not included because the other product is considered to be infinitely separated.

Our EE-MB-CE calculations on neutral, negatively, and positively charged Zn and Cd systems demonstrate, consistently with our previous findings,[218,219] that one must choose a fragmentation scheme where one of the monomers is Zn^{2+} or Cd^{2+} coordinated to at least two ligands. We rationalize this rule in terms of partial atomic charges. In particular, the charge on unligated or monoligated Zn and even on biligated Zn is much larger than the charge on polyligated Zn; thus fragments consisting of unligated, monoligated, or—to a lesser extent—biligated Zn are not precisely representative of a portion of a larger system. But if each fragment already has two ligands on Zn, then even in dimers there are three ligands on Zn. Thus, we only consider fragmentation schemes where one of the fragments is Zn^{2+} or Cd^{2+} with two ligands and the other fragments are individual ligands. Therefore, for pentacoordinate systems, there are ten possible ways to fragment a system. For example, for pentacoordinate system **7.1**, the largest fragments for all ten fragmentation schemes are: ZnAB, ZnAC, ZnAD, ZnAE, ZnBC, ZnBD, ZnBE, ZnCD, ZnCE, and ZnDE. However, there are fifteen different ways to fragment a hexacoordinate system. The largest fragments in each fragmentation for each hexacoordinate system are as follows: ZnAB, ZnAC, ZnAD,

ZnAE, ZnAF, ZnBC, ZnBD, ZnBE, ZnBF, ZnCD, ZnCE, ZnCF, ZnDE, ZnDF, and ZnEF.

Table 7.6 shows the possible ligands of Zn^{2+} and Cd^{2+} in a fragment for each studied system. We performed calculations utilizing all possible fragmentation schemes for all systems.

Table 7.6. Systems considered in this work and types of Zn^{2+} and Cd^{2+}-containing fragments in each[a]

Full system	types of largest fragments[a]
7.1 $([Zn(NH_3)_2(OH)_3]^-)$	$Zn(OH)_2$, $[Zn(OH)(NH_3)]^+$, $[Zn(NH_3)_2]^{2+}$
7.2 $[Zn(NH_3)_3(OH)_2]$	$Zn(OH)_2$, $[Zn(OH)(NH_3)]^+$, $[Zn(NH_3)_2]^{2+}$
7.5 *fac* isomer of $([Zn(NH_3)_3(OH)_3]^-)$	$Zn(OH)_2$, $[Zn(OH)(NH_3)]^+$, $[Zn(NH_3)_2]^{2+}$
7.6 *mer* isomer of $([Zn(NH_3)_3(OH)_3]^-)$	$Zn(OH)_2$, $[Zn(OH)(NH_3)]^+$, $[Zn(NH_3)_2]^{2+}$
7.7 $([Zn(NH_3)_5]^{2+})$	$[Zn(NH_3)_2]2^+$
7.8 $([Zn(NH_3)_5]^{2+})$	$[Zn(NH_3)_2]2^+$
7.9 $([Zn(H_2O)_5]^{2+})$	$[Zn(H_2O)_5]2^+$
7.10 $([Zn(H_2O)_4(OH)]^+)$	$[Zn(OH)(H_2O)]^+$, $[Zn(H_2O)_2]^{2+}$
7.11 $[Cd(H_2O)_2(SH)_2(NH_3)]$	$[Cd(H_2O)_2]^{2+}$, $[Cd(SH)_2]$, $[Cd(H_2O)(SH)]^+$ $[Cd(H_2O)(NH_3)]^{2+}$, $[Cd(SH)(NH_3)]^+$
7.12 $([Zn(OH)_6]^{4-})$	$Zn(OH)_2$

[a]For example, for system **7.1**, there are three possible fragmentations of the type $Zn(OH)_2$, one of the type $[Zn(NH_3)_2]^{2+}$, and six of the type $[Zn(OH)(NH_3)]^+$. In **Table 7.9**, we average results over all ten possible fragmentations, whereas in **Table 7.13**, we give results for the single scheme selected by our new fragmentation strategy.

Each fragmentation scheme included all possible bond dissociation processes. For pentacoordinate systems we consider breaking 5 bonds whereas for hexacoordinate systems we consider breaking 6 bonds. For example, consider system **7.1** where each fragmentation scheme was used within the EE-MB and EE-MB-CE approximations to compute each of the following instantaneous bond dissociation processes:

$$ZnABCDE^- \rightarrow ZnBCDE^- + A \qquad (R1)$$

$$ZnABCDE^- \rightarrow ZnACDE^- + B \qquad (R2)$$

$$ZnABCDE^- \rightarrow ZnABDE^- + C \qquad (R3)$$

$$ZnABCDE^- \rightarrow ZnABCE^- + D \qquad (R4)$$

$$ZnABCDE^- \rightarrow ZnABCD^- + E \qquad (R5)$$
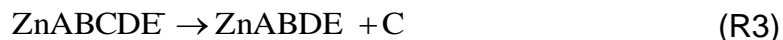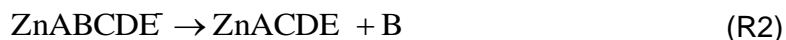
For 10 systems there are a total of 53 bond dissociation processes labeled from R1 to R53. Benchmark values for R1-R53 bond dissociation energies were obtained by full single-point calculations, i.e., without using the many-body approximation (see **Table 7.7**).

Table 7.7. MP2 benchmark bond dissociation energies (kcal/mol) for bonds in each model complex

Reaction	System	Dissociated bond	Bond energy
R1	**7.1**	Zn–A	33.70
R2	**7.1**	Zn–B	28.14
R3	**7.1**	Zn–C	67.19
R4	**7.1**	Zn–D	17.06
R5	**7.1**	Zn–E	14.12
R6	**7.2**	Zn–A	116.45

R7	**7.2**	Zn–B	53.31
R8	**7.2**	Zn–D	7.66
R9	**7.2**	Zn–E	-11.54
R10	**7.2**	Zn–F	8.38
R11	**7.5**	Zn–A	9.67
R12	**7.5**	Zn–B	9.81
R13	**7.5**	Zn–C	9.67
R14	**7.5**	Zn–D	-14.69
R15	**7.5**	Zn–E	-22.45
R16	**7.5**	Zn–F	-20.91
R17	**7.6**	Zn–A	25.89
R18	**7.6**	Zn–B	11.66
R19	**7.6**	Zn–C	32.37
R20	**7.6**	Zn–D	-14.96
R21	**7.6**	Zn–E	-16.91
R22	**7.6**	Zn–F	-8.85
R23	**7.7**	Zn–A	51.27
R24	**7.7**	Zn–B	51.19
R25	**7.7**	Zn–C	45.12
R26	**7.7**	Zn–D	27.18
R27	**7.7**	Zn–E	2.03
R28	**7.8**	Zn–A	33.06
R29	**7.8**	Zn–B	32.86
R30	**7.8**	Zn–C	45.35
R31	**7.8**	Zn–D	45.35
R32	**7.8**	Zn–D´	45.28
R33	**7.9**	Zn–A	34.34
R34	**7.9**	Zn– A´	33.38

R35	**7.9**	Zn–C	45.27
R36	**7.9**	Zn–C´	39.52
R37	**7.9**	Zn–E	43.87
R38	**7.10**	Zn–A	23.25
R39	**7.10**	Zn–A´	23.25
R40	**7.10**	Zn–C	260.83
R41	**7.10**	Zn–D	26.05
R42	**7.10**	Zn–E	8.66
R43	**7.11**	Cd–A	147.65
R44	**7.11**	Cd–B	12.51
R45	**7.11**	Cd–C	144.68
R46	**7.11**	Cd–D	1.08
R47	**7.11**	Cd–E	9.11
R48	**7.12**	Zn–A	-257.79
R49	**7.12**	Zn–B	-257.79
R50	**7.12**	Zn–C	-257.79
R51	**7.12**	Zn–D	-257.79
R52	**7.12**	Zn–E	-257.79
R53	**7.12**	Zn–F	-257.79

Table 7.8 shows CCSD(T) benchmark bond energies for system **7.1**. We measure "errors" as the absolute deviation of the EE-MB and EE-MB-CE results from the full calculations with the same method.

Table 7.8. CCSD(T) benchmark bond dissociation energies (kcal/mol) for every bond in system **7.1**

Reaction	System	Dissociated bond	Bond energy
R1	**7.1**	Zn–A	34.98
R2	**7.1**	Zn–B	29.20
R3	**7.1**	Zn–C	67.48
R4	**7.1**	Zn–D	16.60
R5	**7.1**	Zn–E	14.08

7.5.2 RESULTS AND DISCUSSION

For each method, we consider 350 bond dissociation energy calculations for pentacoordinate systems **7.1, 7.2, 7.7, 7.8, 7.9, 7.10, 7.11** (seven systems, each with five bonds and ten different fragmentation schemes) and 270 bond dissociation energy calculations for hexacoordinate systems **7.5, 7.6,** and **7.12** (three systems, each with six bonds and 15 fragmentation schemes). And we compute final errors averaged over a total of 620 bond dissociation energies for each method. **Table 7.9** shows the EE-MB and the EE-MB-CE mean signed errors (MSEs) and mean unsigned errors (MUEs) in the instantaneous bond energies (kcal/mol) averaged over the 620 cases of bond dissociation.

Table 7.9. EE-MB and EE-MB-CE mean signed and unsigned errors in bond energies (kcal/mol) for all ten Zn and Cd complexes averaged over 10-to-15 fragmentation schemes and averaged over all five or six bonds being broken[a]

System	EE-PA		EE-3B		EE-PA-CE		EE-3B-CE	
	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
7.1	-1.26	7.28	0.64	2.07	-1.35	3.16	0.34	0.95
7.2	0.94	14.97	0.70	6.85	-1.86	8.06	1.14	3.94
7.5	2.41	4.82	-0.53	2.08	0.12	2.64	-0.24	1.54

7.6	0.95	2.73	0.42	1.65	-0.48	1.80	0.17	0.98
7.7	-2.88	4.10	0.49	0.78	-0.39	1.48	0.02	0.42
7.8	-3.71	4.20	0.47	0.74	-0.68	1.75	0.05	0.47
7.9	-0.06	2.69	-0.33	1.30	-0.17	1.10	-0.09	0.45
7.10	0.31	5.31	-0.85	2.64	0.16	2.33	-0.34	0.82
7.11	-17.07	18.69	-0.89	1.40	-1.99	2.39	-0.53	0.65
7.12	6.50	6.52	-1.74	2.16	3.12	3.12	-0.59	0.82
mean[b]	-0.48	6.66	-0.25	2.13	-0.08	2.73	-0.05	1.10

[a]MSE is mean signed error; MUE is mean unsigned error.
[b]averaged over the 620 combinations of system, bond, and fragmentation scheme, not over the ten rows

The EE-3B MUEs range from 0.74 to 2.64 kcal/mol except system **7.2**, which has a MUE of 6.85 kcal/mol. Note that the EE-3B MUE for all ten systems is 2.13 kcal/mol. The EE-3B MUEs in bond dissociation energies of the hexacoordinate systems **7.5** and **7.6**, at 2.08 kcal/mol and 1.65 kcal/mol, are comparable to those of the pentacoordinate systems, which is an encouraging result. Note that the EE-3B MUE for Cd^{2+} system **7.11** is 1.40 kcal/mol. As expected, the EE-PA method is less accurate, resulting in MUEs in bond dissociation energies ranging from 2.69 to 18.69 kcal/mol for the systems studied here. Averaging the MUEs over all ten Zn^{2+} and Cd^{2+} systems gives an overall MUE of 6.66 kcal/mol for the EE-PA method.

Now let us consider the EE-MB-CE method. **Table 7.9** also shows the EE-MB-CE MSEs and MUEs for bond-breaking energies for all possible fragmentation schemes for all ten Zn^{2+} and Cd^{2+} systems. As expected, the EE-3B-CE approximation is more accurate, resulting in MUEs in bond dissociation ranging from 0.42 to 1.54 kcal/mol for all pentacoordinate and hexacoordinate systems except **7.2**, which has a MUE of 3.94 kcal/mol. The EE-3B-CE MUE for Cd^{2+} system **7.11** is 0.65 kcal/mol, which is

comparable to those for the Zn – containing pentacoordinate systems. The EE-3B-CE MUEs in bond dissociation energies for the hexacoordinate systems **7.5** and **7.6** are comparable to those for the pentacoordinate systems, at 1.54 kcal/mol and 0.98 kcal/mol, respectively. As expected, the EE-PA-CE method is less accurate, resulting in MUEs in bond dissociation energies ranging from 1.10 to 3.16 kcal/mol for the systems studied here except system **7.2**, which has a MUE of 8.06 kcal/mol. Averaging the MUEs over all ten Zn^{2+} and Cd^{2+} systems including **7.2** gives an overall MUE of 2.73 kcal/mol for the EE-PA-CE method, but only 1.10 kcal/mol for the EE-3B-CE method.

Based on the results in **Table 7.9**, the many-body methods can be ranked in order of increasing MUE as EE-3B-CE, EE-3B, EE-PA-CE, and EE-PA. Overall, both the EE-MB and the EE-MB-CE methods perform well, even for the hexacoordinate Zn complexes, indicating their usefulness for larger metalloprotein systems for which high-level electronic structure calculations on the entire system might not be feasible.

Table 7.10 provides a more detailed view of the results for system **7.2**, with MSEs and MUEs in bond energies for system **7.2** for all ten possible fragmentations. The ZnFE, ZnDE, and ZnDF fragmentation schemes have the largest errors, with MUEs of 12.47, 10.94, and 9.07 kcal/mol respectively for EE-3B and 6.71, 6.19, and 5.23 kcal/mol respectively for EE-3B-CE. Thus, any combination of Zn with ammonia ligands in a fragment yields poor results. In contrast, the ZnAB fragmentation scheme, where Zn is combined with two hydroxyl ligands in a fragment, yields the best results; note that for this fragmentation, the MUE of EE-3B-CE is only 0.84 kcal/mol, whereas that of the EE-3B is a reasonable 3.28 kcal/mol (**Table 7.10**). Our analysis suggests that the ZnAB fragmentation yields better results than the other fragmentation schemes because there is a large interatomic Coulomb interaction between the O atom of A and the O atom of B,

176

and the ZnAB fragmentation treats these interactions entirely by quantum mechanics without using point charges because it places the A and B groups in a single fragment. This observation was corroborated by additional calculations (not reported in detail here) that resulted in a lower error when the distance between ligands A and B was increased in system **7.2** without altering any other distances when the ZnAB fragmentation scheme was not used, and that resulted in an larger error when decreasing the distance between the ligands A and B when the ZnAB fragmentation scheme was not used. This is consistent with the major part of the error coming from the A–B interaction when A and B are not in a single fragment.

Table 7.10. EE-MB and EE-MB-CE mean signed and unsigned errors in bond energies (kcal/mol) for system **7.2,** for all ten fragmentation schemes

largest fragment	EE-PA		EE-3B		EE-PA-CE		EE-3B-CE	
	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
ZnFE	-7.38	17.38	1.22	12.47	-4.21	10.56	1.63	6.71
ZnDE	-5.81	18.19	0.70	10.94	-3.58	11.77	0.84	6.19
ZnDF	-7.10	17.21	2.70	9.07	-4.26	13.17	2.42	5.23
ZnAD	-11.12	12.36	4.88	6.87	-7.30	7.84	3.31	4.54
ZnBD	17.45	17.45	-3.31	3.31	3.81	5.51	-0.46	1.37
ZnAE	-7.93	11.90	1.99	7.92	-4.41	6.35	1.29	4.77
ZnBE	5.74	10.21	0.87	5.54	1.11	8.28	0.17	3.79
ZnAF	-8.05	10.99	3.62	6.62	-5.03	6.32	2.44	4.21
ZnBF	14.62	14.98	-2.43	2.43	4.09	6.85	-0.75	1.76
ZnAB	19.01	19.01	-3.28	3.28	1.15	3.99	0.49	0.84
mean[c]	0.95	14.97	0.70	6.85	-1.86	8.06	1.14	3.94

[c] mean signed and unsigned errors for the ten rows in the given column, where each row contains an average over five bond dissociation energies

Following these observations, we developed a fragmentation strategy designed to minimize the error. In order to obtain accurate energies for Zn systems with the EE-MB and EE-MB-CE methods, one should locate the two ligands in a system that have the strongest Coulomb interaction with each other, and combine them with Zn^{2+} or Cd^{2+} in a single fragment. In order to quantify "the strongest Coulomb interaction" we used the monomer embedding charges to calculate Coulomb interactions of the partial atomic charges in one ligand with those in another. For example, consider ligands A, B, D, E, F of system **7.2**. First, we calculate the monomeric partial atomic charges using the method described above (that is, we perform an MK partial charge analysis on each isolated ligand in the geometry it has in the cluster). Then those charges are used to calculate absolute maximum atom-atom Coulomb interaction between each pair of ligands. We define the absolute atom-atom Coulomb interaction (V_{ij}^{AB}) between atom i of ligand A and atom j of ligand B as

$$V_{ij}^{AB} = \left| \frac{q_i^A q_j^B}{r_{ij}} \right|, i \in A \text{ and } j \in B, A \neq B$$

where q_i^A is the charge on i, and q_j^B is the charge on j, and r_{ij} is the distance from i to j.

The maximum Coulomb interaction between two fragments is defined as the largest of all of the absolute atom-atom Coulomb interaction energies between a partial atomic charge in one fragment and a partial atomic charge on an atom of a different fragment. For system **7.2**, ligands A and B have the largest absolute maximum Coulomb interaction out of all of the possible pairs of ligands in the complex. Therefore, our strategy places these two ligands together with Zn^{2+} in a fragment; this leaves the other ligands as individual fragments. This strategy produces a preferred fragmentation

scheme for each of the ten complexes. Note that our prescription can be stated in a way that suggests even greater generality: of all possible fragmentation schemes that combine two ligands in the same fragment with the transition metal, one should apply the fragmentation scheme that minimizes the absolute maximum Coulomb interaction between any two remaining ligands that are treated as individual fragments.

Table 7.11 presents maximum Coulomb interactions between every ligand in all ten systems. For example, ligands A and B in system **7.1** have the maximum absolute Coulomb interaction for that system, which is 0.56 e (where e is the charge on a proton). Thus, ligands A and B in system **7.1** should be combined with Zn^{2+} in a fragmentation. However, in hexacoordinate system **7.5**, there are three identical maximum Coulomb interactions: A and C, A and B, and B and C. Any of those ligands can be combined with Zn^{2+} in a fragment. In **Table 7.13** we report the average of the three (ZnAB, ZnAC, and ZnBC) fragmentations for system **7.5**. In system **7.6**, ligands A and B, and B and C have identical maximum Coulomb interactions. Thus, the average of the two (ZnAB and ZnBC) fragmentations was reported in **Table 7.13** for system **7.6**.

Table 7.11. Absolute Coulomb interactions between fragments in all ten systems (in e)

System	Ligand	B or A or A'	C or C'	D or C	E or D'	F
7.1	A	0.56	0.33	0.36	0.33	
	B		0.48	0.28	0.40	
	C			0.35	0.35	
	D				0.27	
7.2	A	0.70		0.40	0.45	0.43
	B			0.37	0.27	0.32
	D				0.35	0.34
	E					0.35
7.5	A	0.57	0.57	0.45	0.32	0.45
	B		0.57	0.45	0.45	0.32
	C			0.32	0.45	0.45
	D				0.36	0.36
	E					0.36
7.6	A	0.57	0.40	0.45	0.45	0.45

179

	B		0.57	0.32	0.45	0.45
	C			0.45	0.45	0.45
	D				0.36	0.36
	E					0.25
7.7	A	0.30	0.29	0.29	0.25	
	B		0.30	0.26	0.28	
	C			0.33	0.21	
	D				0.41	
7.8	A	0.22	0.3221	0.3178	0.3180	
	B		0.3174	0.3218	0.3219	
	C			0.27	0.27	
	D				0.27	
7.9	A'	0.1125	0.1593	0.1593	0.1614	
	A		0.1657	0.1654	0.1626	
	C'			0.1352	0.1356	
	C				0.1354	
7.10	A	0.14	0.316	0.19	0.16	
	A'		0.320	0.19	0.16	
	C			0.22	0.30	
	D				0.18	
7.11	A	0.31	0.27	0.23	0.18	
	B		0.32	0.14	0.21	
	C			0.18	0.18	
	D				0.15	
7.12	A	0.52	0.36	0.52	0.52	0.52
	B		0.52	0.52	0.36	0.52
	C			0.52	0.52	0.52
	D				0.52	0.36
	E					0.52

System **7.12** is a unique symmetric system where only two fragmentation types are possible: *trans* and *cis*. *Trans* fragmentation schemes are those where Zn^{2+} is paired with ligands that are directly across from each other, and *cis* fragmentations are those where Zn^{2+} is paired with ligands that are adjacent to one another. In system **7.12**, ZnAC, ZnBE, and ZnDF are *trans* fragmentations whereas ZnAB, ZnAD, ZnAE, ZnAF, ZnBC, ZnBD, ZnBF, ZnCD, ZnCE, ZnCF, ZnDE, and ZnEF are *cis* fragmentations. **Table 7.12** presents MUEs for *trans* fragmentation ZnAC and *cis*

fragmentation ZnAB. The EE-3B-CE MUEs for *trans* and *cis* fragmentations are 0.58 kcal/mol and 0.88 kcal/mol, respectively.

Table 7.12. EE-MB and EE-MB-CE mean signed and unsigned errors in bond dissociation energies (kcal/mol) for system **7.12**

System	Largest fragment	EE-PA		EE-3B		EE-PA-CE		EE-3B-CE	
		MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
7.12	ZnAB (*cis*)	7.14	7.14	-2.44	2.44	3.45	3.45	-0.88	0.88
7.12	ZnAC (*trans*)	3.92	4.02	1.06	1.06	1.80	1.80	0.58	0.58
mean[d]		6.50	6.52	-1.74	2.16	3.12	3.12	-0.59	0.82

[d]mean signed and unsigned errors for the 15 fragmentation schemes, where 12 are *cis* and 3 are *trans* fragmentations.

Table 7.13 shows results for all ten systems where, for each complex, we apply the new strategic fragmentation scheme rather than averaging over all possible fragmentations. We see that the MUE for the EE-3B method decreases from 2.13 to 1.42 kcal/mol, and the MUE for the EE-3B-CE method decreases from 1.10 to 0.59 kcal/mol, when we use the preferred fragmentation. Note that the EE-PA-CE MUE decreases from 2.73 kcal/mol to 1.93 kcal/mol. The EE-PA MUE decreases from 6.66 kcal/mol to 6.14 kcal/mol.

Table 7.13. EE-MB and EE-MB-CE mean signed and unsigned errors in bond energies (kcal/mol) for Zn and Cd systems using the fragmentation strategy selected by our new criterion

System	Largest fragment	EE-PA		EE-3B		EE-PA-CE		EE-3B-CE	
		MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
7.1	ZnAB	0.44	4.63	0.31	1.41	-0.56	1.74	0.24	0.50
7.2	ZnAB	19.01	19.01	-3.28	3.28	1.15	3.99	0.49	0.84
7.5	Zn(AB, AC, BC)	5.79	5.79	-1.77	2.11	0.82	2.11	-0.28	0.97

7.6	Zn(AB, BC)	3.70	3.80	-1.14	1.83	0.46	1.65	-0.27	0.82
7.7	ZnDE	-3.77	3.99	-0.07	0.58	0.47	1.34	-0.54	0.70
7.8	ZnAC	0.15	2.06	0.40	0.52	-0.07	1.57	-0.14	0.55
7.9	ZnAC'	0.04	2.69	-0.53	1.32	-0.09	1.04	-0.16	0.47
7.10	ZnA'C	0.37	1.34	-0.03	0.62	-0.19	0.65	0.03	0.18
7.11	CdBC	-9.24	11.15	0.09	0.24	-1.45	1.75	0.03	0.03
7.12	ZnAB(*cis*)	7.14	7.14	-2.44	2.44	3.45	3.45	-0.88	0.88
mean[e]		2.36	6.14	-0.86	1.42	0.40	1.93	-0.15	0.59

[e] mean signed and unsigned errors for the ten systems in the given column

Table 7.14 shows MSEs and MUEs for both EE-MB and EE-MB-CE methods obtained by CCSD(T) level calculations on system **7.1**. As expected, the EE-3B-CE approximation is more accurate, resulting in a MUE for bond dissociation of 0.30 kcal/mol, whereas the EE-3B MUE is 1.26 kcal/mol.

Table 7.14. EE-MB and EE-MB-CE mean signed and unsigned errors in bond energies (kcal/mol) for Zn system **7.1** using CCSD(T) level theory and the fragmentation strategy selected by our new criterion

System	Largest fragment	EE-PA		EE-3B		EE-PA-CE		EE-3B-CE	
		MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
7.1	ZnAB	8.54	8.73	0.31	1.26	7.52	7.52	0.21	0.30

Previously,[219] we reported four fragmentation guidelines for Zn systems. According to those rules, Zn should be combined with at least two ligands in a fragment in order to have a representative charge distribution at the Zn center, and three other rules were proposed: one should not dissociate Zn–ligand bonds that are in the same fragment as Zn, one should not include more than one charged fragment, and Zn should not be in a fragment with ligands that are *trans* to each other. In the present work, the

final three of these rules are superseded by one new, more general guideline, i.e., of all possible fragmentation schemes under consideration, one should use the one that minimizes the maximum Coulomb interaction between atoms in two different fragments. Since here we consider only schemes that combine two ligands with Zn^{2+} or Cd^{2+} in one of the fragments and then treat all other ligands as individual fragments, the rule reduces for the present case to: in order to calculate accurate metal-ligand bond dissociation energies, one should find two ligands that exhibit the strongest Coulomb interaction with each other, and combine them with Zn or Cd into a single fragment.

In addition to yielding lower errors, the fragmentation prescription advanced here allows the treatment of systems for which our previous guidelines were ambiguous. For example, consider system **7.7** (see **Figure 7.4**), which features five ammonia ligands and for which none of our previous guidelines offers an optimal fragmentation strategy. Applying our more general rule as described above results in an EE-3B MUE of 0.58 kcal/mol and an EE-3B-CE MUE of 0.70 kcal/mol.

7.5.3 CONCLUSIONS

The EE-MB and the EE-MB-CE calculations were carried out on pentacoordinate and hexacoordinate Zn and Cd complexes, and the results were compared to full calculations at the same correlated level of theory. Both the EE-MB and the EE-MB-CE methods perform well. By using our new prescription for cluster fragmentation, the MUE for the EE-3B method for all 53 bond dissociation energies in all ten complexes is only 1.42 kcal/mol. Also, notably, the MUE for the EE-3B-CE method for all bond energies in all ten complexes was reduced to 0.59 kcal/mol. The average absolute error for the EE-3B-CE method is only 0.93% of the average absolute bond dissociation energy, which is

183

63.29 kcal/mol. These results show that our newly developed fragmentation strategy can be used for various Zn-containing systems representing other zinc-dependent enzyme active sites as well. The fact that EE-3B-CE MUE for Cd^{2+} system **7.11** is 0.03 kcal/mol supports our fragmentation strategy and shows the transferability of the findings to other metals such as Cd. The EE-3B-CE MUE of 0.30 kcal/mol obtained by CCSD(T) calculations on system **7.1** shows the transferability of the findings to other correlated methods. The new fragmentation prescription improves upon our previously published work by replacing four fragmentation guidelines with one simple and unambiguous rule, namely, in order to calculate accurate Zn-ligand or Cd-ligand bond dissociation energies with EE-MB and EE-MB-CE methods, one should find the two ligands that have the strongest Coulomb interaction with each other and combine them with Zn or Cd in one fragment.

The EE-3B-CE method is able to capture most of the correlation energy, requiring only a Hartree–Fock calculation for the full system and correlated calculations only for dimers and trimers of the fragments. It can also be used to obtain instantaneous bond dissociation energies, which are an important indicator of how well a method can capture the wide variety of energetic interactions that occur in a coordination complex, including both electrostatic interactions and electron-electron correlation. The fact that the EE-3B-CE approximation captures bond dissociation energies to within 0.93% of conventionally calculated values when the new fragmentation scheme is used demonstrates that it is possible to apply a relatively simple fragment-based method to obtain an accurate picture of the potential energy surface of a quite complicated system involving coordinate covalent bonds. One could therefore use the EE-MB-CE approximation to parameterize inorganic reactive force fields for use to study

184

macromolecular drug targets, or one could use it directly without force fields to obtain more accurate results that would otherwise be impossible for large systems (for example, systems with large ligands) where full calculations on the whole complex with a reliable method are unaffordable.

7.6 COMPARISON TO OTHER FRAGMENTATION METHODS

In this section we add more discussion of other fragment methods in order to place the EE-MB-CE method and the present work in perspective. Fragment methods may be classified in various ways,[212,235–238] each of which focuses on a different aspect. In discussing these classifications, the language we use is that the smallest subsystems considered (the groups of atoms that are together at all stages of the calculation) are the monomers, the fragments are any group of one or more monomers on which a calculation is carried out, and the extended system is called the entire system (other notation sometimes encountered in the literature is to call the monomers groups, to call the fragments monomers, and to call the entire system the supersystem).

One possible classification is into single-level methods and multi-level methods. In the former, all fragment calculations are carried out at the same level. In the latter, either fragment calculations are carried out with more than one level, or—as here— fragment calculations carried out at a higher level are combined with entire-system calculations carried out at a lower level.

Another possible classification is to distinguish fragment approximations from divide-and-conquer methods. In the latter, one uses fragments as an intermediate part of the calculation, but the goal is to converge to a particular unfragmented calculation. In such methods fragmentation is an algorithmic choice, not an approximation. In fragment approximations, the results converge to a high-level, entire-system result only in the limit

where one of the fragments treated at the high level is so large that it is the same as the entire system. The EE-MB-CE method is a fragment approximation. The rest of this discussion is concerned only with fragment approximations, i.e., with attempts to make useful calculations on larger systems feasible by introducing safe levels of approximation rather than by breaking the full calculation into smaller, more manageable parts without approximations.

Another possible classification is between inclusion–exclusion methods and many-body methods, but this is not unique since an inclusion−exclusion method can be thought of as an untruncated many-body method (with a distance cutoff), and many-body methods can also be thought of as an inclusion−exclusion based method.[236] A more unique classification is to ask whether a given part of the system occurs in one and only fragment (fragments do not overlap) or whether it occurs in more than one fragment (fragments do overlap); this is one way to classify fragment methods into overlapping and nonoverlapping methods (also sometimes called intersecting and non-intersecting). In this sense, many-body methods are overlapping methods because a given monomer occurs in many dimers, many trimers, and so forth. This contrasts with methods like the fragment molecular orbital (FMO) method[239] or the variational explicit polarization (X-Pol) method,[206] which have nonoverlapping fragments, except for a boundary atom in X-Pol. It is easily understood[238] that overlapping methods have faster convergence than nonoverlapping methods with respect to the size of fragments needed to obtain accurate results.

Another possible classification is whether a given fragment is embedded in the electrostatic field of the rest of the entire system. Including this embedding effect leads to more accurate results or to equally accurate results with smaller fragments.

186

In terms of the above classifications, the EE-MB-CE method is a multilevel, overlapping, embedded fragment approximation, and, since it dates to 2007,[220] it may be the first fragment approximation to actually combine all three of these advantages into a single algorithm.[210] Combining these three features makes it very powerful. We note though that the EE-MB method may be considered to be a special case of the electrostatically embedded[238] molecular tailoring[240] approach (EE-MTA), which allows more flexibility in the choice of fragments. We also note that the general energy-based fragment (GEBF) method[241] also shares some of the advantages of EE-MTA. Thus a multilevel version of EE-MTA or GEBF could be more flexible than EE-MB-CE, and a multilevel version of EE-MTA could include EE-MB-CE as a special case. One interesting possible way to make a multilevel extension of the GEBF method has been proposed recently by Li.[242]

Further classification is possible if one uses capping methods on bonds broken at the boundaries of fragments,[243] but we will not discuss that here since no caps are employed in the present work. In particular, even though we do break bonds in forming fragments, these are coordination bonds, and one of the goals of the present work is to show that we can obtain good results for coordination bonds without capping them.

One could also make classifications at a finer level; for example, we do not iterate the background charges to self-consistency (which makes the method simpler and less expensive[202]) whereas several other proposed methods do iterate the background electrostatics.

Discovering the most efficient way to carry out fragmentation calculations is certainly one of the most important challenges in current quantum chemical research, and the answer will almost surely depend on the problem. Further explanation of the accuracies that can be achieved on various kinds of problems by various fragmentation

approaches is therefore very important. The problem treated here is a very hard one. We have fragments that contain only one side of a metal–ligand bond. This shows the advantage of overlapping fragment methods, in that the bond can be split in this way, and yet represented intact in other fragments. Furthermore the electrostatics of the systems treated here are very challenging because the partial atomic charge on the metal atom depends on the number of ligands attached to it in the fragment, but we treat background electrostatics without iterations. The good results achieved here for this very difficult problem are encouraging.

Chapter 8

SYNTHESIS AND SAR OF HIGH-THROUGHPUT SCREENING (HTS) HITS

8.1 ACKNOWLEDGEMENTS

This chapter includes a description of work performed in collaboration with Jonathan Solberg and Jon E. Hawkinson at the Institute for Therapeutics Discovery and Development (ITDD), University of Minnesota. Jonathan Solberg tested the compounds *in vitro* and Jon Hawkinson analyzed the data. Synthetic guidance was provided by Rodney L. Johnson and Michael A. Walters.

8.2 INTRODUCTION

High-throughput screening (HTS) is an efficient drug discovery method that allows for rapid experimental evaluation of large chemical libraries to identify modulators of a validated drug target. In the past, this approach has been used by many pharmaceutical and biotechnology companies to fuel their drug discovery programs.[152] To discover novel small molecule inhibitors of LF, we have screened ~250,000 compounds for LF inhibition at the University of Minnesota ITDD. From this effort, we prioritized two hits, HTS Hits **8.16** and **8.23** (**Figure 8.1**), for further investigation. This chapter will describe the synthesis and biological evaluation of HTS Hits **8.16** and **8.23** and series of analogs of **8.16**.

Figure 8.1. Chemical Structures of HTS Hits **8.16** and **8.23**.
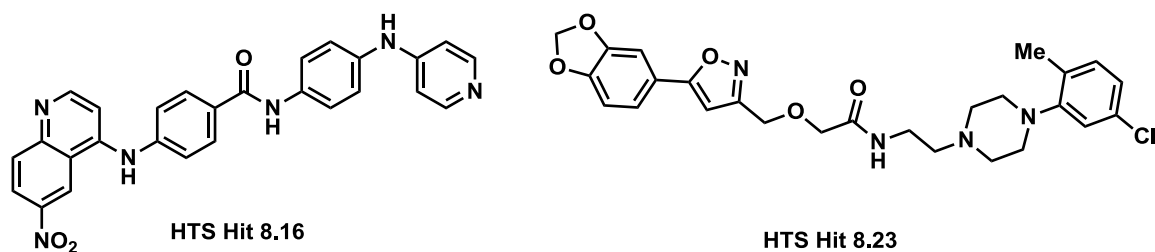
8.3 MATERIALS AND METHODS

8.3.1 Synthesis and Biochemical Evaluation

The synthesis of **8.16** and analogs was accomplished as outlined in **Scheme 8.1**. Compounds **8.5-8.8** were readily synthesized from commercially available anilines **8.1-8.2** and commercially available chloroquinolines **8.3-8.4** through nucleophilic substitution.

Scheme 8.1. Synthesis of compound **8.16** and its analogs.
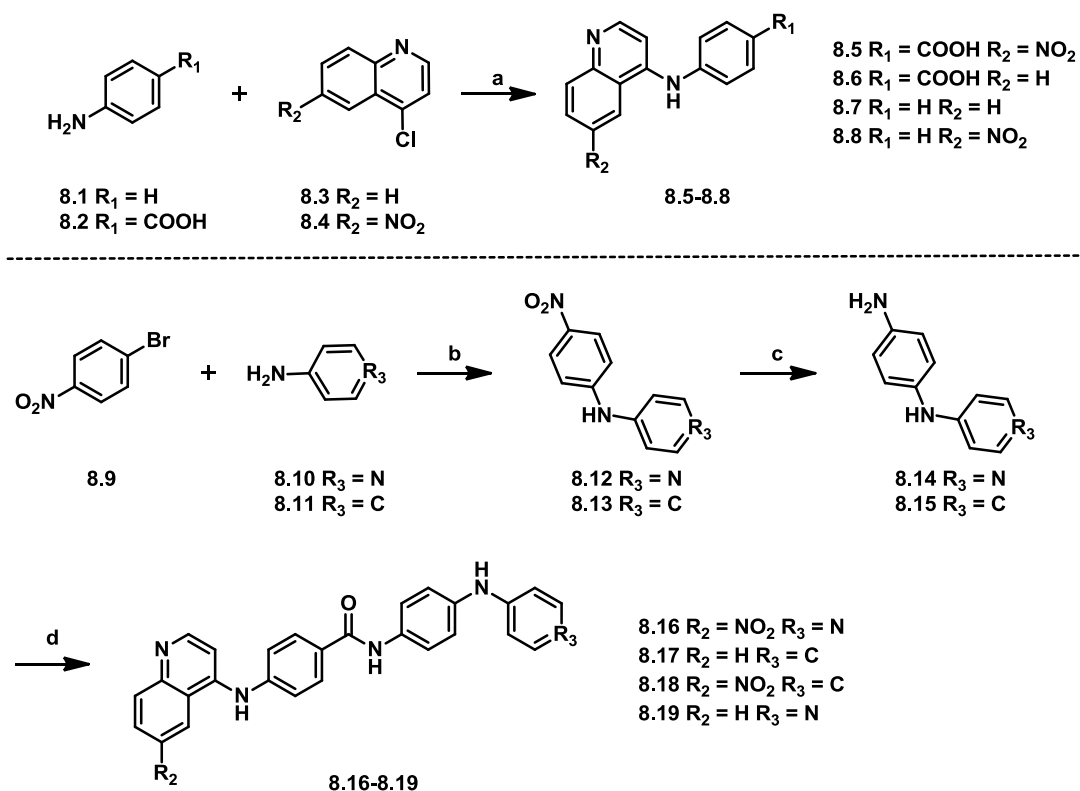
Reagents and Conditions: (a) HCl, Acetone/H_2O, reflux (**8.5**, 99%; **8.6**, 100%; **8.7**, 60%; **8.8**, 97%); (b) $Pd_2(dba)_3$, xantphos, Cs_2CO_3, 1,4-dioxane, 95 °C (**8.12**, 98%; **8.13**, 86%); (c) Fe dust, HCl, EtOH/H_2O, reflux (**8.14**, 73%; **8.15**, 71%); (d) 8.**5** or **8.6**, NMM, HOBt, EDC, DMF, rt (**8.16**, 12%; **8.17**, 82%; **8.18**, 81%; **8.19**, 64%).

Intermediates **8.12-8.13** were synthesized from commercially available starting materials **8.9-8.11** through Buchwald-Hartwig amination in 86-98% yield. Intermediates **8.12-8.13** were reduced to the amine using Fe dust to give **8.14-8.15** in 71-73% yield. Finally, carboxylic acids **8.5-8.6** were submitted to EDC coupling with amines **8.14-8.15** to yield **8.16-8.19** in 12-82% yield. Activities of the synthesized compounds were evaluated using the previously described *in vitro* LF FRET assay.

The synthesis of HTS Hit **8.23** was accomplished as outlined in **Scheme 8.2**. Hit **8.23** was synthesized in a single step from commercially available **8.21** and **8.22** using EDC coupling in 67% yield.

Scheme 8.2. Synthesis of compound **8.23**.
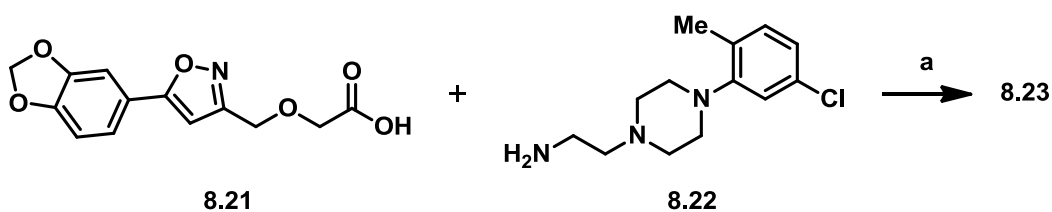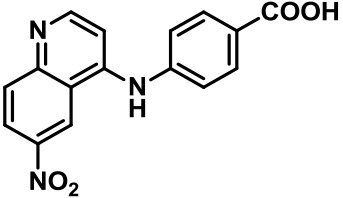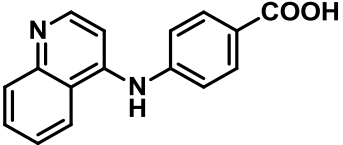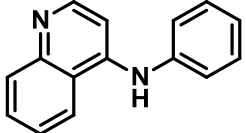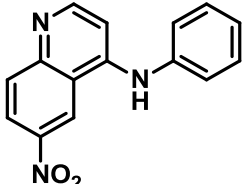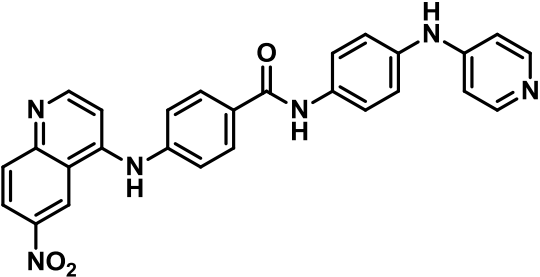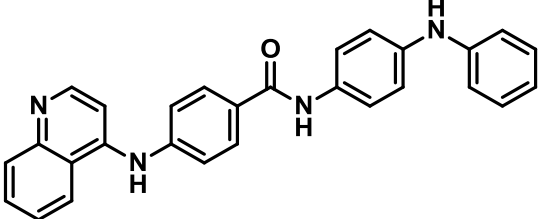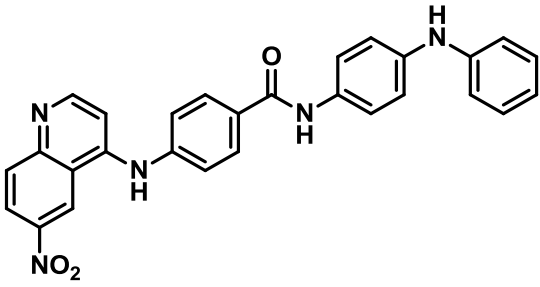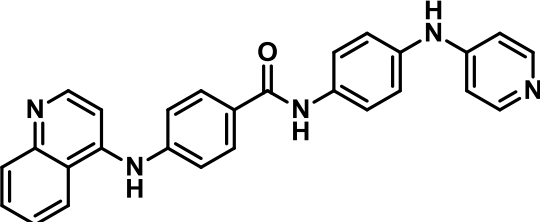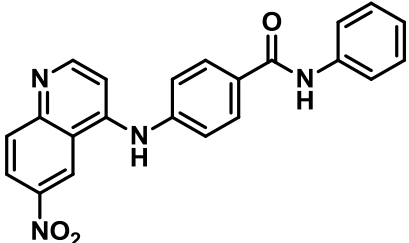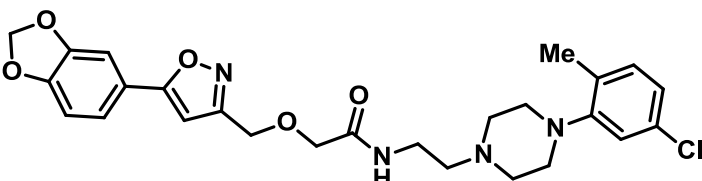
Reagents and Conditions: (a) NMM, HOBt, EDC, DMF, rt, 67%.

8.4 RESULTS AND DISCUSSION

8.4.1 Biochemical Evaluation

Table 8.1. LF FRET assay results for **8.23**, **8.16**, and fragments and analogs of **8.16**.

GPHR #[a]	Cpd #	Structure	LF IC$_{50}$ (μM)[b]
00225284	8.5		>100
00227097	8.6		>100
00227096	8.7		>100
00227095	8.8		>100
00220772-05	8.16		9.3
00278862	8.17		>100

GPHR #[a]		Structure	IC50[b]
00278859	**8.18**		>100
00278861	**8.19**		9.3
00278860	**8.20**		>100
00194983-04	**8.23**		>100

[a]GPHR # is a compound number in our in-house collection of compounds
[b]IC$_{50}$ is a half maximal inhibitory concentration

To our surprise, re-synthesized **8.23** showed no inhibitory activity against LF. To further investigate this observation, we re-purchased **8.23** from the original commercial supplier (ChemDiv) and purified it using column chromatography. Both purified and unpurified batches of **8.23** were evaluated using the FRET assay. The unpurified batch of **8.23** inhibited the enzyme, whereas the purified batch did not. These results indicate that compound **8.23** is likely a false positive as the commercial material contains a number of impurities, one or more of which may be responsible for LF inhibition. Additional investigations to confirm this hypothesis are ongoing.

The re-synthesized HTS Hit **8.16** inhibited the LF enzyme, confirming that it is a true hit. To further investigate the SAR of **8.16**, we have synthesized several fragments and analogs of **8.16** (**Table 8.1**). Fragments **8.5-8.8** and analogs **8.17**, **8.18**, and **8.20** do not show any improvement in activity over that of **8.16**. However, analog **8.19** shows similar activity to **8.16** with an IC_{50} of 9.3 μM. This means that the nitro functionality may not be important for the observed activity of **8.16,** whereas the pyridine ring is likely crucial for the observed activity**.**

8.5 CONCLUSION

In this chapter, we have synthesized HTS Hits **8.16** and **8.23**, and several fragments and analogs of **8.16**. We demonstrated that **8.23** is likely a false positive, whereas **8.16** is a true positive. For preliminary SAR, we synthesized several fragments and analogs of **8.16**. Fragments **8.5-8.8** and analogs **8.17**, **8.18**, and **8.20** did not show any improvement in activity over **8.16**. However, analog **8.19** showed similar LF inhibitory activity to **8.16** with an IC_{50} of 9.3 μM. These results suggest that the nitro functionality may not be important for the observed activity of **8.16,** whereas the pyridine ring is likely crucial for the observed activity of **8.16**.

8.6 EXPERIMENTAL

General Synthesis Information. Chemical reagents were purchased from commercial sources and used without additional purification. Bulk solvents were purchased from Fisher Scientific and anhydrous *N,N'*-dimethylformamide (DMF) was purchased from EMD Chemicals. Reactions were performed under an atmosphere of dry N_2 unless otherwise noted. Silica gel chromatography was performed on self-packed columns with SiliaFlash 60Å silica gel (SiliCycle). Compounds used in biological testing were no less than 95% pure as determined by two-wavelength HPLC analysis (254 and 215 nm). HPLC analyses were performed on an Agilent 1100 series instrument equipped with a diode array detector and a Zorbax SB-C18 column (0.5 x 150 mm, 5 µm, Agilent Technologies). LC-MS analyses were performed on an Agilent 1100 series instrument equipped with an Agilent MSD SL Ion Trap mass spectrometer (positive-ion mode) and a Zorbax SB-C18 column (0.5 x 150 mm, 5 µm, Agilent Technologies). The analysis method (15 µL/min flow rate) involved isocratic 10% MeCN (containing 0.1% TFA) in ddH_2O (containing 0.1% HCO_2H; 0 to 2 mins) followed by a linear gradient of 10% to 90% MeCN (containing 0.1% TFA) in ddH_2O (containing 0.1% HCO_2H; 2 to 24 mins), and isocratic 90% MeCN (containing 0.1% TFA) in ddH_2O (containing 0.1% HCO_2H; 24-26 mins). The column was heated to 40 ^{o}C. Wavelengths monitored = 254 nm and 215 nm. Nuclear magnetic resonance (NMR) spectra were recorded in $CDCl_3$, CD_3OD, or DMSO-d_6 on a Varian instrument operating at 400 MHz (for ^1H) and 100 MHz (for ^{13}C) at ambient temperature. Chemical shifts are reported in parts per million (ppm) and normalized to internal solvent peaks or tetramethylsilane (0 ppm).

4-((6-Nitroquinolin-4-yl)amino)benzoic acid (8.5)
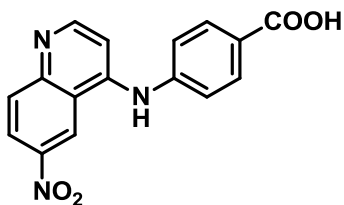
A mixture of **8.1** (0.1 g, 1.1 mmol), **8.3** (0.2 g, 1.0 mmol), and conc. HCl (cat.) in acetone/H_2O (15 mL, 1/2, v/v) was refluxed for 45 min and then cooled to rt. The crude reaction mixture was filtered to yield the desired product as a yellow solid (0.3 g, 99%). ^{1}H NMR (400 MHz, DMSO) δ 13.16, (s, 1H), 11.49 (s, 1H), 9.83 (s, 1H), 8.77-8.7 (m, 2H), 8.28 (d, J = 9.2 Hz, 1H), 8.15 (d, J = 8.8 Hz, 2H), 7.66 (d, J = 8.8 Hz, 2H), 7.16 (d, J = 6.4 Hz, 1H); ^{13}C NMR (100 MHz, DMSO) δ 167.0, 156.0, 145.5, 145.4, 141.8, 141.3, 131.5, 129.9, 127.9, 125.1, 122.7, 121.8, 117.5, 102.5.

4-(Quinolin-4-ylamino)benzoic acid (8.6)
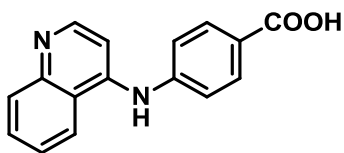
Synthesized according to the procedure reported for **8.5**. Yellow solid (82 mg, 100%). ^{1}H NMR (400 MHz, DMSO) δ 13.11, (s, 1H), 11.09 (s, 1H), 8.83 (d, J = 8.4 Hz, 1H), 8.59 (d, J = 6.8 Hz, 1H), 8.12-8.04 (m, 4H), 7.84 (t, J = 7.6 Hz, 1H), 7.64 (d, J = 8.8 Hz, 2H), 7.05 (d, J = 6.8 Hz, 1H); ^{13}C NMR (100 MHz, DMSO) δ 167.1, 154.8, 143.6, 142.0, 138.8, 134.5, 131.5, 129.4, 127.8, 125.0, 124.2, 120.9, 118.1, 101.1.

N-phenylquinolin-4-amine (8.7)
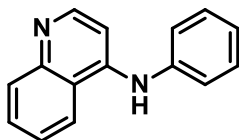
Synthesized according to the procedure reported for **8.5**. The crude material was purified over SiO$_2$ using an eluent of DCM/MeOH (10/1, v/v) to yield the desired product as a white solid (40 mg, 60%). ^{1}H NMR (400 MHz, DMSO) δ 9.0 (brs, 1H), 8.44 (d, J = 5.6 Hz, 1H), 8.38 (d, J = 8.4 Hz, 1H), 7.86 (d, J = 7.6 Hz, 1H), 7.68 (t, J = 7.2 Hz, 1H), 7.52 (t, J = 8.0 Hz, 1H), 7.43-7.35 (m, 4H), 7.13 (t, J = 7.2 Hz, 1H), 6.92 (d, J = 5.6 Hz, 1H); ^{13}C NMR (100 MHz, DMSO) δ 150.9, 149.2, 148.2, 140.9, 129.8, 129.8, 129.4, 125.1, 124.2, 122.8, 122.6, 120.2, 101.9.

6-Nitro-N-phenylquinolin-4-amine (8.8)
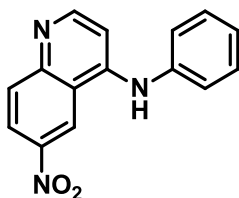
Synthesized according to the procedure reported for **8.5**.The crude material was purified over SiO$_2$ using an eluent of DCM/MeOH (20/1, v/v) to yield the desired product as a yellow solid (60 mg, 97%). ^{1}H NMR (400 MHz, DMSO) δ 9.62 (s, 1H), 9.48 (s, 1H), 8.55 (d, J = 5.2 Hz, 1H), 8.35 (d, J = 9.2 Hz, 1H), 7.97 (d, J = 9.6 Hz, 1H), 7.47-7.38 (m, 4H), 7.2 (t, J = 7.6 Hz, 1H), 6.96 (d, J = 5.6 Hz, 1H); ^{13}C NMR (100 MHz, DMSO) δ 154.6, 152.0, 150.3, 143.9, 134.0, 131.3, 130.0, 125.2, 123.6, 123.1, 120.7, 118.8, 102.8.

N-(4-Nitrophenyl)pyridin-4-amine (8.12)
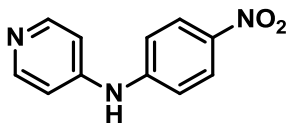
4-Aminopyridine (1.0 g, 10.9 mmol), 1-bromo-4-nitrobenzene (2.0 g, 9.9 mmol), Cs_2CO_3 (8.1 g, 24.8 mmol), xantphos (0.6 g, 1.0 mmol), and $Pd_2(dba)_3$ (0.5 g, 0.5 mmol) were added to a flame dried vial, and the vial was flushed with N_2 for 5 min. After, anhydrous 1,4-dioxane (40.0 mL) was added and the reaction was stirred at 95 °C for 40 h. Upon consumption of the starting material as determined by TLC, the reaction was concentrated under reduced pressure. The resulting solid residue was triturated with diethyl ether (5 × 20 mL), EtOAc (3 × 20 mL), and H_2O (3 × 20 mL) to yield an orange solid (2.1 g, 98%). ^{1}H NMR (400 MHz, DMSO) δ 9.62 (s, 1H), 8.35 (d, J = 5.6 Hz, 2H), 8.15 (d, J = 8.8 Hz, 2H), 7.29 (d, J = 9.2 Hz, 2H), 7.12 (d, J = 5.2 Hz, 2H); ^{13}C NMR (100 MHz, DMSO) δ 151.1, 148.3, 148.0, 140.6, 126.2, 117.1, 112.4.

4-Nitro-N-phenylaniline (8.13)
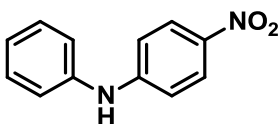
Aniline (0.30 mL, 3.29 mmol), 1-bromo-4-nitrobenzene (0.61 g, 3.00 mmol), Cs_2CO_3 (2.44 g, 7.50 mmol), xantphos (0.17 g, 0.30 mmol), and $Pd_2(dba)_3$ (0.14 g, 0.15 mmol) were added to a flame dried vial, and the vial was flushed with N_2 for 5 min. After, anhydrous 1,4-dioxane (15.00 mL) was added and the reaction was stirred at 95 °C for 40 h. Upon consumption of the starting material as determined by TLC, H_2O (10 mL) was added and the resulting layers were separated. The aq. layer was extracted with EtOAc (3 × 20 mL), and the combined organic layers were washed 1 M HCl (1 × 10 mL),

and brine (1 × 20 mL), dried over MgSO$_4$, filtered, and concentrated under reduced pressure. The crude product was purified over SiO$_2$ using an eluent of DCM/hexanes (1/2, v/v) to yield the product as a yellow solid (0.55 g, 86%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.09 (d, J = 8.8 Hz, 2H), 7.37 (t, J = 7.6 Hz, 1H), 7.20 (d, J = 8.0 Hz, 2H), 7.15 (t, J = 7.6 Hz, 2H), 6.93 (d, J = 9.6 Hz, 2H), 6.46 (s, 1H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 150.3, 139.6, 130.5, 129.7, 126.2, 124.6, 121.9, 113.7.

N-(Pyridin-4-yl)benzene-1,4-diamine hydrochloride (8.14)
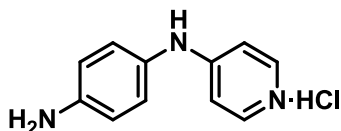
A solution of **8.12** (0.4 g, 2.0 mmol) in EtOH/H$_2$O (9.0 mL, 2/1, v/v) was heated to reflux and Fe dust (0.4 g, 7.9 mmol) and conc. HCl (0.2 mL) were added. The reaction was refluxed for 30 min, then cooled to rt, and filtered to remove solids. The filtrate was concentrated under reduced pressure to yield **8.14** (0.3 g, 73%), which was submitted to the next reaction without further purification.

N^1-Phenylbenzene-1,4-diamine (8.15)
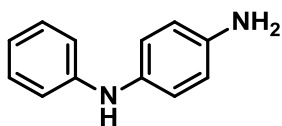
A solution of **8.13** (0.41 g, 1.90 mmol) in EtOH/H$_2$O (9 mL, 2/1, v/v) was heated to reflux and Fe dust (0.43 g, 7.63 mmol) and conc. HCl (0.20 mL) were added. The reaction was refluxed for 30 min, then cooled to rt, and filtered to remove solids. The filtrate was concentrated under reduced pressure, and the crude product was purified over SiO$_2$ using an eluent of DCM/MeOH (25/1, v/v) to yield the desired product as a dark red oil

(0.25 g, 71%).^{1}H NMR (400 MHz, CD$_3$OD) δ 7.09 (t, J = 8.8 Hz, 2H), 6.9 (d, J = 6 Hz, 2H), 6.84 (d, J = 8.0 Hz, 2H), 6.68 (t, J = 7.6 Hz, 1H), 6.4 (d, J = 6.4 Hz, 2H); ^{13}C NMR (100 MHz, CD$_3$OD) δ 146.3, 141.6, 134.8, 128.9, 122.1, 118.4, 116.6, 114.9.

4-((6-Nitroquinolin-4-yl)amino)-N-(4-(pyridin-4-ylamino)phenyl)benzamide (8.16)
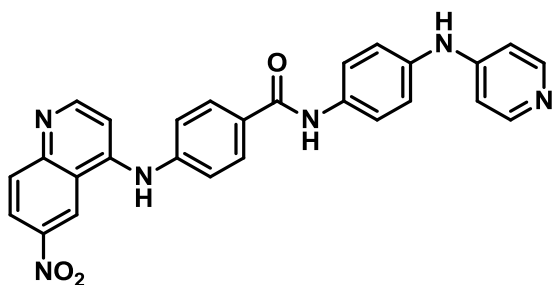
To a solution of **8.5** (0.06 g, 0.18 mmol) in DMF (2.00 mL) was added HOBt (0.03 g, 0.22 mmol), NMM (0.20 mL, 1.80 mmol), **8.14** (0.17 g, 0.90 mmol) and EDC (0.05 g, 0.25 mmol). The mixture was stirred at rt for 48 h. Upon consumption of the starting material as determined by TLC, H$_2$O (10 mL) was added. The resulting ppt was removed by filtration and purified by reverse phase HPLC on an Agilent 1200 series instrument equipped with a diode array detector and Zorbax SB-C18 column (21.2 x 250 mm, 7 μm, Agilent Technologies). The purification method (4.5 mL/min flow rate) involved isocratic 25% MeCN in ddH$_2$O (both containing 0.1% TFA; 0 to 2 mins) followed by a linear gradient of 25% to 95% MeCN in ddH$_2$O (both containing 0.1% TFA; 2 to 35 mins). The wavelengths monitored were 215 nm and 254 nm. The desired product was collected as a yellow solid (0.01 g, 12%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 9.72 (s, 1H), 8.78 (d, J = 9.2 Hz, 1H), 8.56 (d, J = 6.8 Hz, 1H), 8.2-8.14 (m, 5H), 7.9 (d, J = 8.8 Hz, 2H), 7.7 (d, J = 8.4 Hz, 2H), 7.37 (d, J = 8.0 Hz, 2H), 7.18 (d, J = 6.8 Hz, 1H), 7.11 (d, J = 7.2 Hz, 2H); ^{13}C NMR (100 MHz, CD$_3$OD) δ 166.1, 161.6, 157.6, 156.6, 145.8, 144.6, 141.6, 139.9,

137.5, 134.1, 133.0, 129.4, 127.4, 124.7, 124.2, 122.1, 122.0, 120.3, 117.3, 101.6. MS (ESI) 477.0 [M + H]$^+$.

N-(4-(Phenylamino)phenyl)-4-(quinolin-4-ylamino)benzamide (8.17)
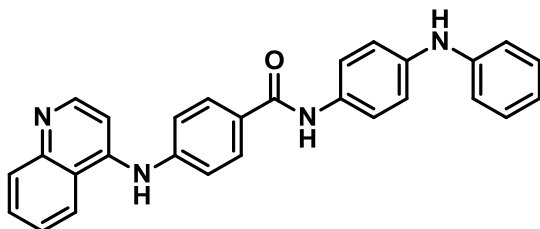
To a solution of **8.6** (0.03 g, 0.12 mmol) in DMF (2.00 mL) was added HOBt (0.02 g, 0.14 mmol), NMM (0.14 mL, 1.2 mmol), **8.15** (0.09 g, 0.46 mmol) and EDC (0.03 g, 0.17 mmol). The mixture was stirred at rt for 24 h. Upon consumption of the starting material as determined by TLC, H$_2$O (10 mL) was added. The aq. layer was extracted with EtOAc (3 × 20 mL), and the combined organic layers were washed with brine (1 × 20 mL) and concentrated under reduced pressure. The crude product was purified over SiO$_2$ using an eluent of DCM/MeOH (4/1, v/v) to yield the desired product as a yellow solid (0.04 g, 82%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.48 (d, *J* = 6 Hz, 1H), 8.42 (d, *J* = 8.8 Hz, 1H), 8.05 (d, *J* = 8.8 Hz, 2H), 7.95 (d, *J* = 8.4 Hz, 1H), 7.86 (t, *J* = 8.4 Hz, 1H), 7.68 (t, *J* = 8.4 Hz, 1H), 7.57-7.54 (m, 4H), 7.24-7.19 (m, 3H), 7.13-7.06 (m, 4H), 6.83 (t, *J* = 7.2 Hz, 1H); ^{13}C NMR (100 MHz, CD$_3$OD) δ 166.3, 150.8, 147.2, 144.8, 143.9, 142.7, 140.9, 131.3, 131.1, 130.8, 128.9, 128.7, 126.1, 125.0, 122.4, 122.0, 121.9, 119.6, 119.5, 117.3, 116.7, 101.9. MS (ESI) 431.1 [M + H]$^+$.

4-((6-Nitroquinolin-4-yl)amino)-N-(4-(phenylamino)phenyl)benzamide (8.18)
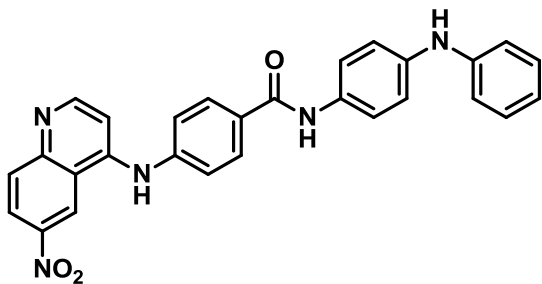
To a solution of **8.5** (0.05 g, 0.14 mmol) in DMF (2.00 mL) was added HOBt (0.02 g, 0.17 mmol), NMM (0.05 mL, 0.43 mmol), **8.15** (0.08 g, 0.43 mmol) and EDC (0.04 g, 0.20 mmol). The mixture was stirred at rt for 24 h. Upon consumption of the starting material as determined by TLC, H_2O (10 mL) was added. The aq. layer was extracted with EtOAc (3 × 20 mL), and the combined organic layers were washed with sat. aq. $NaHCO_3$ (1 × 20 mL) and brine (1 × 20 mL), dried over Na_2SO_4, filtered, and concentrated under reduced pressure to ~10% the original volume. The precipitate that formed was removed by filtration to yield the product as a yellow solid (56 mg, 81%). ^{1}H NMR (400 MHz, DMSO) δ 10.05 (s, 1H), 9.8 (s, 1H), 9.49 (s, 1H), 8.66 (s, 1H), 8.4 (d, J = 9.2 Hz, 1H), 8.03 (t, J = 8 Hz, 4H), 7.62 (d, J = 8.4 Hz, 2H), 7.50 (d, J = 8.0 Hz, 2H), 7.22-7.16 (m, 3H), 7.06-7.0 (m, 4H), 6.75 (t, J = 7.2 Hz, 1H); ^{13}C NMR (100 MHz, DMSO) δ 164.7, 154.4, 1151.9, 149.5, 144.5, 144.1, 139.6, 132.5, 131.1, 130.3, 129.6, 129.4, 123.3, 122.2, 122.1, 121.3, 121.1, 119.6, 119.4, 117.9, 116.3, 104.4. MS (ESI) 476.1 [M + H]$^+$.

N-(4-(Pyridin-4-ylamino)phenyl)-4-(quinolin-4-ylamino)benzamide (8.19)
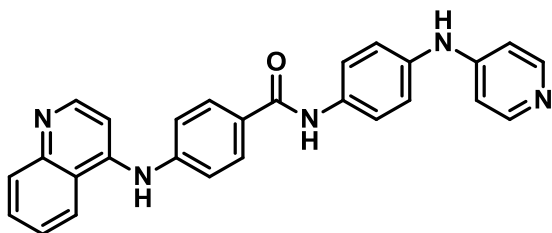
To a solution of **8.6** (0.04 g, 0.16 mmol) in DMF (2.00 mL) was added HOBt (0.03 g, 0.19 mmol), NMM (0.18 mL, 1.6 mmol), **8.14** (0.09 g, 0.49 mmol) and EDC (0.04 g, 0.22 mmol). The mixture was stirred at rt for 24 h. Upon consumption of the starting material as determined by TLC, H_2O (10 mL) was added. The precipitate that formed was removed by filtration to yield the product as a yellow solid (0.05 g, 64%). ^1H NMR (400 MHz, CD_3OD) δ 8.49 (d, J = 5.2 Hz, 1H), 8.30 (d, J = 7.6 Hz, 1H), 8.08 (d, J = 6.0 Hz, 2H), 8.0 (d, J = 9.2 Hz, 2H), 7.92 (d, J = 8.8 Hz, 1H), 7.76-7.69 (m, 3H), 7.57 (t, J = 8.0 Hz, 1H), 7.50 (d, J = 8.8 Hz, 2H), 7.26-7.22 (m, 3H), 6.89 (d, J = 6.4 Hz, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ 165.1, 152.3, 150.9, 149.2, 148.2, 147.0, 144.8, 135.7, 135.4, 130.0, 129.6, 129.5, 129.0, 125.5, 122.8, 122.1, 121.8, 120.8, 120.2, 109.1, 104.1. MS (ESI) 432.1 $[M + H]^+$.

4-((6-Nitroquinolin-4-yl)amino)-N-phenylbenzamide (8.20)
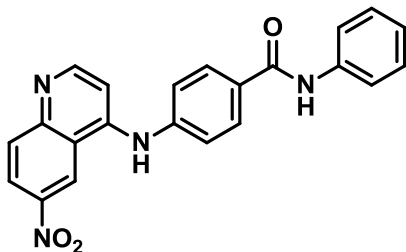
To a solution of **8.5** (0.05 g, 0.16 mmol) in DMF (2.00 mL) was added HOBt (0.03 g, 0.19 mmol), NMM (0.06 mL, 0.48 mmol), aniline (0.05 mL, 0.48 mmol) and EDC (0.04 g, 0.22 mmol). The mixture was stirred at rt for 24 h. Upon consumption of the starting material as determined by TLC, H_2O (10 mL) was added. The aq. layer was extracted with EtOAc (3 × 20 mL), and the combined organic layers were washed with sat. aq. $NaHCO_3$ (1 × 20 mL) and brine (1 × 20 mL), dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude product was purified over SiO_2 using an eluent of

DCM/MeOH (20/1, v/v) to yield the desired product as a yellow solid (0.03 g, 44%). ^{1}H NMR (400 MHz, DMSO) δ 10.17 (s, 1H), 9.82 (s, 1H), 9.49 (s, 1H), 8.67 (d, J = 5.2 Hz, 1H), 8.4 (d, J = 8.8 Hz, 1H), 8.06-8.02 (m, 3H), 7.76 (d, J = 7.6 Hz, 2H), 7.52 (d, J = 8.8 Hz, 2H), 7.33 (t, J = 8.0 Hz, 2H), 7.23 (d, J = 5.2 Hz, 1H), 7.07 (t, J = 7.6 Hz, 1H); ^{13}C NMR (100 MHz, DMSO) δ 165.2, 154.8, 152.0, 149.1, 144.2, 143.6, 139.7, 131.5, 130.3, 129.7, 129.0, 124.0, 123.3, 121.4, 120.8, 120.7, 119.4, 104.5.

2-((5-(benzo[d][1,3]dioxol-5-yl)isoxazol-3-yl)methoxy)-N-(2-(4-(5-chloro-2-methylphenyl)piperazin-1-yl)ethyl)acetamide (8.23)
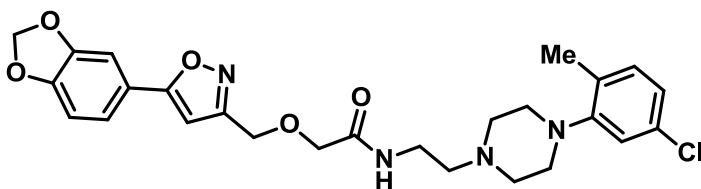
To a solution of **8.21** (0.10 g, 0.36 mmol) in DMF (2.00 mL) was added HOBt (0.06 g, 0.43 mmol), NMM (0.12 mL, 1.1 mmol), **8.22** (0.27 g, 1.10 mmol) and EDC (0.10 g, 0.50 mmol). The mixture was stirred at rt overnight. Upon consumption of the starting material as determined by TLC, H_2O (10 mL) was added. The aq. layer was extracted with EtOAc (3 × 20 mL), and the combined organic layers were washed with sat. aq. $NaHCO_3$ (1 × 20 mL) and brine (1 × 20 mL), dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude product was purified over SiO_2 using an eluent of DCM/MeOH (33/1, v/v) to yield the desired product as a white solid (0.12 mg, 67%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.21 (d, J = 8.0 Hz, 1H), 7.13 (s, 1H), 7.10-7.05 (m, 1H), 7.0 (d, J = 8.0 Hz, 1H), 6.88-6.83 (m, 2H), 6.76 (d, J = 7.6 Hz, 1H), 6.39 (s, 1H), 5.95 (s, 2H), 4.64 (s, 2H), 4.04 (s, 2H), 3.39 (q, J = 6.0 Hz, 2H), 2.82-2.8 (m, 4H), 2.56-2.51 (m, 6H), 2.17 (s, 3H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 170.4, 168.7, 160.9, 152.4, 149.4,

148.2, 131.9, 131.6, 130.7, 122.8, 121.1, 120.5, 119.3, 108.8, 106.0, 101.6, 97.5, 70.2, 64.8, 56.4, 53.2, 51.5, 35.4, 17.5. MS (ESI) 513.10 [M + H]$^+$.

REFERENCES

(1)   Christopher, G. W.; Cieslak, T. J.; Pavlin, J. A.; Eitzen, E. M. J. Biological Warfare. A Historical Perspective. *JAMA* **1997**, *278* (5), 412–417.

(2)   Cieslak, T. J.; Eitzen, E. M. Clinical and Epidemiologic Principles of Anthrax. *Emerg. Infect. Dis.* **1999**, *5* (4), 552–555.

(3)   Inglesby, T. V.; O'Toole, T.; Henderson, D. A.; Bartlett, J. G.; Ascher, M. S.; Eitzen, E.; Friedlander, A. M.; Gerberding, J.; Hauer, J.; Hughes, J.; McDade, J.; Osterholm, M. T.; Parker, G.; Perl, T. M.; Russell, P. K.; Tonat, K. Anthrax as a Biological Weapon, 2002: Updated Recommendations for Management. *JAMA* **2002**, *287* (17), 2236–2252.

(4)   Wilson, J. B.; Russell, K. E. Isolation of Bacillus Anthracis from Soil Stored 60 Years. *J. Bacteriol.* **1964**, *87* (1), 237–238.

(5)   http://emergency.cdc.gov/agent/agentlist-category.asp (accessed Sep 8, 2014).

(6)   Riedel, S. Anthrax: A Continuing Concern in the Era of Bioterrorism. *Proc. Bayl. Univ. Med. Cent.* **2005**, *18* (3), 234–243.

(7)   Meselson, M.; Guillemin, J.; Hugh-Jones, M.; Langmuir, A.; Popova, I.; Shelokov, A.; Yampolskaya, O. The Sverdlovsk Anthrax Outbreak of 1979. *Science* **1994**, *266* (5188), 1202–1208.

(8)   Harris, S. *Factories of Death: Japanese Secret Biological Warfare, 1932- 1945, and the American Cover-Up.*; London: Routledge, 1994.

(9)   Price, L. B.; Vogler, A.; Pearson, T.; Busch, J. D.; Schupp, J. M.; Keim, P. In Vitro Selection and Characterization of Bacillus Anthracis Mutants with High-Level Resistance to Ciprofloxacin. *Antimicrob. Agents Chemother.* **2003**, *47* (7), 2362–2365.

(10)   Athamna, A.; Athamna, M.; Abu-Rashed, N.; Medlej, B.; Bast, D. J.; Rubinstein, E. Selection of Bacillus Anthracis Isolates Resistant to Antibiotics. *J. Antimicrob. Chemother.* **2004**, *54* (2), 424–428.

(11)   Brook, I.; Elliott, T. B.; Pryor, H. I. 2nd; Sautter, T. E.; Gnade, B. T.; Thakar, J. H.; Knudson, G. B. In Vitro Resistance of Bacillus Anthracis Sterne to Doxycycline, Macrolides and Quinolones. *Int. J. Antimicrob. Agents* **2001**, *18* (6), 559–562.

(12)   Choe, C. H.; Bouhaouala, S. S.; Brook, I.; Elliot, T. B.; Knudson, G. B. In Vitro Development of Resistance to Ofloxacin and Doxycycline in Bacillus Anthracis Sterne. *Antimicrob. Agents Chemother.* **2000**, *44* (6).

(13)   Grunow, R.; Verbeek, L.; Jacob, D.; Holzmann, T.; Birkenfeld, G.; Wiens, D.; von Eichel-Streiber, L.; Grass, G.; Reischl, U. Injection Anthrax—a New Outbreak in Heroin Users. *Dtsch. Ärztebl. Int.* **2012**, *109* (49), 843–848.

(14)   Dixon, T. C.; Meselson, M.; Guillemin, J.; Hanna, P. C. Anthrax. *N. Engl. J. Med.* **1999**, *341* (11), 815–826.

(15)   Swartz, M. N. Recognition and Management of Anthrax — An Update. *N. Engl. J. Med.* **2001**, *345* (22), 1621–1626.

(16)   Spencer, R. C. Bacillus Anthracis. *J. Clin. Pathol.* **2003**, *56* (3), 182–187.

(17)   Aksaray, N.; Cinaz, P.; Coskun, U.; Serbest, M.; Koksal, F. Cutaneous Anthrax. *Trop. Geogr. Med.* **1990**, *42* (2), 168–171.

(18)   Holty, J.-E. C.; Bravata, D. M.; Liu, H.; Olshen, R. A.; McDonald, K. M.; Owens, D. K. Systematic Review: A Century of Inhalational Anthrax Cases from 1900 to 2005. *Ann. Intern. Med.* **2006**, *144* (4), 270–280.

(19)   Uchida, I.; Hashimoto, K.; Terakado, N. Virulence and Immunogenicity in Experimental Animals of Bacillus Anthracis Strains Harbouring or Lacking 110 MDa and 60 MDa Plasmids. *J. Gen. Microbiol.* **1986**, *132* (2), 557–559.

(20)   Turnbull, P. C.; Hutson, R. A.; Ward, M. J.; Jones, M. N.; Quinn, C. P.; Finnie, N. J.; Duggleby, C. J.; Kramer, J. M.; Melling, J. Bacillus Anthracis but Not Always Anthrax. *J. Appl. Bacteriol.* **1992**, *72* (1), 21–28.

(21)   Okinaka, R. T.; Cloud, K.; Hampton, O.; Hoffmaster, A. R.; Hill, K. K.; Keim, P.; Koehler, T. M.; Lamke, G.; Kumano, S.; Mahillon, J.; Manter, D.; Martinez, Y.; Ricke, D.; Svensson, R.; Jackson, P. J. Sequence and Organization of pXO1, the Large Bacillus Anthracis Plasmid Harboring the Anthrax Toxin Genes. *J. Bacteriol.* **1999**, *181* (20), 6509–6515.

(22)   Mikesell, P.; Ivins, B. E.; Ristroph, J. D.; Dreier, T. M. Evidence for Plasmid-Mediated Toxin Production in Bacillus Anthracis. *Infect. Immun.* **1983**, *39* (1), 371–376.

(23)   Green, B. D.; Battisti, L.; Koehler, T. M.; Thorne, C. B.; Ivins, B. E. Demonstration of a Capsule Plasmid in Bacillus Anthracis. *Infect. Immun.* **1985**, *49* (2), 291–297.

(24)   Drysdale, M.; Heninger, S.; Hutt, J.; Chen, Y.; Lyons, C. R.; Koehler, T. M. Capsule Synthesis by Bacillus Anthracis Is Required for Dissemination in Murine Inhalation Anthrax. *EMBO J.* **2005**, *24* (1), 221–227.

(25)   Mock, M.; Fouet, A. Anthrax. *Annu. Rev. Microbiol.* **2001**, *55*, 647–671.

(26)   Hambleton, P.; Carman, J. A.; Melling, J. Anthrax: The Disease in Relation to Vaccines. *Vaccine* **1984**, *2* (2), 125–132.

(27)   Cataldi, A.; Mock, M.; Bentancor, L. Characterization of Bacillus Anthracis
        Strains Used for Vaccination. *J. Appl. Microbiol.* **2000**, *88* (4), 648–654.

(28)   STERNE, M. Avirulent Anthrax Vaccine. *Onderstepoort J. Vet. Sci. Anim. Ind.*
        1946, *21*, 41–43.

(29)   Tigertt, W. D. Anthrax. William Smith Greenfield, M.D., F.R.C.P., Professor
        Superintendent, the Brown Animal Sanatory Institution (1878-81). Concerning
        the Priority due to Him for the Production of the First Vaccine against Anthrax. *J.
        Hyg. (Lond.)* **1980**, *85* (3), 415–420.

(30)   Wright, G. G.; Green, T. W.; Kanode, R. G. J. Studies on Immunity in Anthrax. V.
        Immunizing Activity of Alum-Precipitated Protective Antigen. *J. Immunol. Baltim.
        Md 1950* **1954**, *73* (6), 387–391.

(31)   Tresselt, H. B.; Boor, A. K. An Antigen Prepared in Vitro Effective for
        Immunization against Anthrax. III. Immunization of Monkeys against Anthrax. *J.
        Infect. Dis.* **1955**, *97* (2), 207–210.

(32)   PUZISS, M.; MANNING, L. C.; LYNCH, J. W.; BARCLAYE; ABELOW, I.;
        WRIGHT, G. G. Large-Scale Production of Protective Antigen of Bacillus
        Anthracis in Anaerobic Cultures. *Appl. Microbiol.* **1963**, *11*, 330–334.

(33)   http://biothrax.com/whatisbiothrax/howitworks.aspx (accessed Nov 24, 2014).

(34)   Brachman, P. S.; Gold, H.; Plotkin, S. A.; Fekety, F. R.; Werrin, M.; Ingraham, N.
        R. Field Evaluation of a Human Anthrax Vaccine. *Am. J. Public Health Nations
        Health* **1962**, *52* (4), 632–645.

(35)   BioThrax (anthrax Vaccine Adsorbed) [Vaccine Package Insert]. Rockville, MD:
        Emergent BioSolutions; 2008.

(36)   http://biothrax.com/forhealthpros/clinicalInfo/storage.aspx (accessed Nov 24,
        2014).

(37)   Wright, J. G.; Quinn, C. P.; Shadomy, S.; Messonnier, N. Use of Anthrax Vaccine
        in the United States: Recommendations of the Advisory Committee on
        Immunization Practices (ACIP), 2009. *MMWR Recomm. Rep. Morb. Mortal.
        Wkly. Rep. Recomm. Rep. Cent. Dis. Control* **2010**, *59* (RR-6).

(38)   Chames, P.; Van Regenmortel, M.; Weiss, E.; Baty, D. Therapeutic Antibodies:
        Successes, Limitations and Hopes for the Future. *Br. J. Pharmacol.* **2009**, *157*
        (2), 220–233.

(39)   Carter, P. J. Potent Antibody Therapeutics by Design. *Nat Rev Immunol* **2006**, *6*
        (5), 343–357.

(40)   Migone, T.-S.; Subramanian, G. M.; Zhong, J.; Healey, L. M.; Corey, A.;
        Devalaraja, M.; Lo, L.; Ullrich, S.; Zimmerman, J.; Chen, A.; Lewis, M.; Meister,

G.; Gillum, K.; Sanford, D.; Mott, J.; Bolmer, S. D. Raxibacumab for the Treatment of Inhalational Anthrax. *N. Engl. J. Med.* **2009**, *361* (2), 135–144.

(41)   Corey, A.; Migone, T.-S.; Bolmer, S.; Fiscella, M.; Ward, C.; Chen, C.; Meister, G. Bacillus Anthracis Protective Antigen Kinetics in Inhalation Spore-Challenged Untreated or Levofloxacin/ Raxibacumab-Treated New Zealand White Rabbits. *Toxins* **2013**, *5* (1), 120–138.

(42)   Anthrasil [package Insert]. Winnipeg, Canada: Cangene Corporation; 2015.

(43)   Sellman, B. R.; Mourez, M.; Collier, R. J. Dominant-Negative Mutants of a Toxin Subunit: An Approach to Therapy of Anthrax. *Science* **2001**, *292* (5517), 695–697.

(44)   Min, D.-H.; Tang, W.-J.; Mrksich, M. Chemical Screening by Mass Spectrometry to Identify Inhibitors of Anthrax Lethal Factor. *Nat. Biotechnol.* **2004**, *22* (6), 717–723.

(45)   Doganay, M.; Aydin, N. Antimicrobial Susceptibility of Bacillus Anthracis. *Scand. J. Infect. Dis.* **1991**, *23* (3), 333–335.

(46)   Odendaal, M. W.; Pieterson, P. M.; de Vos, V.; Botha, A. D. The Antibiotic Sensitivity Patterns of Bacillus Anthracis Isolated from the Kruger National Park. *Onderstepoort J. Vet. Res.* **1991**, *58* (1), 17–19.

(47)   Inglesby, T. V.; Henderson, D. A.; Bartlett, J. G.; Ascher, M. S.; Eitzen, E.; Friedlander, A. M.; Hauer, J.; McDade, J.; Osterholm, M. T.; O'Toole, T.; Parker, G.; Perl, T. M.; Russell, P. K.; Tonat, K. Anthrax as a Biological Weapon: Medical and Public Health Management. Working Group on Civilian Biodefense. *JAMA* **1999**, *281* (18), 1735–1745.

(48)   Jefferds, M. D.; Laserson, K.; Fry, A. M.; Roy, S.; Hayslett, J.; Grummer-Strawn, L.; Kettel-Khan, L.; Schuchat, A. Adherence to Antimicrobial Inhalational Anthrax Prophylaxis among Postal Workers, Washington, D.C., 2001. *Emerg. Infect. Dis.* **2002**, *8* (10), 1138–1144.

(49)   Pezard, C.; Berche, P.; Mock, M. Contribution of Individual Toxin Components to Virulence of Bacillus Anthracis. *Infect. Immun.* **1991**, *59* (10), 3472–3477.

(50)   Pannifer, A. D.; Wong, T. Y.; Schwarzenbacher, R.; Renatus, M.; Petosa, C.; Bienkowska, J.; Lacy, D. B.; Collier, R. J.; Park, S.; Leppla, S. H.; Hanna, P.; Liddington, R. C. Crystal Structure of the Anthrax Lethal Factor. *Nature* **2001**, *414* (6860), 229–233.

(51)   Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. Design and Therapeutic Application of Matrix Metalloproteinase Inhibitors. *Chem. Rev.* **1999**, *99* (9), 2735–2776.

(52)   Bradley, K. A.; Mogridge, J.; Mourez, M.; Collier, R. J.; Young, J. A. Identification of the Cellular Receptor for Anthrax Toxin. *Nature* **2001**, *414* (6860), 225–229.

(53)   Dalkas, G. A.; Papakyriakou, A.; Vlamis-Gardikas, A.; Spyroulias, G. A. Low Molecular Weight Inhibitors of the Protease Anthrax Lethal Factor. *Mini Rev. Med. Chem.* **2008**, *8* (3).

(54)   Petosa, C.; Collier, R. J.; Klimpel, K. R.; Leppla, S. H.; Liddington, R. C. Crystal Structure of the Anthrax Toxin Protective Antigen. *Nature* **1997**, *385* (6619), 833–838.

(55)   Arun P. Chopra; Sherrie A. Boone; Xudong Liang; Nicholas S. Duesbery. Anthrax Lethal Factor Proteolysis and Inactivation of MAPK Kinase. *J. Biol. Chem.* **2003**, *278* (11), 9402–9406.

(56)   Abrami, L.; Liu, S.; Cosson, P.; Leppla, S. H.; van der Goot, F. G. Anthrax Toxin Triggers Endocytosis of Its Receptor via a Lipid Raft-Mediated Clathrin-Dependent Process. *J. Cell Biol.* **2003**, *160* (3), 321–328.

(57)   Wei, W.; Lu, Q.; Chaudry, G. J.; Leppla, S. H.; Cohen, S. N. The LDL Receptor-Related Protein LRP6 Mediates Internalization and Lethality of Anthrax Toxin. *Cell* **2006**, *124* (6), 1141–1154.

(58)   Koehler, T. M.; Collier, R. J. Anthrax Toxin Protective Antigen: Low-pH-Induced Hydrophobicity and Channel Formation in Liposomes. *Mol. Microbiol.* **1991**, *5* (6), 1501–1506.

(59)   Miller, C. J.; Elliott, J. L.; Collier, R. J. Anthrax Protective Antigen: Prepore-to-Pore Conversion. *Biochemistry (Mosc.)* **1999**, *38* (32), 10432–10441.

(60)   Lacy, D. B.; Wigelsworth, D. J.; Melnyk, R. A.; Harrison, S. C.; Collier, R. J. Structure of Heptameric Protective Antigen Bound to an Anthrax Toxin Receptor: A Role for Receptor in pH-Dependent Pore Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (36), 13147–13151.

(61)   Qa'dan, M.; Christensen, K. A.; Zhang, L.; Roberts, T. M.; Collier, R. J. Membrane Insertion by Anthrax Protective Antigen in Cultured Cells. *Mol. Cell. Biol.* **2005**, *25* (13), 5492–5498.

(62)   Nassi, S.; Collier, R. J.; Finkelstein, A. PA63 Channel of Anthrax Toxin: An Extended Beta-Barrel. *Biochemistry (Mosc.)* **2002**, *41* (5), 1445–1450.

(63)   Vitale, G.; Bernardi, L.; Napolitani, G.; Mock, M.; Montecucco, C. Susceptibility of Mitogen-Activated Protein Kinase Kinase Family Members to Proteolysis by Anthrax Lethal Factor. *Biochem. J.* **2000**, *352 Pt 3*, 739–745.

(64)   Biondi, R. M.; Nebreda, A. R. Signalling Specificity of Ser/Thr Protein Kinases through Docking-Site-Mediated Interactions. *Biochem. J.* **2003**, *372* (1), 1–13.

211

(65)   Tanoue, T.; Nishida, E. Molecular Recognitions in the MAP Kinase Cascades. *Cell. Signal.* **2003**, *15* (5), 455–462.

(66)   Bardwell, A. J.; Abdollahi, M.; Bardwell, L. Anthrax Lethal Factor-Cleavage Products of MAPK (mitogen-Activated Protein Kinase) Kinases Exhibit Reduced Binding to Their Cognate MAPKs. *Biochem. J.* **2004**, *378* (2), 569–577.

(67)   Duesbery, N. S.; Webb, C. P.; Leppla, S. H.; Gordon, V. M.; Klimpel, K. R.; Copeland, T. D.; Ahn, N. G.; Oskarsson, M. K.; Fukasawa, K.; Paull, K. D.; Vande Woude, G. F. Proteolytic Inactivation of MAP-Kinase-Kinase by Anthrax Lethal Factor. *Science* **1998**, *280* (5364), 734–737.

(68)   Pellizzari, R.; Guidi-Rontani, C.; Vitale, G.; Mock, M.; Montecucco, C. Anthrax Lethal Factor Cleaves MKK3 in Macrophages and Inhibits the LPS/IFNγ-Induced Release of NO and TNFα. *FEBS Lett.* **1999**, *462* (1–2), 199–204.

(69)   Agrawal, A.; Lingappa, J.; Leppla, S. H.; Agrawal, S.; Jabbar, A.; Quinn, C.; Pulendran, B. Impairment of Dendritic Cells and Adaptive Immunity by Anthrax Lethal Toxin. *Nature* **2003**, *424* (6946), 329–334.

(70)   Crawford, M. A.; Aylott, C. V.; Bourdeau, R. W.; Bokoch, G. M. Bacillus Anthracis Toxins Inhibit Human Neutrophil NADPH Oxidase Activity. *J. Immunol.* **2006**, *176* (12), 7557–7565.

(71)   Comer, J. E.; Chopra, A. K.; Peterson, J. W.; König, R. Direct Inhibition of T-Lymphocyte Activation by Anthrax Toxins In Vivo. *Infect. Immun.* **2005**, *73* (12), 8275–8281.

(72)   Fang, H.; Xu, L.; Chen, T. Y.; Cyr, J. M.; Frucht, D. M. Anthrax Lethal Toxin Has Direct and Potent Inhibitory Effects on B Cell Proliferation and Immunoglobulin Production. *J. Immunol. Baltim. Md 1950* **2006**, *176* (10), 6155–6161.

(73)   Moayeri, M.; Leppla, S. H. The Roles of Anthrax Toxin in Pathogenesis. *Curr. Opin. Microbiol.* **2004**, *7* (1).

(74)   Warfel, J. M.; Steele, A. D.; D'Agnillo, F. Anthrax Lethal Toxin Induces Endothelial Barrier Dysfunction. *Am. J. Pathol.* **2005**, *166* (6), 1871–1881.

(75)   Liu, T.; Milia, E.; Warburton, R. R.; Hill, N. S.; Gaestel, M.; Kayyali, U. S. Anthrax Lethal Toxin Disrupts the Endothelial Permeability Barrier through Blocking p38 Signaling. *J. Cell. Physiol.* **2012**, *227* (4), 1438–1445.

(76)   Matthews, B. W. Structural Basis of the Action of Thermolysin and Related Zinc Peptidases. *Acc. Chem. Res.* **1988**, *21* (9), 333–340.

(77)   Tonello, F.; Naletto, L.; Romanello, V.; Molin, F. D.; Montecucco, C. Tyrosine-728 and Glutamic Acid-735 Are Essential for the Metalloproteolytic Activity of the Lethal Factor of Bacillus Anthracis. *Biochem. Biophys. Res. Commun.* **2004**, *313* (3), 496–502.

(78)   Bode, W.; Gomis-Ruth, F. X.; Stockler, W. Astacins, Serralysins, Snake Venom and Matrix Metalloproteinases Exhibit Identical Zinc-Binding Environments (HEXXHXXGXXH and Met-Turn) and Topologies and Should Be Grouped into a Common Family, the "Metzincins". *FEBS Lett.* **1993**, *331* (1-2), 134–140.

(79)   Marie-Claire, C.; Ruffet, E.; Tiraboschi, G.; Fournie-Zaluski, M. C. Differences in Transition State Stabilization between Thermolysin (EC 3.4.24.27) and Neprilysin (EC 3.4.24.11). *FEBS Lett.* **1998**, *438* (3), 215–219.

(80)   Xiong, Y.; Wiltsie, J.; Woods, A.; Guo, J.; Pivnichny, J. V.; Tang, W.; Bansal, A.; Cummings, R. T.; Cunningham, B. R.; Friedlander, A. M.; Douglas, C. M.; Salowe, S. P.; Zaller, D. M.; Scolnick, E. M.; Schmatz, D. M.; Bartizal, K.; Hermes, J. D.; MacCoss, M.; Chapman, K. T. The Discovery of a Potent and Selective Lethal Factor Inhibitor for Adjunct Therapy of Anthrax Infection. *Bioorg. Med. Chem. Lett.* **2006**, *16* (4), 964–968.

(81)   Forino, M.; Johnson, S.; Wong, T. Y.; Rozanov, D. V.; Savinov, A. Y.; Li, W.; Fattorusso, R.; Becattini, B.; Orry, A. J.; Jung, D.; Abagyan, R. A.; Smith, J. W.; Alibek, K.; Liddington, R. C.; Strongin, A. Y.; Pellecchia, M. Efficient Synthetic Inhibitors of Anthrax Lethal Factor. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (27), 9499–9504.

(82)   Johnson, S. L.; Jung, D.; Forino, M.; Chen, Y.; Satterthwait, A.; Rozanov, D. V.; Strongin, A. Y.; Pellecchia, M. Anthrax Lethal Factor Protease Inhibitors: Synthesis, SAR, and Structure-Based 3D QSAR Studies. *J. Med. Chem.* **2006**, *49* (1).

(83)   Goldman, E.; Cregar, L.; Nguyen, D.; Simo, O.; O'Malley, S.; Humphreys, T. Cationic Polyamines Inhibit Anthrax Lethal Factor Protease. *BMC Pharmacol.* **2006**, *6* (1), 1–8.

(84)   Schepetkin, I. A.; Khlebnikov, A. I.; Kirpotina, L. N.; Quinn, M. T. Novel Small-Molecule Inhibitors of Anthrax Lethal Factor Identified by High-Throughput Screening. *J. Med. Chem.* **2006**, *49* (17), 5232–5244.

(85)   Karginov, V. A.; Nestorovich, E. M.; Moayeri, M.; Leppla, S. H.; Bezrukov, S. M. Blocking Anthrax Lethal Toxin at the Protective Antigen Channel by Using Structure-Inspired Drug Design. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (42), 15075–15080.

(86)   Jiao, G.-S.; Kim, S.; Moayeri, M.; Crown, D.; Thai, A.; Cregar-Hernandez, L.; McKasson, L.; Sankaran, B.; Lehrer, A.; Wong, T.; Johns, L.; Margosiak, S. A.; Leppla, S. H.; Johnson, A. T. Antidotes to Anthrax Lethal Factor Intoxication. Part 3: Evaluation of Core Structures and Further Modifications to the C2-Side Chain. *Bioorg. Med. Chem. Lett.* **2012**, *22* (6), 2242–2246.

(87)   Verma, R. P.; Hansch, C. Combating the Threat of Anthrax: A Quantitative Structure-Activity Relationship Approach. *Mol. Pharm.* **2008**, *5* (5), 745–759.

(88)   Gaddis, B. D.; Avramova, L. V.; Chmielewski, J. Inhibitors of Anthrax Lethal Factor. *Bioorg. Med. Chem. Lett.* **2007**, *17* (16), 4575–4578.

(89)   Turk, B. E. Discovery and Development of Anthrax Lethal Factor Metalloproteinase Inhibitors. *Curr. Pharm. Biotechnol.* **2008**, *9* (1).

(90)   Tonello, F.; Seveso, M.; Marin, O.; Mock, M.; Montecucco, C. Pharmacology: Screening Inhibitors of Anthrax Lethal Factor. *Nature* **2002**, *418* (6896), 386–386.

(91)   Turk, B. E.; Wong, T. Y.; Schwarzenbacher, R.; Jarrell, E. T.; Leppla, S. H.; Collier, R. J.; Liddington, R. C.; Cantley, L. C. The Structural Basis for Substrate and Inhibitor Selectivity of the Anthrax Lethal Factor. *Nat. Struct. Mol. Biol.* **2004**, *11* (1), 60–66.

(92)   Jiao, G.-S; Kim, S.; Moayeri, M.; Cregar-Hernandez, L.; McKasson, L.; Margosiak, S. A.; Leppla, S. H.; Johnson, A. T. Antidotes to Anthrax Lethal Factor Intoxication. Part 1: Discovery of Potent Lethal Factor Inhibitors with in Vivo Efficacy. *Bioorg. Med. Chem. Lett.* **2010**, *20* (22), 6850–6853.

(93)   Moayeri, M.; Crown, D.; Jiao, G.-S.; Kim, S.; Johnson, A.; Leysath, C.; Leppla, S. H. Small-Molecule Inhibitors of Lethal Factor Protease Activity Protect against Anthrax Infection. *Antimicrob. Agents Chemother.* **2013**, *57* (9), 4139–4145.

(94)   Puerta, D. T.; Cohen, S. M. A Bioinorganic Perspective on Matrix Metalloproteinase Inhibition. *Curr. Top. Med. Chem.* **2004**, *4* (15), 1551–1573.

(95)   Summers, J. B.; Gunn, B. P.; Mazdiyasni, H.; Goetze, A. M.; Young, P. R.; Bouska, J. B.; Dyer, R. D.; Brooks, D. W.; Carter, G. W. In Vivo Characterization of Hydroxamic Acid Inhibitors of 5-Lipoxygenase. *J. Med. Chem.* **1987**, *30* (11), 2121–2126.

(96)   Tegoni, M.; Dallavalle, F.; Amélia Santos, M. Succinylhydroxamic Derivatives of A-Amino Acids as MMP Inhibitors. Study of Complex-Formation Equilibria with Cu2+, Ni2+ and Zn2+. *J. Inorg. Biochem.* **2004**, *98* (2), 209–218.

(97)   Konstantinopoulos, P. A.; Karamouzis, M. V.; Papatsoris, A. G.; Papavassiliou, A. G. Matrix Metalloproteinase Inhibitors as Anticancer Agents. *Int. J. Biochem. Cell Biol.* **2008**, *40* (6-7), 1156–1168.

(98)   Levitt, N. C.; Eskens, F. A.; O'Byrne, K. J.; Propper, D. J.; Denis, L. J.; Owen, S. J.; Choi, L.; Foekens, J. A.; Wilner, S.; Wood, J. M.; Nakajima, M.; Talbot, D. C.; Steward, W. P.; Harris, A. L.; Verweij, J. Phase I and Pharmacological Study of the Oral Matrix Metalloproteinase Inhibitor, MMI270 (CGS27023A), in Patients with Advanced Solid Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2001**, *7* (7), 1912–1922.

(99)   Farkas, E.; Enyedy, É. A.; Csóka, H. A Comparison between the Chelating Properties of Some Dihydroxamic Acids, Desferrioxamine B and Acetohydroxamic Acid. *Polyhedron* **1999**, *18* (18), 2391–2398.

(100)   Panchal, R. G.; Hermone, A. R.; Nguyen, T. L.; Wong, T. Y.; Schwarzenbacher, R.; Schmidt, J.; Lane, D.; McGrath, C.; Turk, B. E.; Burnett, J.; Aman, M. J.; Little, S.; Sausville, E. A.; Zaharevitz, D. W.; Cantley, L. C.; Liddington, R. C.; Gussio, R.; Bavari, S. Identification of Small Molecule Inhibitors of Anthrax Lethal Factor. *Nat. Struct. Mol. Biol.* **2004**, *11* (1).

(101)   Lee, L. V.; Bower, K. E.; Liang, F.-S.; Shi, J.; Wu, D.; Sucheck, S. J.; Vogt, P. K.; Wong, C.-H. Inhibition of the Proteolytic Activity of Anthrax Lethal Factor by Aminoglycosides. *J. Am. Chem. Soc.* **2004**, *126* (15), 4774–4775.

(102)   Kocer, S. S.; Walker, S. G.; Zerler, B.; Golub, L. M.; Simon, S. R. Metalloproteinase Inhibitors, Nonantimicrobial Chemically Modified Tetracyclines, and Ilomastat Block Bacillus Anthracis Lethal Factor Activity in Viable Cells. *Infect. Immun.* **2005**, *73* (11), 7548–7557.

(103)   Johnson, S. L.; Chen, L.-H.; Harbach, R.; Sabet, M.; Savinov, A.; Cotton, N. J. H.; Strongin, A.; Guiney, D.; Pellecchia, M. Rhodanine Derivatives as Selective Protease Inhibitors against Bacterial Toxins. *Chem. Biol. Drug Des.* **2008**, *71* (2), 131–139.

(104)   Houseman, B. T.; Huh, J. H.; Kron, S. J.; Mrksich, M. Peptide Chips for the Quantitative Evaluation of Protein Kinase Activity. *Nat. Biotechnol.* **2002**, *20* (3), 270–274.

(105)   Min, D.-H.; Su, J.; Mrksich, M. Profiling Kinase Activities by Using a Peptide Chip and Mass Spectrometry. *Angew. Chem. Int. Ed.* **2004**, *43* (44), 5973–5977.

(106)   Su, J.; Mrksich, M. Using Mass Spectrometry to Characterize Self-Assembled Monolayers Presenting Peptides, Proteins, and Carbohydrates. *Angew. Chem. Int. Ed.* **2002**, *41* (24), 4715–4718.

(107)   Dell'Aica, I. D., Massimo Tonello, Fiorella Piris, Alejandro Mock, Michèle Montecucco, Cesare Garbisa, Spiridione. Potent Inhibitors of Anthrax Lethal Factor from Green Tea. *EMBO Rep.* **2004**, *5* (4), 418–422.

(108)   Kim, C.; Gajendran, N.; Mittrucker, H.-W.; Weiwad, M.; Song, Y.-H.; Hurwitz, R.; Wilmanns, M.; Fischer, G.; Kaufmann, S. H. E. Human Alpha-Defensins Neutralize Anthrax Lethal Toxin and Protect against Its Fatal Consequences. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (13), 4830–4835.

(109)   Johnson, S. L.; Chen, L.-H.; Pellecchia, M. A High-Throughput Screening Approach to Anthrax Lethal Factor Inhibition. *Bioorganic Chem.* **2007**, *35* (4), 306–312.

(110) Shoop, W. L.; Xiong, Y.; Wiltsie, J.; Woods, A.; Guo, J.; Pivnichny, J. V.; Felcetto, T.; Michael, B. F.; Bansal, A.; Cummings, R. T.; Cunningham, B. R.; Friedlander, A. M.; Douglas, C. M.; Patel, S. B.; Wisniewski, D.; Scapin, G.; Salowe, S. P.; Zaller, D. M.; Chapman, K. T.; Scolnick, E. M.; Schmatz, D. M.; Bartizal, K.; MacCoss, M.; Hermes, J. D. Anthrax Lethal Factor Inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (22), 7958–7963.

(111) Chiu, T.-L.; Solberg, J.; Patil, S.; Geders, T. W.; Zhang, X.; Rangarajan, S.; Francis, R.; Finzel, B. C.; Walters, M. A.; Hook, D. J.; Amin, E. A. Identification of Novel Non-Hydroxamate Anthrax Toxin Lethal Factor Inhibitors by Topomeric Searching, Docking and Scoring, and in Vitro Screening. *J. Chem. Inf. Model.* **2009**, *49* (12), 2726–2734.

(112) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53* (7), 2719–2740.

(113) Tomasic, T.; Peterlin Masic, L. Rhodanine as a Scaffold in Drug Discovery: A Critical Review of Its Biological Activities and Mechanisms of Target Modulation. *Expert Opin. Drug Discov.* **2012**, *7* (7), 549–560.

(114) Glide, Version 5.9, Schrödinger, LLC, New York, NY, 2014.

(115) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749.

(116) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196.

(117) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750–1759.

(118) Maestro, Version 9.4, Schrödinger, LLC, New York, NY, 2013.

(119) http://www.sigmaaldrich.com/chemistry/chemistry-services/selected-structure.html (accessed Sep 8, 2014).

(120) LigPrep, Version 2.6, Schrödinger, LLC, New York, NY, 2013.

(121) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236.

(122) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput. Aided Mol. Des.* **2013**, *27* (3), 221–234.

(123) Schrödinger Suite 2013 Protein Preparation Wizard; Epik Version 2.4, Schrödinger, LLC, New York, NY, 2013; Impact Version 5.9, Schrödinger, LLC, New York, NY, 2013; Prime Version 3.2, Schrödinger, LLC, New York, NY, 2013.

(124) Prime, Version 3.2, Schrödinger, LLC, New York, NY, 2013.

(125) Maize, K. M.; Kurbanov, E. K.; De La Mora-Rey, T.; Geders, T. W.; Hwang, D.-J.; Walters, M. A.; Johnson, R. L.; Amin, E. A.; Finzel, B. C. Anthrax Toxin Lethal Factor Domain 3 Is Highly Mobile and Responsive to Ligand Binding. *Acta Crystallogr. Sect. D* **2014**, *70* (11), 2813–2822.

(126) Stols, L.; Zhou, M.; Eschenfeldt, W. H.; Millard, C. S.; Abdullah, J.; Collart, F. R.; Kim, Y.; Donnelly, M. I. New Vectors for Co-Expression of Proteins: Structure of Bacillus Subtilis ScoAB Obtained by High-Throughput Protocols. *Protein Expr. Purif.* **2007**, *53* (2), 396–403.

(127) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *High Throughput Protein Expression and Purification*; Methods in Molecular Biology; 2009; Vol. 498, pp 105–115.

(128) Cormier, C. Y.; Mohr, S. E.; Zuo, D.; Hu, Y.; Rolfs, A.; Kramer, J.; Taycher, E.; Kelley, F.; Fiacco, M.; Turnbull, G.; LaBaer, J. Protein Structure Initiative Material Repository: An Open Shared Public Resource of Structural Genomics Plasmids for the Biological Community. *Nucleic Acids Res.* **2010**, *38* (Database), D743–D749.

(129) Seiler, C. Y.; Park, J. G.; Sharma, A.; Hunter, P.; Surapaneni, P.; Sedillo, C.; Field, J.; Algar, R.; Price, A.; Steel, J.; Throop, A.; Fiacco, M.; LaBaer, J. DNASU Plasmid and PSI:Biology-Materials Repositories: Resources to Accelerate Biological Research. *Nucleic Acids Res.* **2014**, *42* (D1), D1253–D1260.

(130) Cormier, C. Y.; Park, J. G.; Fiacco, M.; Steel, J.; Hunter, P.; Kramer, J.; Singla, R.; LaBaer, J. PSI:Biology-Materials Repository: A Biologist's Resource for Protein Expression Plasmids. *J. Struct. Funct. Genomics* **2011**, *12* (2), 55–62.

(131) Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (2), 125–132.

(132) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62* (1), 72–82.

(133) Pflugrath, J. The Finer Things in X-Ray Diffraction Data Collection. *Acta Crystallogr. Sect. D* **1999**, *55* (10), 1718–1725.

(134) Shoop, W. L.; Xiong, Y.; Wiltsie, J.; Woods, A.; Guo, J.; Pivnichny, J. V.; Felcetto, T.; Michael, B. F.; Bansal, A.; Cummings, R. T.; Cunningham, B. R.; Friedlander, A. M.; Douglas, C. M.; Patel, S. B.; Wisniewski, D.; Scapin, G.; Salowe, S. P.; Zaller, D. M.; Chapman, K. T.; Scolnick, E. M.; Schmatz, D. M.; Bartizal, K.; MacCoss, M.; Hermes, J. D. Anthrax Lethal Factor Inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (22), 7958–7963.

(135) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674.

(136) 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. W.; 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. Sect. D* **2011**, *67* (4), 235–242.

(137) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D* **2011**, *67* (4), 355–367.

(138) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D* **2010**, *66* (2), 213–221.

(139) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.refine. *Acta Crystallogr. Sect. D* **2012**, *68* (4), 352–367.

(140) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D* **2004**, *60* (12 Part 1), 2126–2132.

(141) Finzel, B. C.; Akavaram, R.; Ragipindi, A.; Van Voorst, J. R.; Cahn, M.; Davis, M. E.; Pokross, M. E.; Sheriff, S.; Baldwin, E. T. Conserved Core Substructures in the Overlay of Protein–Ligand Complexes. *J. Chem. Inf. Model.* **2011**, *51* (8), 1931–1941.

(142) Otwinowski, Z.; Minor, W. [20] Processing of X-Ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*; Elsevier, 1997; Vol. 276, pp 307–326.

(143) Desmond Molecular Dynamics System, Version 3.7, D. E. Shaw Research, New York, NY, 2014. Maestro-Desmond Interoperability Tools, Version 3.7, Schrödinger, New York, NY, 2014.

(144) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies Using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6* (5), 1509–1519.

(145) Guo, Z.; Mohanty, U.; Noehre, J.; Sawyer, T. K.; Sherman, W.; Krilov, G. Probing the Alpha-Helical Structural Stability of Stapled p53 Peptides: Molecular Dynamics Simulations and Analysis. *Chem. Biol. Drug Des.* **2010**, *75* (4), 348–359.

(146) Bowers, K. J.; Chow, E.; Huageng Xu; Dror, R. O.; Eastwood, M. P.; Gregersen, B. .; Klepeis, J. L.; Kolossvary, I.; Moraes, M. .; Sacerdoti, F. D.; Salmon, J. K.; Yibing Shan; Shaw, D. E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *SC 2006 Conference, Proceedings of the ACM/IEEE*; 2006; pp 43–43.

(147) Maestro, Version 9.7, Schrödinger, LLC, New York, NY, 2014.

(148) Jain, A. N. Surflex-Dock 2.1: Robust Performance from Ligand Energetic Modeling, Ring Flexibility, and Knowledge-Based Search. *J. Comput. Aided Mol. Des.* **2007**, *21* (5).

(149) Jain, A. N. Surflex: Fully Automatic Flexible Molecular Docking Using a Molecular Similarity-Based Search Engine. *J. Med. Chem.* **2003**, *46* (4).

(150) 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* (16), 2785–2791.

(151) Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015.

(152) Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S. Impact of High-Throughput Screening in Biomedical Research. *Nat Rev Drug Discov* **2011**, *10* (3), 188–195.

(153) Perola, E. An Analysis of the Binding Efficiencies of Drugs and Their Leads in Successful Drug Discovery Programs. *J. Med. Chem.* **2010**, *53* (7), 2986–2997.

(154) Lahana, R. How Many Leads from HTS? *Drug Discov. Today* **1999**, *4* (10), 447–448.

(155) Malo, N.; Hanley, J. A.; Cerquozzi, S.; Pelletier, J.; Nadon, R. Statistical Practice in High-Throughput Screening Data Analysis. *Nat Biotech* **2006**, *24* (2), 167–175.

(156) McInnes, C. Virtual Screening Strategies in Drug Discovery. *Anal. Tech. Mech.* **2007**, *11* (5), 494–502.

(157) Klebe, G. Virtual Ligand Screening: Strategies, Perspectives and Limitations. *Drug Discov. Today* **2006**, *11* (13–14), 580–594.

(158) Kubinyi, H. Success Stories of Computer-Aided Design. In *Computer Applications in Pharmaceutical Research and Development*; John Wiley & Sons, Inc., 2006; pp 377–424.

(159) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757–1768.

(160) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **2001**, *46* (1-3).

(161) http://zinc.docking.org/browse/subsets/, Accessed 03/03/2015.

(162) Kim, S.; Jiao, G.-S.; Moayeri, M.; Crown, D.; Cregar-Hernandez, L.; McKasson, L.; Margosiak, S. A.; Leppla, S. H.; Johnson, A. T. Antidotes to Anthrax Lethal Factor Intoxication. Part 2: Structural Modifications Leading to Improved in Vivo Efficacy. *Bioorg. Med. Chem. Lett.* **2011**, *21* (7), 2030–2033.

(163) Pearlman, D. A.; Charifson, P. S. Improved Scoring of Ligand-Protein Interactions Using OWFEG Free Energy Grids. *J. Med. Chem.* **2001**, *44* (4), 502–511.

(164) Jacobsen, J. A.; Fullagar, J. L.; Miller, M. T.; Cohen, S. M. Identifying Chelators for Metalloprotein Inhibitors Using a Fragment-Based Approach. *J. Med. Chem.* **2011**, *54* (2).

(165) www.emolecules.com (accessed May 15, 2015).

(166) http://www.rcsb.org/pdb/home/home.do (accessed Jul 15, 2015).

(167) Finzel, B. C.; Akavaram, R.; Ragipindi, A.; Van Voorst, J. R.; Cahn, M.; Davis, M. E.; Pokross, M. E.; Sheriff, S.; Baldwin, E. T. Conserved Core Substructures in the Overlay of Protein-Ligand Complexes. *J. Chem. Inf. Model.* **2011**, *51* (8), 1931–1941.

(168) Schrödinger Suite 2011 Schrödinger Suite; Epik Version 2.2, Schrödinger, LLC, New York, NY, 2011; Impact Version 5.7, Schrödinger, LLC, New York, NY, 2011; Prime Version 2.3, Schrödinger, LLC, New York, NY, 2011.

(169) Small-Molecule Drug Discovery Suite 2014-2: QikProp, Version 4.0, Schrödinger, LLC, New York, NY, 2014.

(170) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45* (12), 2615–2623.

(171) Van der Vaart, A.; Merz, K. M. The Role of Polarization and Charge Transfer in the Solvation of Biomolecules. *J. Am. Chem. Soc.* **1999**, *121* (39), 9182–9190.

(172) Sakharov, D. V.; Lim, C. Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **2005**, *127* (13), 4921–4929.

(173) Khandelwal, A.; Lukacova, V.; Comez, D.; Kroll, D. M.; Raha, S.; Balaz, S. A Combination of Docking, QM/MM Methods, and MD Simulation for Binding Affinity Estimation of Metalloprotein Ligands. *J. Med. Chem.* **2005**, *48* (17), 5437–5447.

(174) http://www.schrodinger.com/kb/793 (accessed May 14, 2015).

(175) Sandstead, H. H. Understanding Zinc: Recent Observations and Interpretations. *J. Lab. Clin. Med.* **1994**, *124* (3), 322–327.

(176) Prasad, A. S. Zinc: An Overview. *Nutr. Burbank Los Angel. Cty. Calif* **1995**, *11* (1 Suppl), 93–99.

(177) Solomons, N. W. Mild Human Zinc Deficiency Produces an Imbalance between Cell-Mediated and Humoral Immunity. *Nutr. Rev.* **1998**, *56* (1 Pt 1), 27–28.

(178) MacDonald, R. S. The Role of Zinc in Growth and Cell Proliferation. *J. Nutr.* **2000**, *130* (5S Suppl).

(179) Heyneman, C. A. Zinc Deficiency and Taste Disorders. *Ann. Pharmacother.* **1996**, *30* (2), 186–187.

(180) Henkin, R. I. Editorial: Zinc in Wound Healing. *N. Engl. J. Med.* **1974**, *291* (13), 675–676.

(181) Wallwork, J. C.; Duerre, J. A. Effect of Zinc Deficiency on Methionine Metabolism, Methylation Reactions and Protein Synthesis in Isolated Perfused Rat Liver. *J. Nutr.* **1985**, *115* (2), 252–262.

(182) Amin, E. A.; Truhlar, D. G. Zn Coordination Chemistry: Development of Benchmark Suites for Geometries, Dipole Moments, and Bond Dissociation Energies and Their Use To Test and Validate Density Functionals and Molecular Orbital Theory. *J. Chem. Theory Comput.* **2008**, *4* (1), 75–85.

(183) Cramer, C. J.; Truhlar, D. G. Density Functional Theory for Transition Metals and Transition Metal Chemistry. *Phys. Chem. Chem. Phys. PCCP* **2009**, *11* (46), 10757–10816.

(184) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46* (7), 618–622.

(185)   Čížek, J. On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules. In *Advances in Chemical Physics*; John Wiley & Sons, Inc., 1969; pp 35–89.

(186)   Purvis, G. D.; Bartlett, R. J. A Full Coupled-cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76* (4), 1910–1918.

(187)   Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, *157* (6), 479–483.

(188)   Raghavachari, K.; Anderson, J. B. Electron Correlation Effects in Molecules. *J. Phys. Chem.* **1996**, *100* (31), 12960–12973.

(189)   Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. Pseudospectral Localized Mo/ller–Plesset Methods: Theory and Calculation of Conformational Energies. *J. Chem. Phys.* **1995**, *103* (4), 1481–1490.

(190)   Nakao, Y.; Hirao, K. A Local Second-Order Moller-Plesset Method with Localized Orbitals: A Parallelized Efficient Electron Correlation Method. *J. Chem. Phys.* **2004**, *120* (14), 6375–6380.

(191)   Saebo/, S.; Pulay, P. Fourth-order Mo/ller–Plessett Perturbation Theory in the Local Correlation Treatment. I. Method. *J. Chem. Phys.* **1987**, *86* (2), 914–922.

(192)   Galli; Parrinello. Large Scale Electronic Structure Calculations. *Phys. Rev. Lett.* **1992**, *69* (24).

(193)   Nielsen, I. M. B.; Janssen, C. L. Local Møller−Plesset Perturbation Theory: A Massively Parallel Algorithm. *J. Chem. Theory Comput.* **2007**, *3* (1), 71–79.

(194)   Assfeld, X.; Rivail, J.-L. Quantum Chemical Computations on Parts of Large Molecules: The Ab Initio Local Self Consistent Field Method. *Chem. Phys. Lett.* **1996**, *263* (1–2), 100–106.

(195)   Sugiki, S.; Kurita, N.; Sengoku, Y.; Sekino, H. Fragment Molecular Orbital Method with Density Functional Theory and DIIS Convergence Acceleration. *Chem. Phys. Lett.* **2003**, *382* (5–6), 611–617.

(196)   Zhang, D. W.; Zhang, J. Z. H. Molecular Fractionation with Conjugate Caps for Full Quantum Mechanical Calculation of Protein–molecule Interaction Energy. *J. Chem. Phys.* **2003**, *119* (7), 3599–3605.

(197)   Fedorov, D. G.; Kitaura, K. The Importance of Three-Body Terms in the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2004**, *120* (15), 6832–6840.

(198)   Li, S.; Li, W.; Fang, T. An Efficient Fragment-Based Approach for Predicting the Ground-State Energies and Structures of Large Molecules. *J. Am. Chem. Soc.* **2005**, *127* (19), 7215–7226.

(199) Jiang, N.; Ma, J.; Jiang, Y. Electrostatic Field-Adapted Molecular Fractionation with Conjugated Caps for Energy Calculations of Charged Biomolecules. *J. Chem. Phys.* **2006**, *124* (11).

(200) Bettens, R. P. A.; Lee, A. M. A New Algorithm for Molecular Fragmentation in Quantum Chemical Calculations. *J. Phys. Chem. A* **2006**, *110* (28), 8777–8785.

(201) Collins, M. A.; Deev, V. A. Accuracy and Efficiency of Electronic Energies from Systematic Molecular Fragmentation. *J. Chem. Phys.* **2006**, *125* (10).

(202) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for Large Systems, with Applications to Water Clusters. *J. Chem. Theory Comput.* **2007**, *3* (1), 46–53.

(203) Fedorov, D. G.; Kitaura, K. Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2007**, *111* (30), 6904–6914.

(204) Kobayashi, M.; Imamura, Y.; Nakai, H. Alternative Linear-Scaling Methodology for the Second-Order Moller-Plesset Perturbation Calculation Based on the Divide-and-Conquer Method. *J. Chem. Phys.* **2007**, *127* (7).

(205) Hirata, S.; Yagi, K. Predictive Electronic and Vibrational Many-Body Methods for Molecules and Macromolecules. *Chem. Phys. Lett.* **2008**, *464* (4–6), 123–134.

(206) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. The Variational Explicit Polarization Potential and Analytical First Derivative of Energy: Towards a next Generation Force Field. *J. Chem. Phys.* **2008**, *128* (23), 234108.

(207) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. Accurate Methods for Large Molecular Systems. *J. Phys. Chem. B* **2009**, *113* (29), 9646–9663.

(208) Soderhjelm, P.; Aquilante, F.; Ryde, U. Calculation of Protein-Ligand Interaction Energies by a Fragmentation Approach Combining High-Level Quantum Chemistry with Classical Many-Body Effects. *J. Phys. Chem. B* **2009**, *113* (32), 11085–11094.

(209) Li, W.; Piecuch, P. Multilevel Extension of the Cluster-in-Molecule Local Correlation Methodology: Merging Coupled-Cluster and Moller-Plesset Perturbation Theories. *J. Phys. Chem. A* **2010**, *114* (24), 6721–6727.

(210) Řezáč, J.; Salahub, D. R. Multilevel Fragment-Based Approach (MFBA): A Novel Hybrid Computational Method for the Study of Large Molecules. *J. Chem. Theory Comput.* **2010**, *6* (1), 91–99.

(211) Rahalkar, A. P.; Katouda, M.; Gadre, S. R.; Nagase, S. Molecular Tailoring Approach in Conjunction with MP2 and Ri-MP2 Codes: A Comparison with

Fragment Molecular Orbital Method. *J. Comput. Chem.* **2010**, *31* (13), 2405–2418.

(212) Mayhall, N. J.; Raghavachari, K. Molecules-in-Molecules: An Extrapolated Fragment-Based Approach for Accurate Calculations on Large Molecules and Materials. *J. Chem. Theory Comput.* **2011**, *7* (5), 1336–1343.

(213) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. Application of the Electrostatically Embedded Many-Body Expansion to Microsolvation of Ammonia in Water Clusters. *J. Chem. Theory Comput.* **2008**, *4* (5), 683–688.

(214) Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. Evaluation of the Electrostatically Embedded Many-Body Expansion and the Electrostatically Embedded Many-Body Expansion of the Correlation Energy by Application to Low-Lying Water Hexamers. *J. Chem. Theory Comput.* **2008**, *4* (1), 33–41.

(215) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for Simulations. *J. Chem. Theory Comput.* **2008**, *4* (1), 1–6.

(216) Leverentz, H. R.; Truhlar, D. G. Electrostatically Embedded Many-Body Approximation for Systems of Water, Ammonia, and Sulfuric Acid and the Dependence of Its Performance on Embedding Charges. *J. Chem. Theory Comput.* **2009**, *5* (6), 1573–1584.

(217) ErinDahlke Speetzen; HannahR Leverentz; Hai Lin; DonaldG Truhlar. Electrostatically Embedded Many-Body Expansion for Large Systems. In *Accurate Condensed-Phase Quantum Chemistry*; Computation in Chemistry; CRC Press, 2010; pp 105–127.

(218) Hua, D.; Leverentz, H. R.; Amin, E. A.; Truhlar, D. G. Assessment and Validation of the Electrostatically Embedded Many-Body Expansion for Metal−Ligand Bonding. *J. Chem. Theory Comput.* **2011**, *7* (2), 251–255.

(219) Kurbanov, E. K.; Leverentz, H. R.; Truhlar, D. G.; Amin, E. A. Electrostatically Embedded Many-Body Expansion for Neutral and Charged Metalloenzyme Model Systems. *J. Chem. Theory Comput.* **2012**, *8* (1).

(220) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Correlation Energy, with Applications to the Calculation of Accurate Second-Order Møller−Plesset Perturbation Theory Energies for Large Water Clusters. *J. Chem. Theory Comput.* **2007**, *3* (4), 1342–1348.

(221) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11* (4), 431–439.

(222) Zhao, Y.; Truhlar, D. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing

of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120* (1-3), 215–241.

(223) Sorkin, A.; Truhlar, D. G.; Amin, E. A. Energies, Geometries, and Charge Distributions of Zn Molecules, Clusters, and Biocenters from Coupled Cluster, Density Functional, and Neglect of Diatomic Differential Overlap Models. *J. Chem. Theory Comput.* **2009**, *5* (5), 1254–1265.

(224) Kaupp, M.; Stoll, H.; Preuss, H. Pseudopotential Calculations for Methyl Compounds of Zinc and Magnesium. *J. Comput. Chem.* **1990**, *11* (9), 1029–1037.

(225) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-adjusted Abinitio Pseudopotentials for the First Row Transition Elements. *J. Chem. Phys.* **1987**, *86* (2), 866–872.

(226) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision A.01, Gaussian, Inc.: Wallingford, CT, 2009.

(227) Dahlke, E. E.; Lin, H.; Leverentz, H.; Truhlar, D. G. MBPAC 2011-2; University of Minnesota: Minneapolis, MN, 2011.

(228) Becker, J. W.; Marcy, A. I.; Rokosz, L. L.; Axel, M. G.; Burbaum, J. J.; Fitzgerald, P. M.; Cameron, P. M.; Esser, C. K.; Hagmann, W. K.; Hermes, J. D. Stromelysin-1: Three-Dimensional Structure of the Inhibited Catalytic Domain and of the C-Truncated Proenzyme. *Protein Sci. Publ. Protein Soc.* **1995**, *4* (10), 1966–1976.

(229) Dennington, R.; Keith, T.; Millam, J. Semichem Inc., Shawnee Mission KS, 2009.

(230) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys. PCCP* **2005**, *7* (18), 3297–3305.

(231) Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; Preuß, H. Energy-Adjustedab Initio Pseudopotentials for the Second and Third Row Transition Elements. *Theor. Chim. Acta* **1990**, *77* (2), 123–141.

(232) Dahlke, E. E.; Lin, H.; Leverentz, H.; Truhlar, D. G. MBPAC 2011-5; University of Minnesota: Minneapolis, MN, 2011.

(233) Browner, M. F.; Smith, W. W.; Castelhano, A. L. Matrilysin-Inhibitor Complexes: Common Themes among Metalloproteases. *Biochemistry (Mosc.)* **1995**, *34* (20), 6602–6610.

(234) Xu, Y.; Feng, L.; Jeffrey, P. D.; Shi, Y.; Morel, F. M. M. Structure and Metal Exchange in the Cadmium Carbonic Anhydrase of Marine Diatoms. *Nature* **2008**, *452* (7183).

(235) Suárez, E.; Díaz, N.; Suárez, D. Thermochemical Fragment Energy Method for Biomolecules: Application to a Collagen Model Peptide. *J. Chem. Theory Comput.* **2009**, *5* (6), 1667–1679.

(236) Mayhall, N. J.; Raghavachari, K. Many-Overlapping-Body (MOB) Expansion: A Generalized Many Body Expansion for Nondisjoint Monomers in Molecular Fragmentation Calculations of Covalent Molecules. *J. Chem. Theory Comput.* **2012**, *8* (8), 2669–2675.

(237) Richard, R. M.; Herbert, J. M. A Generalized Many-Body Expansion and a Unified View of Fragment-Based Methods in Electronic Structure Theory. *J. Chem. Phys.* **2012**, *137* (6).

(238) Isegawa, M.; Wang, B.; Truhlar, D. G. Electrostatically Embedded Molecular Tailoring Approach and Validation for Peptides. *J. Chem. Theory Comput.* **2013**, *9* (3), 1381–1393.

(239) Fedorov, D.; Kitaura, K. Theoretical Development of the Fragment Molecular Orbital (FMO) Method. In *Modern Methods for Theoretical Physical Chemistry of Biopolymers*; Elsevier: Amsterdam, 2006; pp 3–38.

(240) Babu, K.; Gadre, S. R. Ab Initio Quality One-Electron Properties of Large Molecules: Development and Testing of Molecular Tailoring Approach. *J. Comput. Chem.* **2003**, *24* (4), 484–495.

(241) Li, W.; Li, S.; Jiang, Y. Generalized Energy-Based Fragmentation Approach for Computing the Ground-State Energies and Properties of Large Molecules. *J. Phys. Chem. A* **2007**, *111* (11), 2193–2199.

(242) Li, W. Linear Scaling Explicitly Correlated MP2-F12 and ONIOM Methods for the Long-Range Interactions of the Nanoscale Clusters in Methanol Aqueous Solutions. *J. Chem. Phys.* **2013**, *138* (1).

(243)   Wang, B.; Truhlar, D. G. Tuned and Balanced Redistributed Charge Scheme for Combined Quantum Mechanical and Molecular Mechanical (QM/MM) Methods and Fragment Methods: Tuning Based on the CM5 Charge Model. *J. Chem. Theory Comput.* **2013**, *9* (2), 1036–1042.