EVALUATING THE INTERPLAY BETWEEN VIRAL PROTEASES AND HOST UBIQUITIN-LIKE IMMUNOREGULATORY PROTEINS

by

BRENDAN TAYLOR FREITAS

(Under the Direction of Scott Pegan)

ABSTRACT

The 2003-2005 SARS epidemic and the currently ongoing COVID-19 pandemic highlight the threat that zoonotic coronaviruses pose to humans. Both SARS-CoV-1 and SARS-CoV-2 belong to a closely related subgroup of betacoronaviruses known as 2b that primarily infect bats throughout southeast Asia. Coronaviruses encode a papain-like protease (PLpro) that is critical to their success. PLpros are required for the assembly of the viral replicase complex, without which the RNA genome cannot be replicated. Additionally, PLpros suppress the host's innate immune response through the reversal of post-translational modifications by ubiquitin (Ub), and the Ub-like protein (Ubl) interferon (IFN)-stimulated gene product 15 (ISG15). Due to the importance of PLpros, they are common targets of drug discovery research, and some small molecules have been designed with the ability to inhibit SARS-CoV-1 PLpro activity. The PLpros of subgroup 2b coronaviruses are highly conserved, sharing upwards of 80 percent sequence identity. Here we characterize the activity of the PLpros of SARS-CoV-2 as well as a subgroup 2b virus isolated from the greater horseshoe bat, BtSCoV-Rf1.2004, against Ub and ISG15 from various species. We then investigate the ability of naphthalene-based compounds to non-covalently inhibit the PLpros of SARS-CoV-1, SARS-CoV-2, and BtSCoV-Rf1.2004, and their potential for use as pan subgroup 2b coronavirus therapeutics. To this end, we obtained the X-ray crystal structures BtSCoV-Rf1.2004 in complex with two different series of naphthalene-based compounds for use in structure-based drug design. Furthermore, we structurally characterize the Ubl domain of the immunoregulatory protein oligoadenylate synthetase-like (OASL) and analyze its interactions with various coronavirus PLpros as well as the ovarian tumor domain proteases (OTUs) of multiple Nairoviruses. This is the first comparative investigation into the activity and inhibition of subgroup 2b PLpros. It validates the use of naphthalene-based inhibitors against novel subgroup 2b coronaviruses and provides insight that will be useful in optimizing these compounds for use as therapeutics.

INDEX WORDS: Ubiquitin, Deubiquitinase, Interferon Stimulated Gene Product 15, ISG15, 2’-5’ Oligoadenylate Synthetase-Like Protein, OASL, Severe Acute Reparatory Syndrome, SARS, Coronavirus, Papain-Like Protease, PLpro, Nairovirus, Ovarian Tumor Domain Protease, OTU

EVALUATING THE INTERPLAY BETWEEN VIRAL PROTEASES AND HOST UBIQUITIN-LIKE IMMUNOREGULATORY PROTEINS

by

BRENDAN TAYLOR FREITAS

B.S., Virginia Tech, 2017

B.S., Virginia Tech, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

EVALUATING THE INTERPLAY BETWEEN VIRAL PROTEASES AND HOST UBIQUITIN-LIKE IMMUNOREGULATORY PROTEINS

by

BRENDAN TAYLOR FREITAS

Major Professor:	Scott Pegan
Committee:	Éric Bergeron Ralph Tripp William Lanzilotta

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2021

DEDICATION

I would like to dedicate my dissertation to my family, to whom I owe so much of my success. You have all been loving and supportive not only over the course of my PhD, but throughout my entire education process. Mom and Dad, I know I was not always the easiest child or the most driven student. Thank you for not allowing me to squander my potential when I was younger and school did not seem like a priority. Thank you for your guidance along the way. I knew from an early age that I would be choosing a career that you are unfamiliar with, but you always provided me with advice that was beneficial regardless of the path I chose. Jonathan and Alison, thank you for being amazing siblings, and for putting up with all of my science talk over the years. I love and miss you both. I don't get to see any of you as much as I would like, but I look forward to a time hopefully not too far in the future when we can see each other more often.

To my wonderful fiancée Stephanie, you are my best friend and the love of my life. Throughout graduate school you have been my greatest support and a constant source of inspiration. You have always been there for me when I was having a hard time with a project, or needed to practice a presentation, or just needed to decompress because graduate school is stressful. You make every moment better, and thanks to you graduate school was not only bearable, but some of the best years of my life. I cannot thank you enough for all the help you have given me over the past four years. I am so excited that with the conclusion of my PhD we can move on to the next chapter of our adventure together.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor Dr. Scott Pegan. The last four years have been the most challenging of my life, but also some of the most rewarding. When I joined this lab, I knew it would be harder than most, but that I was making an investment in my future and I feel that it is definitely paying off. You have been a great mentor through every step and taught me so much about science and about the career of a scientist. Thank you for pushing me to work harder than I previously thought I was capable. To my committee members, both current and former, Dr. Bergeron, Dr. Lanzilotta, Dr. Tripp, and Dr. Strauch. Thank you all for your valuable input and guidance. Committee meetings were never fun, but they were always helpful in my development as a scientist.

I would also like to acknowledge my current and former lab mates: Dr. Courtney Daczkowski, Dr. Stephanie Bester, Dr. John Dzimianski, Caroline Langley, Jack McGuire, Ian Durie, Justin Shepard, and Brady O'Boyle. I have learned a lot from those of you that came before me, and to those of you that came after I hope I have been able to pass on just as much. To all the undergraduate students I have worked with over the years, particularly Kailee Baker and Jaron Longo, thank you for all your help. All the work you put in is a big reason why graduate students in this lab can be so productive.

TABLE OF CONTENTS

Page

CHAPTER

APPENDICES

LIST OF TABLES

Page

LIST OF FIGURES

LIST OF SCHEMES

CHAPTER 1

INTRODUCTION

Coronaviruses in the 21st century

As of this writing, the most recent coronavirus outbreak has claimed the lives of over 2.7 million people and infected 125 million worldwide. This is the 3rd recorded instance in the last 20 years of a zoonotic coronavirus crossing to humans to cause severe disease[1-3]. The earliest of these outbreaks can be traced to November 2002, although the first cases of Severe Acute Respiratory Syndrome coronavirus 1 (SARS-CoV-1) were reported in February 2003 in Asia[4]. The severe viral respiratory illness spread to twenty-six countries, infecting over eight thousand people with approximately 10 percent mortality before containment was reached within about 4 months[5]. As of 2004, no further SARS-CoV-1 cases have been reported[6]. However, in 2012 the first case of Middle East Respiratory Syndrome coronavirus (MERS-CoV) was confirmed in Saudi Arabia[7]. Since 2012, cases have arisen in twenty-seven countries including a significant outbreak in 2015 in the Republic of Korea[8-10]. Outbreaks of MERS-CoV continue to this day with the majority of MERS-CoV cases occurring in the Middle East, primarily in Saudi Arabia. Despite MERS-CoV having a mortality rate exceeding 35 percent, there is still a lack of effective treatments and vaccines[11]. Subsequently, in December 2019, the first cases of Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2), the causative agent of the COVID-19 pandemic, were recognized in Wuhan city, Hubei province, China[12, 13]. SARS-CoV-2 spread rapidly, leading to a worldwide pandemic. The mortality rate of

SARS-CoV-2 is lower than MERS-CoV and SARS-CoV-1, but the virus is highly contagious and spread rapidly[14]. This pattern of zoonotic coronaviruses crossing to humans suggests that more outbreaks of novel coronaviruses will occur in the coming decades.

Viruses of the subfamily *Orthocoronavirinae* in the family *Coronaviridae* infect a wide range of species including humans and can be divided into four genera: alpha, beta, gamma, and delta. In humans, coronaviruses generally cause mild to moderate upper respiratory tract illnesses, such as the common cold. Despite the periodic emergence of new human coronaviruses, it can be difficult to identify when less pathogenic spillover events happen as infections with mild symptoms largely go unreported [13, 15]. For instance, some human coronaviruses (HCoV)s such as HCoV-229E and HCoV-OC43 have long been known, while others such as HCoV-NL63 and HCoV-HKU1 were identified more recently[16]. In total, seven HCoVs have been identified so far, including two alpha and five betacoronaviruses[7, 17]. Alpha and betacoronaviruses, such as porcine epidemic diarrhea virus (PEDV) and mouse hepatitis virus (MHV), primarily replicate in bats but infect many diverse mammalian hosts[18]. The three recent spillovers, MERS-CoV, SARS-CoV-1, and SARS-CoV-2, all belong to the betacoronavirus genus. Specifically, SARS-CoV-1 and SARS-CoV-2 belong to betacoronavirus subgroup 2b, a family of over 30 SARS-like viruses[19, 20]. On the other hand, gamma and deltacoronaviruses, such as avian infectious bronchitis virus (IBV) and porcine deltacoronavirus (pΔCoV), primarily replicate in birds but infect some mammalian hosts such as pigs, beluga whales, and Asian leopard cats[18]. The ability to infect a wide range of host species is likely a contributing factor to the widespread prevalence and recurring emergence of coronaviruses.

Dual functionality of viral proteases

Coronaviruses have a non-segmented positive sense, single-stranded RNA genome, which is surrounded by nucleocapsid proteins (N) and together they appear as a coiled tubular helix within the viral envelope [21, 22]. The viral envelope is made up of a lipid bilayer with anchored envelope (E), membrane (M), and spike (S) structural proteins. E proteins are only a minor component of virions but show significant sequence variation across coronavirus groups and in particular cases, within a single group[23]. They are however responsible for virion assembly, intracellular trafficking, and morphogenesis. Meanwhile, M proteins are the major structural protein of the envelope, providing for its overall shape[23]. M proteins are also crucial during several stages of the viral lifecycle including the assembly, budding, envelope formation, and pathogenesis stages[24, 25]. The S glycoprotein is the most prominent of the three viral envelope proteins as it assembles into trimers to form the distinctive surface spikes of coronaviruses[26]. As a class I fusion protein, it is vital for viral infection as it mediates receptor attachment and membrane fusion between the virus and host cell[23, 24]. The S protein is composed of an S1 and S2 subunit, where the S1 subunit forms the head of the spike and contains the receptor-binding domains (S1-NTD and S1-CTD) and the S2 subunit anchors the spike to the viral envelope and enables fusion upon protease activation[26]. As infection of a cell is triggered by an S protein binding to a complementary host cell surface receptor or sugar, S1 proteins demonstrate substantial sequence diversity as they are responsible for host cell specificity[23]. Several HCoV receptors have been identified, including angiotensin-converting enzyme 2 (ACE2) for SARS-CoV-1, SARS-CoV-2, and HCoV-NL63, dipeptidyl peptidase 4 (DPP4) for MERS-CoV, and human aminopeptidase N (APN) for HCoV-229E[27-29]. After receptor binding, a protease enzyme cleaves and activates the S protein allowing the virus to enter the host cell where the virus particle is uncoated and the viral genome is released before being transcribed by host machinery into two polyproteins, pp1a and pp1ab[24, 26]. The polyproteins encode up to 16 nonstructural protein (NSP)s, which must be processed by viral proteases into their functional units to form the viral replicase complex[26]. The replicase complex is responsible for transcribing the viral RNA genome before translation of the virus' N protein and S, E, and M structural proteins[13]. These components are ultimately converted into complete virions inside the endoplasmic reticulum (ER)-Golgi intermediate compartment[30]. Mature virions are then transported to the cell surface in vesicles and released to infect further host cells.

The processing of the polyproteins into functional units by two viral cysteine proteases is essential for successful viral replication. The first of which is the 3-chymotrypsin like protease (3CLpro), also known as the main protease (Mpro). It is located on NSP5 and is responsible for cleaving NSPs 4-16[31]. The other is the papain-like protease (PLpro), known as PLPs in viruses that encode two, which is located on NSP3[32]. PLpros cleave NSPs 1-3 at ELNGG, RLKGG, and SLKGG motifs respectively[32]. PLpros target Gly-Gly motifs within the viral polyprotein which bind at the active site, also known as the P1 site, and a site of interaction one residue upstream of the C-terminal glycine known as the P2. PLpro active sites typically contain a Cys-His-Asp catalytic triad along with an oxyanion hole, which for subgroup 2b PLpros contains a tryptophan[33] (Figure 1.1). Residue interaction sites adjacent to the active site are named PX to represent that they interact with an amino acid X number of residues upstream of the cleavage site. The P3/P4 pocket is adjacent to the active site and also contains the flexible BL2 loop, which can adopt open and closed conformations to accommodate substrate peptide chains. The P1-P5 sites recognize and bind the XLXGG cleavage sites found between NSPs 1-3.

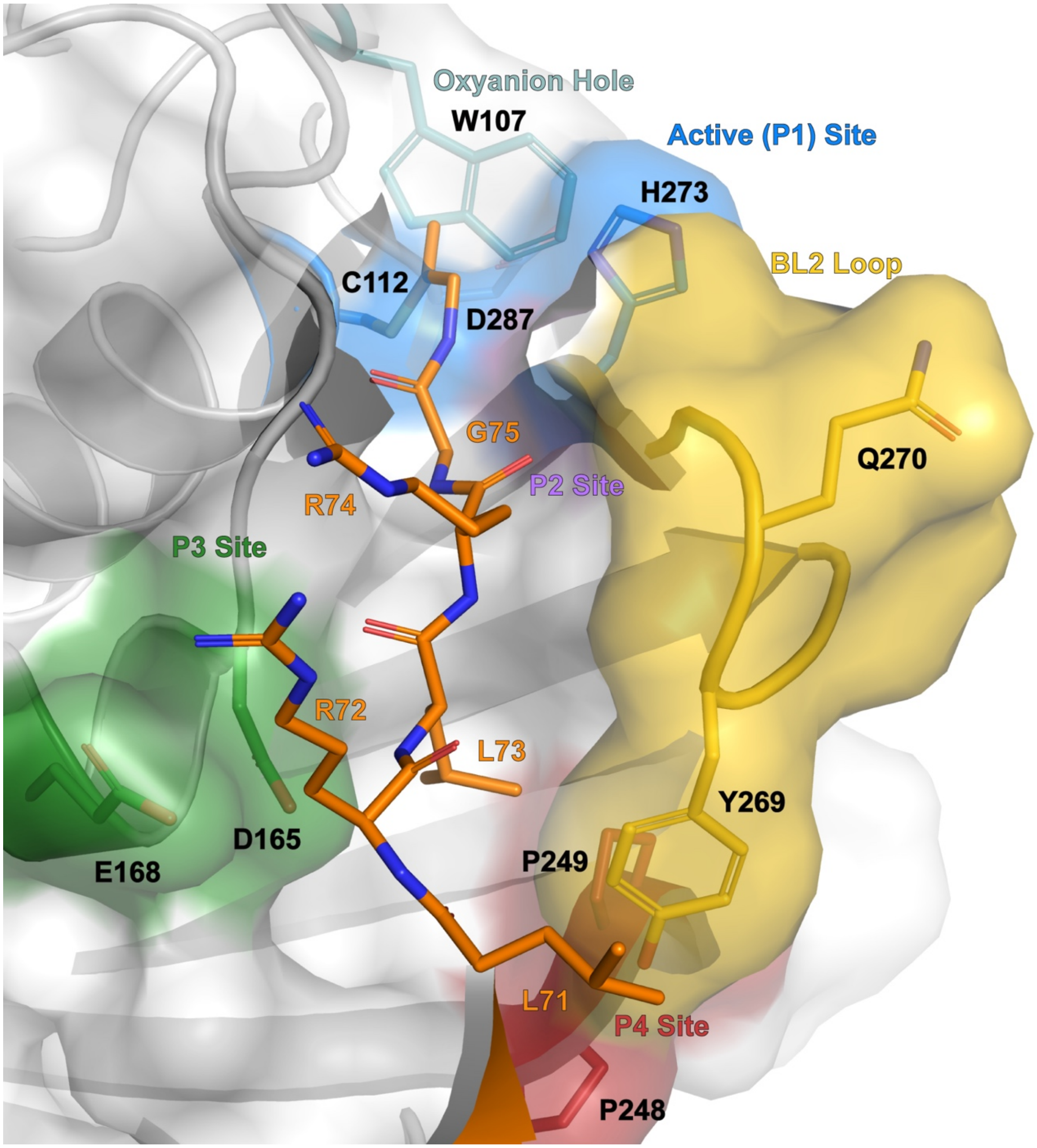


Figure 1.1. Critical sites of interaction between PLpros and protein substrates. Cartoon representation of SARS-CoV-1 PLpro (grey) bound to K48 linked di-Ub (orange) with interacting residues shown in stick form. Active site (blue), oxyanion hole tryptophan (teal), P2 site (purple), P3 site (green), and P4 site (red) are denoted with colored sticks and surfaces. The BL2 loop (yellow) is shown in an open conformation to accommodate the LRLRGG peptide chain. The surface of L163 and N110 are hidden to show the active site.

In addition to cleaving the viral polyprotein, PLpros serve a secondary, immunosuppressive role by reversing post-translational modifications by Ub and the Ubl interferon (IFN) stimulated gene product 15 (ISG15)[34, 35]. The protease cleavage sites between NSPs 1-3 that PLpros target are similar enough to the conserved LRLRGG motifs at which Ubls are conjugated to targets that PLpros have adapted to process both. PLpros are approximately 315 residues long and are comprised of (from N-terminal to C-terminal) a ubiquitin (Ub)-like protein (Ubl) domain, thumb domain, palm domain, and zinc finger domain. The purpose of the Ubl domain remains unclear and some PLpros have been shown to function without one[36]. The thumb, palm, and zinc finger domains each contain part of the Ub interacting motif (UIM), which is the site of interaction between PLpros and their Ubl substrates[37]. The UIM, also known as the proximal Ub or S1 binding pocket, is located on the zinc finger and palm domains and binds to the Ubl domain being cleaved. The distal Ub or S2 binding pocket is located on the thumb domain and binds the Ubl domain immediately upstream of the Ubl being cleaved. Binding poly-Ub or ISG15 at these pockets causes the C-terminal Ubl to be oriented in a way that its C-terminus is within the active site, located within the palm domain. Because of this, the surface residues within these pockets are critical determinants of Ubl substrate specificity[37].

Proteases capable of cleaving Ub and ISG15 are known as deubiquitinases (DUB)s and deISGylases, and do not cleave all Ubls equally. Many viral proteases are adapted to specifically cleave Ubls involved in the pathways most detrimental to their virus' ability to replicate. This typically manifests in preferences in poly-Ub chain linkage type or species of ISG15. Commonly, viral proteases display preferences towards ISG15s from their natural host species, along with off-target activity towards ISG15s of closely related species. Owing to the wide range of species coronaviruses infect, and therefore the diversity in immune responses they have to circumvent, PLpros can vary considerably in both their substrate preferences and kinetic properties. PLpros of betacoronaviruses, in particular, tend to have a strong preference for ISG15 over Ub, and several have low enough DUB efficiency that they can be saturated by mono-Ub at baseline cellular concentrations[38]. The effects of this trait are compounded in PLpros from subgroup 2b viruses which have a distinct di-distributive cleavage pattern, which drastically reduces their ability to process Ub chains comprised of less than three Ub monomers[39].

PLpro functionality has been shown to be vital to the success of coronaviruses[35, 40-42]. While the specific immunosuppressive function that deISGylase activity serves for coronaviruses is still unknown[43, 44], a recent study demonstrated that when the PLP2 of the subgroup 2a MHV had its DUB activity ablated through mutation, viral titers were markedly reduced[45]. The reduced pathogenicity was attributed, at least in part, to the host IFN response being triggered faster than it was by wild-type MHV. Despite the remaining uncertainties about the mechanisms of PLpro immunosuppression, its dual functionality makes it a highly sought-after target for the development of anti-coronavirus therapeutics[33, 46, 47].

Inhibiting PLpros as a means of treating coronavirus infections

Several strategies have been investigated to treat and prevent coronavirus infections including vaccines and small molecule inhibitors of their viral proteases, the 3CLpro and PLpro. In response to the emergence of a novel virus, priority is often placed on the development of a vaccine to immunize recipients and reduce viral spread through herd immunity. While vaccines are a crucial part of responding to the outbreak of novel pathogens, even under ideal conditions the development and large-scale distribution of a safe and effective vaccine can take over a year, as has been the case with the COVID-19 pandemic. Furthermore, populations that are often most susceptible to infectious diseases such as the elderly, newborns, or the immunocompromised often cannot receive vaccines depending on the platform used[48]. Vaccines are also specifically tailored to particular antigens and only confer immunity to the pathogen for which they were designed. RNA viruses are particularly prone to mutation, increasing the likelihood that a vaccine-resistant strain could emerge. Due to these factors, it is important to also develop small-molecule antiviral therapeutics that target highly conserved viral features. If there had been greater investment in this endeavor after the 2003 SARS epidemic, we may have been able to mitigate much of the damage from the 2019 COVID-19 pandemic.

Due to its role in the cleavage of NSPs 4-16, the 3CLpro has often been a target for small molecule inhibitor development[49, 50]. Immediately following the emergence of SARS-CoV-2 there was a flood of high throughput, small molecule molecular docking papers published focusing on many known coronavirus enzymes, with the 3CLpro being among the most common. A considerable number of high throughput screening (HTS) assays were performed as well, identifying compounds such as Quercetin, which is expected to non-covalently bind at the active site[51]. However, this compound, like many others, would require further optimization before it could be used as a therapeutic. There have also been attempts made to use compounds with efficacy towards other related viral proteases such as AG-7088 which has already been shown to inhibit rhinovirus 3C protease[52]. Many of these methods are being used on both the 3CLpro and PLpro of SARS-

CoV-2, and with enough interest and investment could deliver a drug or combination of drugs that will be effective at treating COVID-19.

As the PLpro plays a vital role in viral replication, it is also often targeted during coronavirus drug development[53-56]. One method employed was the covalent modification of the catalytic cysteine by electrophilic warheads. This method is effective at inhibiting PLpro activity but faces two key shortcomings in applicability as a therapeutic. Primarily, covalently binding inactivators are often less specific than competitive inhibitors, resulting in more off-target effects[57]. Secondly, as these electrophilic warheads are covalent modifiers, the effects are irreversible, which can lead to higher toxicity[57]. As a result, non-covalent competitive inhibitors are often a preferable alternative as they tend to have greater specificity and lower toxicity than covalent modifiers. Additionally, they can target conserved regions, which allows them to be effective against multiple closely related viruses. Furthermore, small molecules are often cheaper and faster to produce than vaccines. Moreover, scaling up the production and distribution of an approved antiviral therapeutic with proven efficacy against a novel virus can occur more rapidly than with a vaccine.

One such group of compounds are the naphthalene-based SARS-CoV-1 PLpro inhibitors. Originally identified in 2008, series I naphthalene inhibitors are comprised of a naphthyl group connected to an arene group by an amide bond[58]. The lead compound of the 2008 study, GRL0617, was shown to inhibit SARS-CoV-1 PLpro activity at high nanomolar concentrations and suppress SARS-CoV replication at low micromolar concentrations with negligible cytotoxicity[58]. The inhibitors bind to the active site adjacent, P3/P4 pocket of SARS-CoV-1 PLpro that binds the XLX residues of the polyprotein or

Ubls immediately upstream of the Gly-Gly cleavage site. Later studies have introduced a second series of naphthalene-based inhibitors that utilize piperidine-based backbones in place of amide bonds[46]. Series II compounds bind to the same pocket as those of series I, however, their extended backbones cause their eastern arene group to wrap around the BL2 loop rather than interact at the P3 site. By blocking peptide chains from binding to the P3 and P4 sites series I and II compounds inhibit both polyprotein cleavage as well as DUB and deISGylase activity.

The antiviral innate immune response

Ub and Ubls play a critical role within the antiviral innate immune response. The response is comprised of several signaling cascades that result in the upregulation of IFNs, cytokines, chemokines, and ISGs which induce an antiviral state[59-61]. The innate immune response is first triggered when pattern recognition receptors (PRR)s detect pathogen-associated molecular patterns (PAMP)s within the cytosol[62]. PAMPs can include double- or single-stranded RNA (dsRNA/ssRNA) as well as viral proteins, which are recognized by different classes of PRRs. Three key PRRs are the toll-like receptors (TLR)s, oligoadenylate synthetases (OAS)s, and retinoic acid-inducible gene I (RIG-I)-like receptors (RLR)s (Figure 1.2).

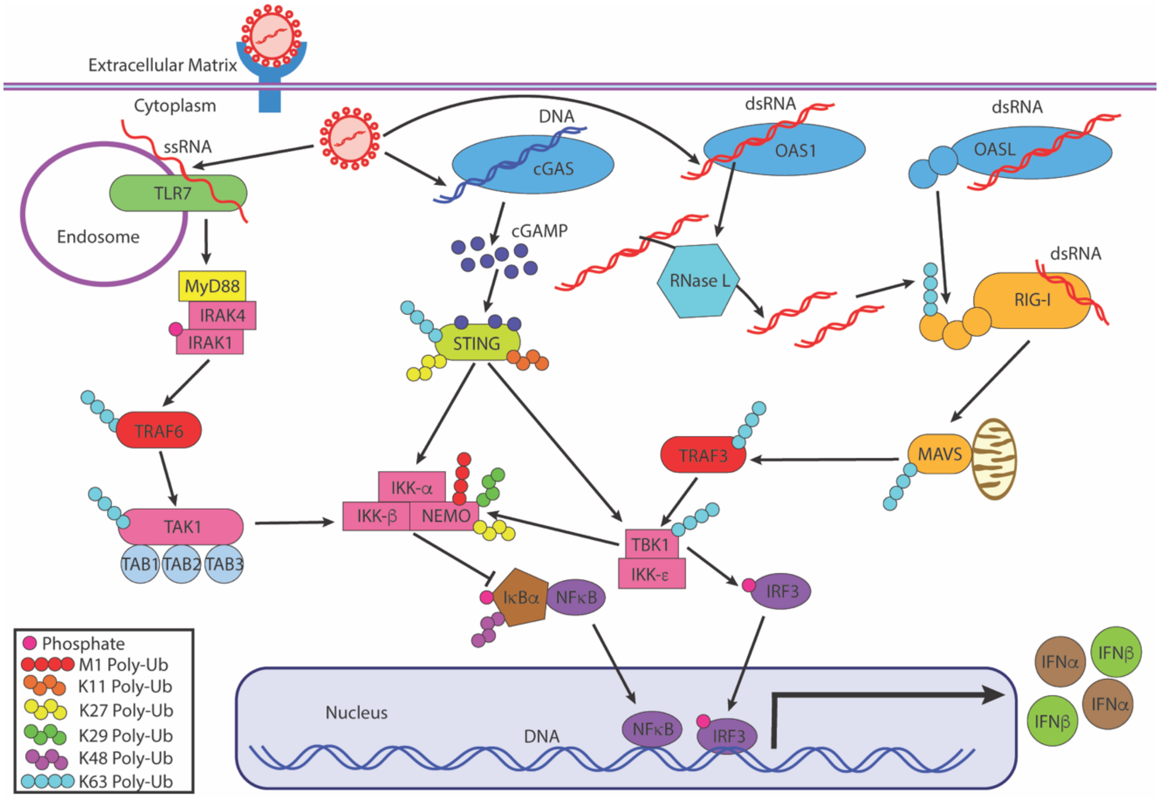


Figure 1.2. Ubiquitin in antiviral innate immune response pathways. Overview of the MyD88dependant TLR7 pathway, cGAS/STING pathway, OAS/RNase L pathway, and RIG-I pathway.

TLRs are expressed in immune cells and can generally be grouped into cell surface TLRs and endosomal TLRs[63-65]. TLRs recognize a wide variety of PAMPs with those on the cell surface primarily recognizing microbial membrane components such as proteins and lipids, while those found in the endosome typically respond to nucleic acids[64, 66]. TLR7 in particular recognizes viral ssRNA and leads to increased production of type-I IFNs, IFN-α, and IFN-β, as well as proinflammatory cytokines through the NF-κB pathway[67]. TLR7 initiates a signaling cascade through myeloid differentiation primary response gene 88 (MyD88) which forms a complex with interleukin-I receptor kinase 4 (IRAK4) called a Myddosome and activates IRAK1[68]. Autophosphorylation causes IRAK1 to dissociate from MyD88 and bind to the E3 ligase tumor necrosis factor (TNF) receptor-associated factor 6 (TRAF6)[69-71]. Together with the E2 Ub conjugating enzymes Ubc13 and Uev1A, TRAF6 promotes K63 linked polyubiquitination of itself and TGF-β-activated kinase 1 (TAK1)[72]. TAK1 forms a complex with TAK1-binding proteins 1, 2, and 3 (TAB1/2/3), and then activates the IKK complex pathways. The IκB kinase (IKK) complex is comprised of IKKα, IKKβ, and IKKγ (also known as NEMO) and activates NF-κB by phosphorylating the NF-κB inhibitory protein IκBα, causing it to be degraded[73-76]. Activated NF-κB translocates to the nucleus and induces the production of proinflammatory cytokines, such as TNFα[77].

OASs are a family of nucleotidyltransferase (NTase) enzymes that respond to the presence of dsRNA[78, 79]. OAS proteins are comprised of 1 or more OAS domains, each containing an NTase active site as well as a dsRNA binding site[80]. Upon binding to viral RNA, the oligomerized OAS enzyme catalyzes the synthesis of 2'-5'-linked oligoadenylate (2-5A). 2-5A in turn activates the latent endoribonuclease, RNase L, which degrades cellular and viral RNA[81-83]. In addition to preventing the translation of viral proteins, the cleaved dsRNA then goes on to activate RLRs such as RIG-I and melanoma differentiation-associated protein 5 (MDA5)[84]. Another NTase enzyme, cyclic GMP-AMP (cGAMP) synthetase (cGAS), responds to DNA in the cytosol is an upstream activator of NF-κB and IFN regulatory factor 3 (IRF3). cGAS is activated when bound to DNA and begins synthesis of cGAMP, which oligomerize with stimulator of IFN genes (STING) at the ER[85]. Oligomerized STING translocates to the Golgi to interact with TANK-binding kinase 1 (TBK1), which in turn phosphorylates IRF3[86]. Similar to TAK1, STING also activates an IKK, resulting in degradation of IκBα and NF-κB translocation[86]. In addition to the

OAS/RNase L and cGAS/STING pathways, a unique member of the OAS family is capable of activating the RLR pathway as well. OAS-like protein (OASL) contains two C-terminal Ubl domains in addition to a single N-terminal OAS domain[87]. Much is still unknown about OASL, and its function varies between species, but in humans, the OASL Ubl domains interact with RIG-I to activate the RLR signaling cascade, which is the pathway K63 poly-Ub is involved in as well [80].

In response to binding short dsRNA, RLRs such as RIG-I and MDA5 undergo a conformational change that allows their caspase activation and recruitment domains (CARD)s to be K63 ubiquitinated. Polyubiquitination of RLRs induces an interaction with the mitochondrial antiviral signaling (MAVS) protein[88, 89]. MAVS activates TRAF3, which recruits TBK1and IKKε, ultimately activating IRF3 and NF-κB through the same signaling cascade as cGAS[90]. The PRR/PAMP system is a robust method of detecting many different pathogens or endogenous disease states to upregulate the production of type-I IFNs, proinflammatory response molecules, and ISGs. All of the pathways discussed so far rely heavily on posttranslational modification by Ub and Ubls. This versatility and ubiquity make Ubls invaluable to our ability to mount effective antiviral defenses, and therefore make them common targets of viral immunosuppressive mechanisms.

Ub structure enables signaling diversity

Despite its simple structure, Ub is capable of regulating so many vital immune pathways due to its ability to form poly-Ub chains. Ub is one of the most abundant proteins within the cell, with concentrations ranging from 10-23 uM[38]. It is also one of the most highly conserved proteins, with no variation between mammals and greater than 90 percent conservation across all plants, animals, and fungi[91]. Ub is a 76 amino acid protein comprised of a single β-grasp domain ending in an LRLRGG motif. β-grasps contain three N-terminal and two C-terminal β-sheets wrapped around a central α-helix, resulting in a somewhat spherical fold[92]. The N-terminal methionine (M1) along with seven lysines (K6, K11, K27, K29, K33, K48, and K63) are solvent-exposed and can be covalently bonded to the C-terminal glycine of another Ub[92]. Because of this Ub can form homotypic, heterotypic, or branched poly-Ub chains, as well as post-translationally modified chains and Ub-Ubl mixed chains[93]. Mono-Ub along with the various poly-Ub chains are recognized as unique signals within the body, allowing Ub to mediate many different pathways[93].

Each chain type is recognized differently within the body and corresponds to specific signals. The two most common targets of viral DUBs, K48 and K63 linkages, are also the best studied. The main function of K48 poly-Ub is marking proteins for proteasomal degradation[94]. This includes maintenance of homeostasis through the elimination of misfolded proteins but also has antiviral effects. The proteasomal degradation of IκBα mentioned previously that allows for NF-κB translocation is signaled for by K48 polyubiquitination of IκBα[95]. Additionally, viral proteins can be degraded when marked with K48 poly-Ub, leading to reduced viral replication[96, 97]. K63 linked poly-Ub, similar to M1 poly-Ub, is involved with the regulation of NF- κB signaling. As was mentioned previously, the RLR signaling cascade is also initiated following K63 polyubiquitination. The prevalence of viral DUBs that target K48 and K63 poly-Ub suggests these two linkage types are particularly detrimental to viral success.

The wide range of functions Ub serves via post-translational modification are made possible by a group of proteins called activating (E1), conjugating (E2), and ligating (E3) enzymes[98]. These enzymes work in consort to covalently bond the C-terminal glycine to a primary amine, either through a peptide bond at the N-terminus of a protein or via an iso-peptide bond to a lysine side chain[98]. This occurs in a three-step reaction in which the E1 forms a thioester linkage with Ub, after which it is transferred to the active site cysteine of the E2, before finally being bound to the target by the E3[98]. This mechanism can be downregulated by endogenous DUBs that reverse ubiquitination to suppress various signals[99]. E3 ligases are typically associated with specific linkage types and therefore specific pathways, such as Ubc13 and Uev1A[98, 100]. In addition to Ub, some E3 ligases such as TRIM25 and HERC5 have been associated with other Ubls[101, 102]. One such Ubl that plays a significant part in the interplay between viruses and the host's innate immune system is ISG15.

CHAPTER 2

HOW ISG15 COMBATS VIRAL INFECTION[1]

---

[1] Freitas, B.T., F.E.M. Scholte, E. Bergeron, and S.D. Pegan. 2020. *Virus Research*. 286, 198036. doi: 10.1016/j.virusres.2020.198036.

https://www.sciencedirect.com/science/article/pii/S0168170220300939?via%3Dihub

Reprinted here with permission of the publisher.

Abstract

Interferon (IFN)-stimulated gene product 15 (ISG15) is a ubiquitin-like protein critical for the control of microbial infections. ISG15 appears to serve a wide variety of functions, which regulate multiple cellular responses contributing to the development of an antiviral state. ISG15 is a versatile molecule directly modulating both host and virus protein function which regulate many signaling pathways, including its own synthesis. Here we review the various roles ISG15 plays in the antiviral immune response, and examine the mechanisms by which viruses attempt to mitigate or exploit ISG15 activity.

Introduction

One of the first lines of defense against viral pathogens is the innate immune response restricting virus replication early after infection. Immune cells such as neutrophils and macrophages, as well as inflammatory cytokines such as interferons (IFNs) are critical elements of the innate immunity. One highly upregulated IFN-stimulated gene product (ISG) protein in the initial stages of the innate response to viral infection is ISG15. Human ISG15 is a 15kDa Ubiquitin (Ub)-like (Ubl) protein that structurally resembles two β-grasp Ub domains connected by a linker sequence(Narasimhan et al., 2005). ISG15 can act as both an effector and a signaling molecule in various phases of the innate immune response. ISG15 was first discovered in 1979 with references to a 15-kDa protein induced by IFN treatment(Farrell et al., 1979). ISG15's Ubl structure was first discovered in 1987 when it cross-reacted with anti-Ub antibodies(Haas et al., 1987). It was quickly suspected of serving an immunological function due to ISG15 being one of the most highly upregulated genes during viral infections(Der et al., 1998; González-Sanz et al., 2016; Meraro et al., 2002).

Similar to Ub and other Ubls, ISG15 is able to regulate a wide range of cellular pathways tied to host immune responses. ISG15 can be post-translationally linked to a wide array of target proteins, both cellular and viral(Huang et al., 2014; Loeb and Haas, 1992; Okumura et al., 2007; Zhang et al., 2019; Zhao et al., 2016). The ISG15 conjugation mechanism mirrors that of ubiquitination as it requires three enzymes; the E1 activating enzyme (Ube1L), E2 conjugating enzyme (UbcH8), and an E3 ligase enzyme (HERC5, EFP, or TRIM25)(Zhang and Zhang, 2011). These enzymes work together to covalently bond the C-terminal glycine of the ISG15 LRLRGG motif to a lysine of the target protein(Berndsen and Wolberger, 2014; Wong et al., 2006). This process is initiated when ISG15 is converted from its 17kDa precursor form by currently unidentified cellular proteases into its 15kDa mature form. ISGylation is principally reversed by one host protease, ubiquitin-specific protease 18 (USP18)(Potter et al., 1999). Beyond acting through conjunction with other proteins, ISG15 has also been found to have numerous immunological roles in its free form, both intra- and extracellularly(Baldanta et al., 2017; Campbell and Lenschow, 2013; Napolitano et al., 2018; Recht et al., 1991; Swaim et al., 2017; Yeung et al., 2018).

In this review we describe the specific roles ISG15 plays in the antiviral immune response. ISG15 can vary considerably between species in sequence, structure and function(Deaton et al., 2016). ISG15 has been shown to be a vital part of the immune response in mice. The absence of ISG15 in knockout mice and mouse cell lines results in a significant reduction in the ability to mount a viable defense against a variety of viral infections(Lenschow et al., 2007; Speer et al., 2016). In contrast, human patients lacking ISG15 do not appear to be more susceptible to viral infection, but are more susceptible to some bacterial infections(Bogunovic et al., 2012). Various viruses have developed methods to counteract the antiviral effects of ISG15, including the reversal of ISGylation, sequestering of ISGylated proteins, and interfering with ISG15 synthesis(Daczkowski et al., 2017b; Gargan et al., 2018; Zhao et al., 2016).

Induction and regulation of ISG15 signaling

As the name implies, ISG15 is an IFN-stimulated gene, and its expression is therefore upregulated upon IFN stimulation (FIG. 2.1). Type I IFN is induced when Pathogens-Associated Molecular Patterns (PAMPs), including double-stranded viral RNA (dsRNA), are detected in the cytoplasm by retinoic acid-inducible gene 1 protein (RIG-I) or other RIG-I-like receptors (RLRs) such as melanoma differentiation-associated protein 5 (MDA5). Upon binding to dsRNA RLRs undergo a conformational change that allows ubiquitination of their caspase recruitment domains (CARDs). The ubiquitinated CARDs are recruited to the mitochondria by mitochondrial antiviral-signaling protein (MAVS) where they initiate a signaling cascade leading to the downstream nuclear translocation of interferon regulatory factor 3 (IRF3) and NF-κB. These transcription factors are important activators of the type I IFN promoters, resulting in increased production of IFN-α and IFN-β(Kalvakolanu and Borden, 1996; Meraro et al., 2002). IRF3 can directly activate the transcription of several ISGs including ISG15(Grandvaux et al., 2002). Type I IFNs are secreted and subsequently bind the IFN-α/β receptor (IFNAR) on the cell membranes of the infected cell and neighboring cells. Activated IFNAR signals for Janus kinase 1 (JAK1) to phosphorylate signal transducer and activator of transcription (STAT) proteins, which form a complex with IRF9 called IFN-stimulated gene factor 3 (ISGF3). ISGF3 binds the IFN-sensitive response element (ISRE) within ISG promoters, increasing transcription and expression of hundreds of ISGs, including ISG15 and its conjugating enzymes Ube1L, UbcH8, EFP, TRIM25, and HERC5, as well as USP18 (FIG. 2.1)(Zhang and Zhang, 2011). IFN-mediated upregulation of ISGs results in an antiviral state, thus reducing viral spread.

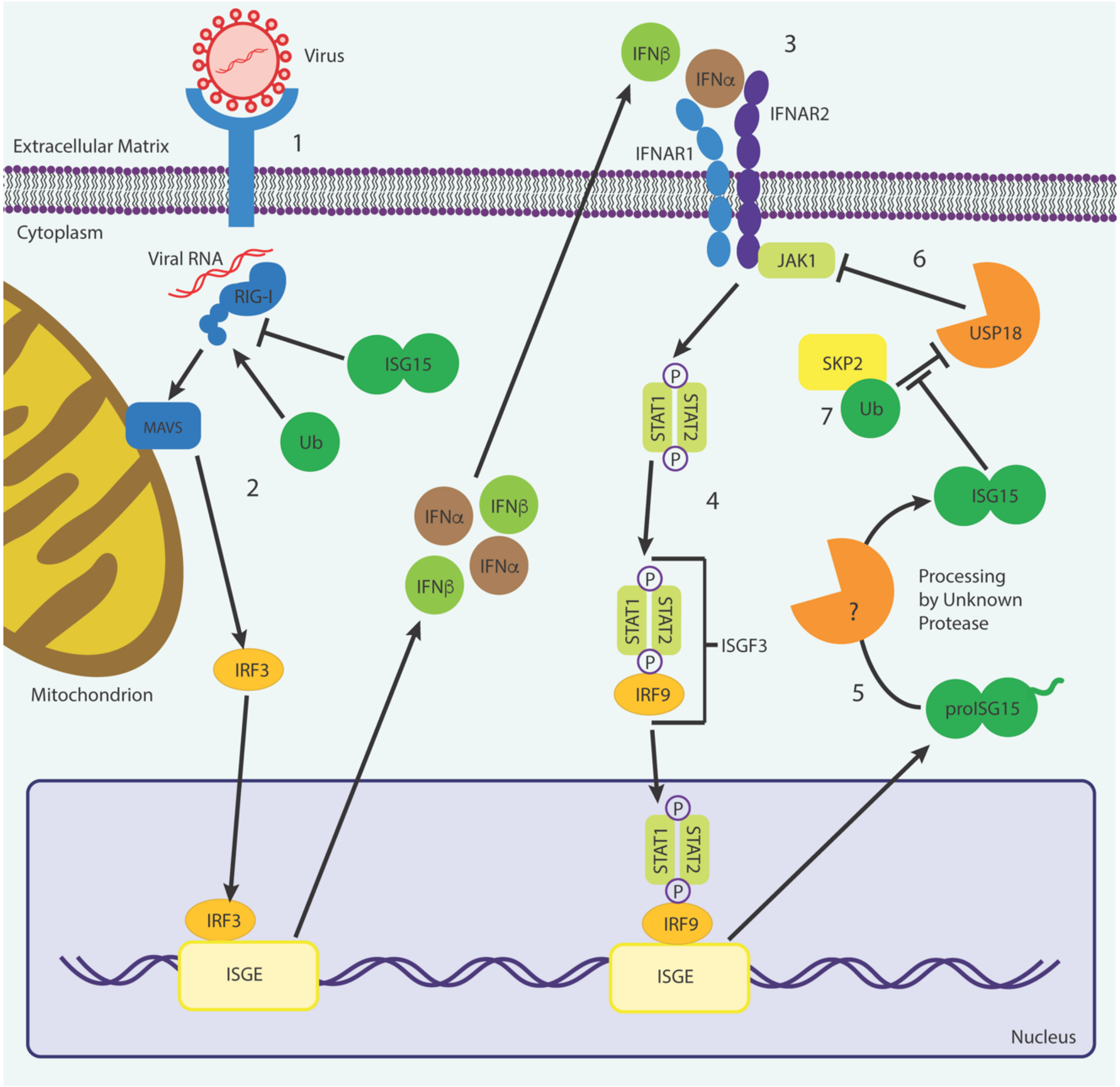


Figure 2.1. (1) Upon infection viral RNA or DNA is released into the cytoplasm where it is detected by RIG-I or other RLRs. After binding to the nucleic acid, the RLR will undergo a conformational change, exposing the CARDs. (2) The exposed CARDs can be ubiquitinated or ISGylated. Ubiquitinated CARDs activate MAVS, which leads to downstream activation of IRF3. Activated IRF3 increases transcription at the ISREs for IFNα and IFNβ. ISGylation of RIG-I inhibits this pathway by marking RIG-I for proteasomal degradation. (3) Type 1 IFNs bind extracellularly to IFNAR, activating JAK1. (4) Activated JAK1 phosphorylates STAT1 and STAT2, which bind to IRF9 to form ISGF3. ISGF3 binds to the ISREs of many ISGs including ISG15, Ube1L, UbcH8, and HERC5. (5) proISG15 is processed by an unknown protease into functional free ISG15. (6)USP18 can inhibit activation of the JAK/STAT pathway by binding competitively with IFNAR2 preventing JAK1 activation. (7) ISG15 stabilizes the interaction between IFNAR2 and USP18 by inhibiting SKP2 mediated Ub conjugation. This prevents proteasomal degradation of USP18, effectively down regulating production of ISG15. ISGylation of RIG-I CARDs downregulates MAVS activation and IFN production.

ISGylation of cellular proteins interferes with viral replication

ISG15 can hinder viral replication by interfering with the endogenous translation and exocytosis machinery that viruses hijack in order to replicate. Eukaryotic initiation factor 4E (eIF4E) facilitates translation initiation. eIF4E homologous protein (4EHP) binds to the cap of both cellular and viral mRNAs preventing translation by competing with eIF4E. ISGylation of 4EHP enhances this inhibitory effect on translation, likely due to a stabilization of the 4EHP-cap interaction. It has been proposed that ISG15 may selectively conjugate 4EHP-capped viral mRNAs as a mechanism of selective inhibition of viral RNA translation(Okumura et al., 2007). Potential targets include flaviviruses such as Dengue, West Nile, yellow fever, Kunjin, and Japanese encephalitis viruses, which contain capped positive sense RNA.

Furthermore, ISG15 has been described to interfere with the budding of virus-like particles (VLPs) by inhibiting endogenous enzymes required for this process. Ebola relies on ubiquitination of its VP40 matrix protein by E3 ligase Nedd4 to facilitate virion egress. ISGylation of Nedd4 prevents VP40 ubiquitination and inhibits budding(Malakhova and Zhang, 2008; Okumura et al., 2008; Pincetic et al., 2010; Seo and Leis, 2012). Budding of Avian Sarcoma Leukosis virus (ASLV) and human immunodeficiency virus (HIV)-1 requires recruitment of ESCRT-III complex proteins such as Vps4 and LIP5. ISGylation of the ESCTRT-III protein CHMP5, which regulates LIP5, prevents oligomerization of Vps4 and LIP5, halting the budding process(Pincetic et al., 2010).

ISGylation of viral proteins

One of the best-studied ways ISG15 affects viral replication is through conjugation to viral proteins. ISGylation can hinder protein function, mark proteins for degradation, affect protein localization, and prevent the formation of protein complexes(Nakashima et al., 2015; Tang et al., 2010; Villarroya-Beltri et al., 2016; Zhang et al., 2019; Zhao et al., 2016). Unlike Ub and other Ubl proteins, ISG15 has a conjugation preference towards newly synthesized proteins(Durfee et al., 2010). The proposed model suggests that HERC5 associates with the 60S ribosomal subunit, which contains the exit tunnel, and allows for preferential ISGylation of newly synthesized proteins. However, only a small percentage of newly synthesized proteins is ISGylated. What drives certain proteins to become ISGylated over others during translation is currently unknown. However, many proteins produced during viral infection are viral proteins and ISGs, potentially allowing ISG15 to maximize its impact on viral replication.

ISGylation interferes with viral protein function

One of the first identified viral targets of ISG15 is the NS1 protein of Influenza A virus (IAV)(Tang et al., 2010). IAV NS1 can inhibit the production of type I IFNs, thereby preventing activation of ISGs including protein kinase R (PKR) and oligoadenylate synthetase (OAS), which results in increased translation of viral RNA(Bergmann et al., 2000; de la Luna et al., 1995; Min and Krug, 2006; Zhang et al., 2015). In addition, ISGylation inhibits the ability of NS1 to bind, among other things, dsRNA and PKR, reducing viral suppression of the pathways in which these factors are involved(Tang et al., 2010). ISGylated NS1 is also prevented from associating with importin-α, which mediates the translocation of NS1 into the nucleus(Zhao et al., 2010). ISGylation of NS1 also impacts its ability to form homodimers, which is required for many NS1 functions. The Coxsackie B3 virus 2A protease inhibits host translation by inhibiting the same eIF4E pathway as 4EHP. By cleaving eIF4G - the binding partner of eIF4E - Coxsackie B3 virus

2A shuts off host translation. ISGylation of the 2A protease restores cellular translation by inhibiting cleavage of eIF4G(Rahnefeld et al., 2014). Another viral ISGylation target is the pUL26 protein from Human cytomegalovirus (HCMV). Active pUL26 suppresses NF-κB signaling, but ISG15 conjugation inhibits this activity and results in decreased HCMV replication(Kim et al., 2016).

ISG15 interferes with oligomerization of viral proteins

Many viral proteins rely on forming oligomers or complexes to perform their functions. ISGylation of these viral proteins as an antiviral strategy is particularly effective. Conjugated ISG15 causes a steric hindrance, preventing further oligomerization once an ISGylated protein is incorporated. Therefore, ISGylation of a relatively small percentage of viral proteins can achieve a dominant inhibitory effect(Durfee et al., 2010; Zhao et al., 2016). A well-studied example of this is the nucleoprotein (NP) of Influenza B virus (IBV). NP is a component of the IBV ribonucleoprotein (RNP), and is critical for synthesizing viral RNA(Zhao et al., 2016). Because the incorporation of a single ISGylated NP into an RNP oligomer results in the entire oligomer being non-functional, ISGylated NPs have been observed to have a dominant effect over unmodified NPs. ISGylation of IBV NP results in reduced viral RNA synthesis and ultimately reduced viral replication(Zhao et al., 2016). Similarly, ISGylation of 10-30 percent of human papillomavirus (HPV) L1 capsid proteins results in a 70 percent decrease in infectivity(Durfee et al., 2010).

ISGylation can potentially target proteins for degradation

It has been suggested that ISG15 may serve a similar or redundant role to Ub in the autophagosomal and proteasomal degradation pathways. When stimulated by type I IFN, both free ISG15 and ISGylated proteins localize with histone deacetylase 6 (HDAC6) and

Ub-binding protein p62(Nakashima et al., 2015). HDAC6 controls autophagosome maturation in the process of clearing ubiquitinated protein aggregates, and p62 acts as a major chaperone protein(Lee et al., 2010; Nakashima et al., 2015). p62 is also critical to the proteasomal degradation pathway(Liu et al., 2016). In the presence of the proteasome inhibitor MG132 p62-linked autophagy increases, as does co-localization between p62, HDAC6, and ISG15(Nakashima et al., 2015). The autophagy and proteasomal degradation pathways are utilized to degrade ISGylated and ubiquitinated proteins. They are effective at lowering levels of viral proteins as well as misfolded endogenous proteins. Infection by HIV-1 induces accumulation of misfolded tumor suppressor p53, an antiviral factor that is suspected to inhibit HIV-1 long terminal repeat promoter activity(Cooper et al., 2013; GENINI et al., 2001; Osei Kuffour et al., 2019). ISG15 conjugation to the misfolded p53 by HERC5 and TRIM25 leads to p53 degradation, and in the absence of ISG15 misfolded p53 accumulates and enhances HIV-1 replication(Huang et al., 2014; Osei Kuffour et al., 2019; Park et al., 2016). One example of ISGylation leading to p62-mediated degradation is found in the negative regulation of RIG-I signaling. LRRC25 binds to ISGylated RIG-I and promotes association with p62(Du et al., 2018). ISG15 has also been linked to increased basal and infection-induced autophagy during Listeria monocytogenes infection by modifying mTOR, WIPI2, AMBRA1, and RAB7(Zhang et al., 2019). Despite both Ub and ISG15 being associated with protein degradation, ISG15-Ub mixed chains are not degradation signals. Furthermore, ISG15 can conjugate to Lys-29 of Ub, and ISGylation of cellular ubiquitinated proteins appears to slow turnover(Fan et al., 2015).

Free ISG15 negatively regulates type I IFN signaling in humans

Despite the fact that ISG15 is generally considered an antiviral protein, ISG15 has also been described to negatively regulate type I IFN signaling, at various places in this pathway. Type I IFN signaling is critical for antiviral innate immune responses, but excessive IFN signaling can result in autoinflammatory pathogenesis(Tokarz et al., 2004; Zhang et al., 2015). To mitigate this the human JAK/STAT pathway is regulated through a negative feedback loop, in which ISG15 and USP18 play an important role. STAT2 recruits USP18 to IFNAR2. USP18 binds competitively at the IFNAR2-JAK1 binding site, displacing JAK1. In humans ISG15 binds to USP18, inhibiting SKP2-mediated ubiquitination, which would result in proteasomal degradation of USP18(Vuillier et al., 2019). By preventing USP18 degradation ISG15 stabilizes the interaction between IFNAR2 and USP18 (FIG. 2.1). Bound USP18 interferes with receptor dimerization and JAK activation, preventing the formation of ISGF3(Arimoto et al., 2017). This ultimately results in decreased expression of ISGs and moderating the IFN response. The role of ISG15 in this negative feedback regulation system appears to be specific to humans. In mice USP18 is a negative regulator of the JAK/STAT pathway but is not ISGylated to prevent degradation(Speer et al., 2016). This may account for part of the difference in ISG15 function between species(Speer et al., 2016). In addition, ISG15 acts as a negative regulator of RIG-I activation. As opposed to increasing MAVS and IRF3 activation as seen with ubiquitination, ISGylation of RIG-I CARDs results in decreased activation of these proteins and decreased production of type I IFNs(Kim et al., 2008).

ISG15 as an extracellular signaling molecule

In addition to contributing to intracellular immune responses, ISG15 also plays an important role in the extracellular immune response. ISG15 can initiate the production and secretion of a wide variety of antiviral and anti-bacterial factors such as type III IFNs, reactive oxygen species (ROS), and nitric oxide (NO), as well as act as an extracellular cytokine itself (FIG. 2.2). Type III IFNs are proinflammatory, but also upregulate ISG15 production in a similar manner to type I IFNs. The mechanism of ISG15 secretion or release is still uncertain, but it does not involve a hydrophobic leader protein and is not inhibited by blocking classical secretion mediated by the Endoplasmic Reticulum-Golgi pathway(D'Cunha et al., 1996). There is speculation that it could involve exosomal trafficking, or apoptosis(Campbell and Lenschow, 2013; Dos Santos and Mansur, 2017; Perng and Lenschow, 2018; Sun et al., 2016). Extracellular free ISG15 binds to lymphocyte function-associated antigen 1 (LFA-1), a surface receptor on dendritic cells (DCs), natural killer (NK) cells, T cells, and macrophages(Baldanta et al., 2017; Napolitano et al., 2018; Recht et al., 1991; Swaim et al., 2017; Yeung et al., 2018). Depending on the cell type this can trigger proliferation and maturation. Upon co-stimulation with IL-12, ISG15 binding to LFA-1 can initiate production of IFN-γ and IL-10 (FIG. 2.2)(Baldanta et al., 2017; Napolitano et al., 2018; Padovan et al., 2002; Recht et al., 1991; Swaim et al., 2017; Yeung et al., 2018).

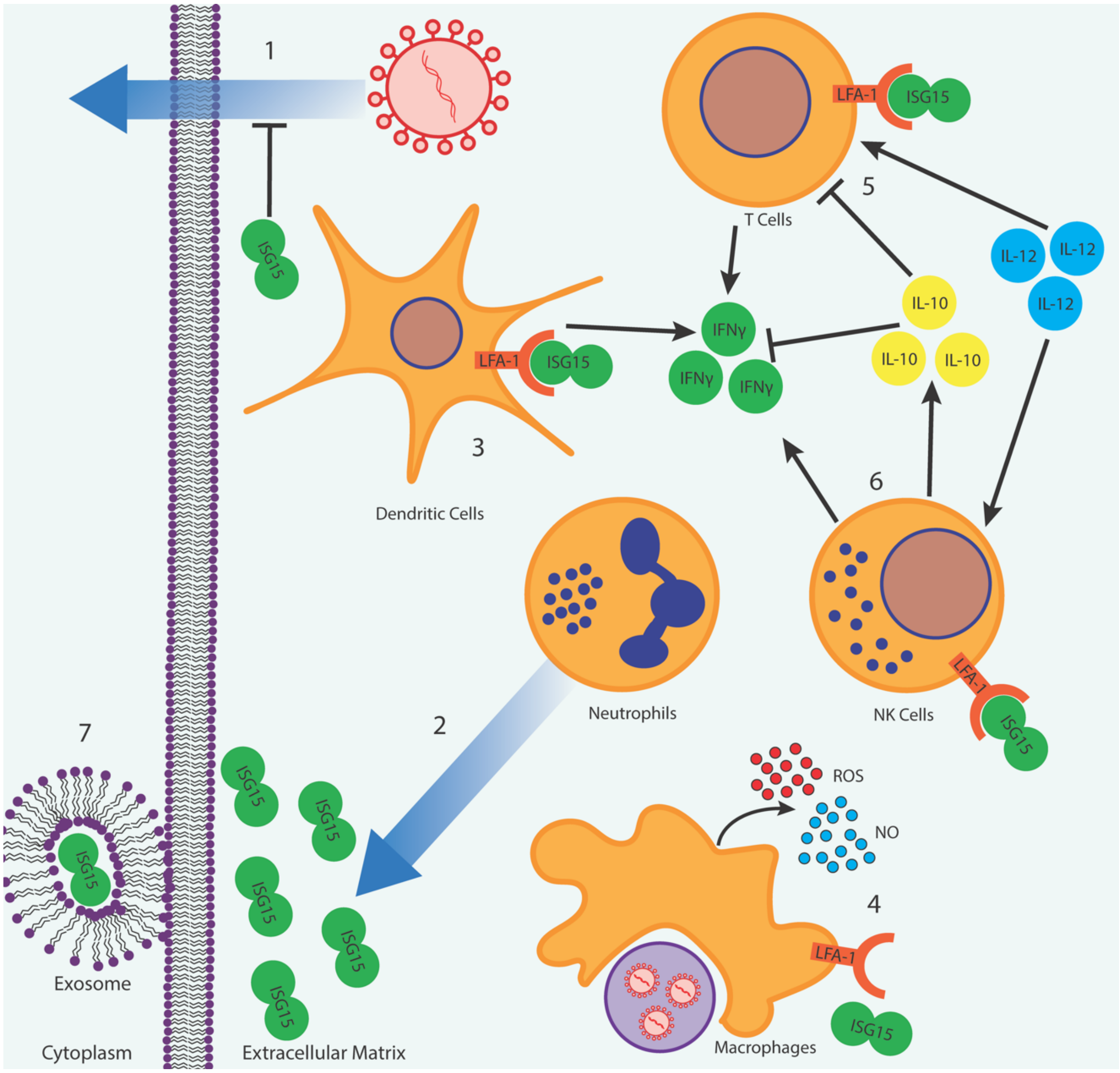


Figure 2.2. Extracellular ISG15 directly affects pathogenesis or activates various immune cell types. (1) It inhibits the virus replication at the earliest stages, potentially by preventing entry into the cell. (2) It is an effective chemoattractant for neutrophils. It binds to the LFA-1 receptor on dendritic cells, NK cells, and T cells. This receptor is present in macrophages but has yet to be confirmed as the mechanism by which ISG15 activates macrophages. (3) In dendritic cells it initiates maturation and IFN-γ production. (4) In macrophages it causes polarization to the M1 phenotype, resulting in production of reactive oxygen species and nitric oxide. It also increased autophagy and mitophagy of infected cells and organelles. (5) In T cells and NK cells upon co-stimulation with IL-12 it stimulates production of IFN-γ. (6) In NK cells it also stimulates production of IL-10, which inhibits IFN-γ production in T cells. (7) Exosomal trafficking is one of the proposed methods by which ISG15 may be secreted from a cell but that mechanism is still unclear.

ISG15 induces an antiviral state in immune cells

ISG15 is a potent activator and recruiter of immune cells. In the presence of ISG15 macrophages increase in polarization towards the proinflammatory and antiviral M1 phenotype. M1 macrophages display higher production of antiviral factors such as ROS and NO(FIG. 2.2). In addition, ISG15-stimulated macrophages display increased autophagy and mitophagy of infected cells and organelles (Baldanta et al., 2017). The mechanism of by which ISG15 induces M1 polarization in macrophages is currently unknown, but the expression of LFA-1 receptors by macrophages suggest it may be through a similar mechanism to NK cells.

NK cells respond to ISG15 binding LFA-1 by proliferating and increasing production of IFN-γ and IL-10(Swaim et al., 2017). Increased production of the pro-inflammatory cytokine IFN-γ is a common effect of activation by ISG15, occurring in NK cells, DCs, and T cells(Napolitano et al., 2018; Recht et al., 1991; Swaim et al., 2017). Increased production of IL-10 in response to ISG15 signaling appears to be negative feedback inhibition to prevent damage from an overactive inflammation response. IL-10 inhibits T cell activation and downregulates proinflammatory cytokine production(Couper et al., 2008). In CD8+ T cells ISG15 binding increases activation(Yeung et al., 2018). Both CD4+ and CD8+ T cells increase IFN-γ production in the presence of free ISG15 but only after being primed with IL-2(Recht et al., 1991).

ISG15 causes CD8α+ DCs to increase production of IL-1β and IFN-γ as well(Napolitano et al., 2018). Additionally, the presence of free ISG15 induces dendritic cell phenotypic maturation. Incubation with ISG15 in growth media induced production of E-cadherin, CD15, and CD86 by dendritic cells. Expression of these genes is typically associated with a mature phenotype, and CD86 acts as a co-stimulatory signal to activate T cells. The presence of anti-ISG15 antibodies in the media nullified these effects(Padovan et al., 2002).

Aside from binding to the LFA-1 receptor, free ISG15 can act as a chemoattractant for neutrophils(Owhashi et al., 2003). Because ISG15 is highly concentrated at sites of infection and apoptotic cells it is effective in recruiting neutrophils to areas of need. When left unchecked this elevated proinflammatory state can cause damage. Therefore, the IFN regulatory activity of USP18 is critically important to maintain signaling levels under normal conditions and restoring homeostasis after an infection. By inhibiting the JAK/STAT pathway and suppressing the downstream effects of type I IFN signaling USP18 effectively reduces auto-inflammatory pathogenesis. The absence of USP18 in humans has been observed to result in severe interferonopathy(Gruber et al., 2020; Zhang et al., 2015).

Interference with viral cell entry

In addition to activating immune cells and extracellular pathways, ISG15 can interfere with viral infection at its earliest stages. ISG15 interferes with Norovirus entry or uncoating. ISG15 is a known inhibitor of Norovirus replication and this inhibition occurs upstream of virus transcription(Rodriguez et al., 2014). It has been proposed that ISG15 reduces Zika virus (ZIKV) infection by preventing viral entry as well. The presence of free extracellular ISG15 reduces intracellular viral titers during ZIKV infection (Singh et al., 2019).

Viral countermeasures to ISG15 signaling

Various viruses have developed strategies to circumvent ISG15 interference. Most strategies to suppress ISG15 signaling principally appear to fall along two lines, viral protease mediated deISGylating and sequestering ISGylated proteins. However, some viruses have additional mechanisms that allow them to suppress ISG15 antiviral activities.

Viral deISGylases

The initial suggestion that viruses could be encoding proteases with deubiquitinase function (DUBs) to counter innate immune responses came from structural similarities between papain-like proteases (PLP) from coronaviruses to host USPs such as HAUSP(Barretto et al., 2005; Sulea et al., 2005). Shortly thereafter, these proteases were among the first to be found to be multifunctional in their ability to reverse ubiquitination, as well as ISGylation(Lindner et al., 2005). Prominent examples of coronaviruses that encode PLPs possessing deISGylase activity include Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) and Middle East Respiratory Syndrome coronavirus (MERS-CoV), as well as the lesser-known mouse hepatitis virus (MHV)(Mielech et al., 2015; Mielech et al., 2014; Ratia et al., 2014). However, the full breath of PLPs from coronaviruses that can robustly process ISG15 is currently unknown and has been somewhat impeded by the revelation that viral PLPs are specific for certain host's ISG15(Daczkowski et al., 2017a). In contrast to the highly conserved Ub, ISG15 displays significantly more sequence variation between host species(Dzimianski et al., 2019a). Sequence variation within the C-terminal β-grasp Ubl domain, or differences that impact the domain-domain interactions within ISG15 have been observed to impact the ability of coronavirus PLPs to engage with ISG15(Daczkowski et al., 2017a; Langley et al., 2019).

This phenomenon has been suggested to occur because PLPs are likely specifically adapted to cleave the ISG15 of their predominate host reservoirs, but lack deISGylase activity in other species(Langley et al., 2019). As unclear about how many coronaviruses encode for PLPs that have deISGylase activity, so is the exact role and impact viral deISGylase activity plays in modulating host immune function. The combination of both deISGylase and DUB activity in coronaviruses have been observed to contribute to the suppression of the innate immune response by acting in part on IFN-β and NF-κB signaling pathways(Bailey-Elkin et al., 2014; Clementz et al., 2010; Mielech et al., 2014; Ratia et al., 2014). Yet the individual contribution and role of viral deISGylase activity remains unclear. However, the recent development of several new molecular tools in the form of altered MERS-CoV PLPs that lack DUB, deISGylase, or both activities may help elucidate the exact role of ISG15 during CoV infection(Clasman et al., 2019; Daczkowski et al., 2017b).

While the exact contribution of PLP deISGylase activity on immune suppression remains unknown, this is not entirely the case with the other well-studied family of viral deISGylases, viral ovarian tumor (OTU) domain proteases. Viral deISGylases belonging to the OTU superfamily are encoded by multiple viral families, including Nairoviruses, arteriviruses and tenuiviruses(Bester et al., 2018; Frias-Staheli et al., 2007; Zhang et al., 2007). Like PLPs, OTUs in general have shown potential to possess the ability to reverse ISGylation in addition to a DUB role(Frias-Staheli et al., 2007). The Erve Nairovirus even possesses an OTU domain containing deISGylase activity, but lacking DUB activity, underscoring the importance of viral deISGylase function for particular viruses(Deaton et al., 2016). Similar to viral PLPs, viral OTUs have shown to be specific for ISG15 from certain species, including those they productively infect(Dzimianski et al., 2019b).

Specific mutations within the OTU domain from the Crimean-Congo hemorrhagic fever (CCHFV) Nairovirus that silenced DUB activity, or both DUB and deISGylase activities allowed narrowing down the mechanism by which viral OTU deISGylase function manipulates host immunity. This has suggested a distinct function for viral deISGylase activity for Nairoviruses. Whereas DUB activity was shown to downregulate type I IFN signaling, OTU deISGylase activity appeared to stabilize CCHFV L-protein levels at a later stage in infection(Scholte et al., 2017).

Sequestering proteins

Another countermeasure viruses employ against ISG15 signaling is the production of proteins that sequester ISG15 and ISGylated proteins. The Vaccinia virus E3 protein sequesters free ISG15(Guerra et al., 2008; Smith et al., 2013). By reducing unconjugated ISG15 Vaccinia E3 prevents interference with viral proteins or propagation of antiviral signals. NS1B from IBV on the other hand sequesters ISGylated proteins in a species dependent manner, particularly ISGylated viral proteins(Zhao et al., 2016). This species dependent behavior has been observed to influence the replication of IBV in humans over their mouse and other animal counterparts(Jiang and Wang, 2019; Sridharan et al., 2010). As previously stated, ISGylation of viral proteins that need to form complexes to function is a particularly effective method of viral suppression due to the relatively small percentage of proteins that need to be ISGylated to elicit an antiviral effect. By sequestering ISGylated components of these complexes IBV drastically reduces the efficacy of that method of viral suppression (FIG. 2.3). For both of these viruses expression of their respective sequestering proteins appears to be critical to preventing viral clearance(Guerra et al., 2008; Zhao et al., 2016).

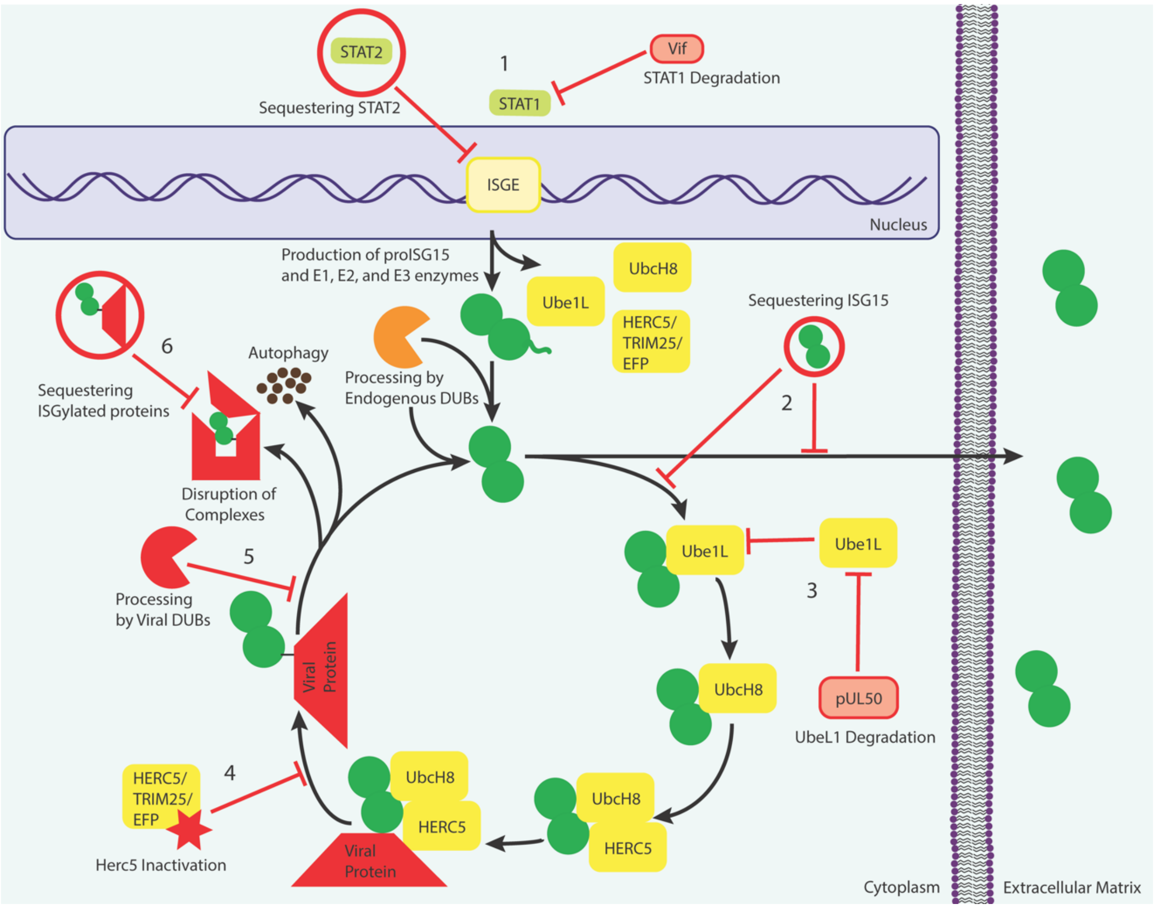


Figure 2.3. Viral countermeasures to ISG15 production and activity. (1) The Vif protein from HIV-1 degrades STAT1 and IE1 from HCMV sequesters STAT2, both of which are critical to upregulating ISG15 synthesis. (2) Vaccinia virus E3 protein sequesters free ISG15, preventing it from conjugating to targets or acting as a signaling protein. (3) pUL50 of HCMV binds to and causes proteasomal degradation of Ube1L, an essential E1 activator protein that facilitates ISGylation. (4) KSHV and HCMV both produce proteins that interfere with the HERC5 E3 ligase. Both reduce ISG15 conjugation to target proteins. (5) Viral DUBs and deISGylases such as OTUs and PLPs cleave the conjugation between ISG15 and target proteins, returning ISGylated proteins to their normal state. (6) NS1B from IBV sequesters ISGylated proteins, preventing them from being incorporated into oligomers.

Inhibition of ISG15 production and conjugation

Some other less common methods viruses employ to counteract ISG15 activity include inhibition of ISG15 conjugation to targets and interfering with the ISG15 synthesis and conjugation pathways (FIG. 2.3). Kaposi's sarcoma-associated herpesvirus (KSHV) and HCMV both encode proteins that interfere with the HERC5 E3 ligase(Jacobs et al., 2015; Kim et al., 2016). HCMV also encodes pUL50, a protein that binds to Ube1L leading to proteasomal degradation, resulting in reduced ISG15 activation(Lee et al., 2018). In addition, HCMV also produces an immediate-early protein 1(IE1) that reduces overall ISG15 production(Bianco and Mohr, 2017; Kim et al., 2016). It does so by sequestering STAT2, preventing formation of the ISGF3 complex and inhibiting ISRE activation(Huh et al., 2008; Krauss et al., 2009; Paulus et al., 2006). This reduces not only the amount of ISG15 available, but also the amount of ISG15 conjugating enzymes (Ube1L, UbcH8, HERC5). HIV-1 is capable of degrading components of the JAK/STAT pathway to reduce the effectiveness of type 1 IFN signaling(Gargan et al., 2018). The HIV-1 protein Vif inhibits IFN-α signaling by degrading STAT1 and STAT3. By reducing STAT1 HIV-1 effectively reduces the amount of ISG15 produced to fight infection.

Conclusions

ISG15 is a functionally versatile Ubl protein that affects many aspects of the antiviral immune response. It is important as both a signaling molecule and effector protein. Intracellularly, ISG15 interferes with many viral processes directly, and is involved in many antiviral-signaling pathways. Extracellularly, ISG15 can directly interact with viruses to prevent infection, as well as activate various immune cells to promote viral clearance. ISG15 can also promote the production of many extracellular antiviral cytokines. Aligning with ISG15 being one of the most highly upregulated proteins during viral pathogenesis, not surprising the information known to date on ISG15 highlights it as having central role in myriad of host immune mechanisms. Mechanisms that viruses from diverse families seek to stymie in different ways to favor virulence.

Surprisingly, what is known about ISG15 and its immunological role is likely still in its infancy. While some mechanisms such as the interplay between ISG15 and IBV are well understood, others such as how ISGylation targets are determined and which viruses have evolved to target ISG15 activity over Ub activity remain unclear. Despite ISG15's importance as an extracellular signaling molecule the process by which ISG15 is secreted is still largely unknown. Additionally, how 4EHP and ISG15 work synergistically to downregulate viral mRNAs over host mRNAs translation is limited. This limited knowledge also extends to what degree species to species variation within ISG15s contribute to protection from viruses spilling over into another host. Already ISG15 species-species variation appears to impact the effectiveness of immune evasion mechanisms of IBV, Nairoviruses, and coronaviruses. Then there are the species-to-species variations in how hosts utilize ISG15 as a regulator when it comes to USP18. How all of these species-species differences translate to disease outcomes is not well understood. So as much as many things are known about ISG15's roles, there seems to be as much yet to be revealed. Taking into account that numerous host and viral pathways that have been already been identified that could lend themselves to therapeutic intervention, the benefit of filling in the knowledge gaps regarding ISG15 should not be underestimated.

References

Arimoto, K.I., Lochte, S., Stoner, S.A., Burkart, C., Zhang, Y., Miyauchi, S., Wilmes, S., Fan, J.B., Heinisch, J.J., Li, Z., Yan, M., Pellegrini, S., Colland, F., Piehler, J., Zhang, D.E., 2017. STAT2 is an essential adaptor in USP18-mediated suppression of type I interferon signaling. Nat Struct Mol Biol 24(3), 279-289.

Bailey-Elkin, B.A., Knaap, R.C., Johnson, G.G., Dalebout, T.J., Ninaber, D.K., van Kasteren, P.B., Bredenbeek, P.J., Snijder, E.J., Kikkert, M., Mark, B.L., 2014. Crystal structure of the Middle East respiratory syndrome coronavirus (MERS-CoV) papain-like protease bound to ubiquitin facilitates targeted disruption of deubiquitinating activity to demonstrate its role in innate immune suppression. J Biol Chem 289(50), 34667-34682.

Baldanta, S., Fernandez-Escobar, M., Acin-Perez, R., Albert, M., Camafeita, E., Jorge, I., Vazquez, J., Enriquez, J.A., Guerra, S., 2017. ISG15 governs mitochondrial function in macrophages following vaccinia virus infection. PLoS Pathog 13(10), e1006651.

Barretto, N., Jukneliene, D., Ratia, K., Chen, Z., Mesecar, A.D., Baker, S.C., 2005. The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. J Virol 79(24), 15189-15198.

Bergmann, M., Garcia-Sastre, A., Carnero, E., Pehamberger, H., Wolff, K., Palese, P., Muster, T., 2000. Influenza Virus NS1 Protein Counteracts PKR-Mediated Inhibition of Replication. Journal of Virology 74(13), 6203-6206.

Berndsen, C.E., Wolberger, C., 2014. New insights into ubiquitin E3 ligase mechanism. Nature Structural & Molecular Biology 21(4), 301-307.

Bester, S.M., Daczkowski, C.M., Faaberg, K.S., Pegan, S.D., 2018. Insights into the Porcine Reproductive and Respiratory Syndrome Virus Viral Ovarian Tumor Domain Protease Specificity for Ubiquitin and Interferon Stimulated Gene Product 15. ACS Infect Dis 4(9), 1316-1326.

Bianco, C., Mohr, I., 2017. Restriction of Human Cytomegalovirus Replication by ISG15, a Host Effector Regulated by cGAS-STING Double-Stranded-DNA Sensing. Journal of virology 91(9), e02483-02416.

Bogunovic, D., Byun, M., Durfee, L.A., Abhyankar, A., Sanal, O., Mansouri, D., Salem, S., Radovanovic, I., Grant, A.V., Adimi, P., Mansouri, N., Okada, S., Bryant, V.L., Kong, X.-F., Kreins, A., Velez, M.M., Boisson, B., Khalilzadeh, S., Ozcelik, U., Darazam, I.A., Schoggins, J.W., Rice, C.M., Al-Muhsen, S., Behr, M., Vogt, G., Puel, A., Bustamante, J., Gros, P., Huibregtse, J.M., Abel, L., Boisson-Dupuis, S., Casanova, J.-L., 2012. Mycobacterial Disease and Impaired IFN-γ Immunity in Humans with Inherited ISG15 Deficiency. Science 337(6102), 1684-1688.

Campbell, J.A., Lenschow, D.J., 2013. Emerging roles for immunomodulatory functions of free ISG15. J Interferon Cytokine Res 33(12), 728-738.

Clasman, J.R., Everett, R.K., Srinivasan, K., Mesecar, A.D., 2019. Decoupling deISGylating and deubiquitinating activities of the MERS virus papain-like protease. Antiviral Res 174, 104661.

Clementz, M.A., Chen, Z., Banach, B.S., Wang, Y., Sun, L., Ratia, K., Baez-Santos, Y.M., Wang, J., Takayama, J., Ghosh, A.K., Li, K., Mesecar, A.D., Baker, S.C., 2010. Deubiquitinating and interferon antagonism activities of coronavirus papain-like proteases. J Virol 84(9), 4619-4629.

Cooper, A., García, M., Petrovas, C., Yamamoto, T., Koup, R.A., Nabel, G.J., 2013. HIV-1 causes CD4 cell death through DNA-dependent protein kinase during viral integration. Nature 498(7454), 376-379.

Couper, K.N., Blount, D.G., Riley, E.M., 2008. IL-10: The Master Regulator of Immunity to Infection. The Journal of Immunology 180(9), 5771-5777.

D'Cunha, J., Ramanujam, S., Wagner, R.J., Witt, P.L., Knight, E., Jr., Borden, E.C., 1996. In vitro and in vivo secretion of human ISG15, an IFN-induced immunomodulatory cytokine. J Immunol 157(9), 4100-4108.

Daczkowski, C.M., Dzimianski, J.V., Clasman, J.R., Goodwin, O., Mesecar, A.D., Pegan, S.D., 2017a. Structural Insights into the Interaction of Coronavirus Papain-Like Proteases and Interferon-Stimulated Gene Product 15 from Different Species. J Mol Biol 429(11), 1661-1683.

Daczkowski, C.M., Goodwin, O.Y., Dzimianski, J.V., Farhat, J.J., Pegan, S.D., 2017b. Structurally Guided Removal of DeISGylase Biochemical Activity from Papain-Like Protease Originating from Middle East Respiratory Syndrome Coronavirus. J Virol 91(23).
de la Luna, S., Fortes, P., Beloso, A., Ortín, J., 1995. Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs. Journal of Virology 69(4), 2427-2433.

Deaton, M.K., Dzimianski, J.V., Daczkowski, C.M., Whitney, G.K., Mank, N.J., Parham, M.M., Bergeron, E., Pegan, S.D., 2016. Biochemical and Structural Insights into the Preference of Nairoviral DeISGylases for Interferon-Stimulated Gene Product 15 Originating from Certain Species. J Virol 90(18), 8314-8327.

Der, S.D., Zhou, A., Williams, B.R.G., Silverman, R.H., 1998. Identification of genes differentially regulated by interferon α, β, or γ using oligonucleotide arrays. Proceedings of the National Academy of Sciences 95(26), 15623.

Dos Santos, P.F., Mansur, D.S., 2017. Beyond ISGlylation: Functions of Free Intracellular and Extracellular ISG15. J Interferon Cytokine Res 37(6), 246-253.

Du, Y., Duan, T., Feng, Y., Liu, Q., Lin, M., Cui, J., Wang, R.F., 2018. LRRC25 inhibits type I IFN signaling by targeting ISG15-associated RIG-I for autophagic degradation. EMBO J 37(3), 351-366.

Durfee, L.A., Lyon, N., Seo, K., Huibregtse, J.M., 2010. The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. Mol Cell 38(5), 722-732.

Dzimianski, J.V., Scholte, F.E.M., Bergeron, E., Pegan, S.D., 2019a. ISG15: It's Complicated. J Mol Biol 431(21), 4203-4216.

Dzimianski, J.V., Scholte, F.E.M., Williams, I.L., Langley, C., Freitas, B.T., Spengler, J.R., Bergeron, E., Pegan, S.D., 2019b. Determining the molecular drivers of species-specific interferon-stimulated gene product 15 interactions with nairovirus ovarian tumor domain proteases. PLoS One 14(12), e0226415.

Fan, J.B., Arimoto, K., Motamedchaboki, K., Yan, M., Wolf, D.A., Zhang, D.E., 2015. Identification and characterization of a novel ISG15-ubiquitin mixed chain and its role in regulating protein homeostasis. Sci Rep 5, 12704.

Farrell, P.J., Broeze, R.J., Lengyel, P., 1979. Accumulation of an mRNA and protein in interferon-treated Ehrlich ascites tumour cells. Nature 279(5713), 523-525.

Frias-Staheli, N., Giannakopoulos, N.V., Kikkert, M., Taylor, S.L., Bridgen, A., Paragas, J., Richt, J.A., Rowland, R.R., Schmaljohn, C.S., Lenschow, D.J., Snijder, E.J., Garcia-Sastre, A., Virgin, H.W.t., 2007. Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. Cell Host Microbe 2(6), 404-416.

Gargan, S., Ahmed, S., Mahony, R., Bannan, C., Napoletano, S., O'Farrelly, C., Borrow, P., Bergin, C., Stevenson, N.J., 2018. HIV-1 Promotes the Degradation of Components of the Type 1 IFN JAK/STAT Pathway and Blocks Anti-viral ISG Induction. EBioMedicine 30, 203-216.

GENINI, D., SHEETER, D., ROUGHT, S., ZAUNDERS, J.J., SUSIN, S.A., KROEMER, G., RICHMAN, D.D., CARSON, D.A., CORBEIL, J., LEONI, L.M., 2001. HIV induces lymphocyte apoptosis by a p53-initiated, mitochondrial-mediated mechanism. The FASEB Journal 15(1), 5-6.

González-Sanz, R., Mata, M., Bermejo-Martín, J., Álvarez, A., Cortijo, J., Melero, J.A., Martínez, I., 2016. ISG15 Is Upregulated in Respiratory Syncytial Virus Infection and Reduces Virus Growth through Protein ISGylation. Journal of Virology 90(7), 3428-3438.
Grandvaux, N., Servant, M.J., tenOever, B., Sen, G.C., Balachandran, S., Barber, G.N., Lin, R., Hiscott, J., 2002. Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. J Virol 76(11), 5532-5539.

Gruber, C., Martin-Fernandez, M., Ailal, F., Qiu, X., Taft, J., Altman, J., Rosain, J., Buta, S., Bousfiha, A., Casanova, J.L., Bustamante, J., Bogunovic, D., 2020. Homozygous STAT2 gain-of-function mutation by loss of USP18 activity in a patient with type I interferonopathy. J Exp Med 217(5).

Guerra, S., Cáceres, A., Knobeloch, K.-P., Horak, I., Esteban, M., 2008. Vaccinia virus E3 protein prevents the antiviral action of ISG15. PLoS pathogens 4(7), e1000096-e1000096.

Haas, A.L., Ahrens, P., Bright, P.M., Ankel, H., 1987. Interferon induces a 15-kilodalton protein exhibiting marked homology to ubiquitin. J Biol Chem 262(23), 11315-11323.

Huang, Y.-F., Wee, S., Gunaratne, J., Lane, D.P., Bulavin, D.V., 2014. Isg15 controls p53 stability and functions. Cell Cycle 13(14), 2199-2209.

Huh, Y.H., Kim, Y.E., Kim, E.T., Park, J.J., Song, M.J., Zhu, H., Hayward, G.S., Ahn, J.-H., 2008. Binding STAT2 by the Acidic Domain of Human Cytomegalovirus IE1 Promotes Viral Growth and Is Negatively Regulated by SUMO. Journal of Virology 82(21), 10444.
Jacobs, S.R., Stopford, C.M., West, J.A., Bennett, C.L., Giffin, L., Damania, B., 2015. Kaposi's Sarcoma-Associated Herpesvirus Viral Interferon Regulatory Factor 1 Interacts with a Member of the Interferon-Stimulated Gene 15 Pathway. J Virol 89(22), 11572-11583.

Jiang, Y., Wang, X., 2019. Structural insights into the species preference of the influenza B virus NS1 protein in ISG15 binding. Protein Cell 10(9), 681-687.

Kalvakolanu, D.V., Borden, E.C., 1996. An overview of the interferon system: signal transduction and mechanisms of action. Cancer Invest 14(1), 25-53.

Kim, M.J., Hwang, S.Y., Imaizumi, T., Yoo, J.Y., 2008. Negative feedback regulation of RIG-I-mediated antiviral signaling by interferon-induced ISG15 conjugation. J Virol 82(3), 1474-1483.

Kim, Y.J., Kim, E.T., Kim, Y.E., Lee, M.K., Kwon, K.M., Kim, K.I., Stamminger, T., Ahn, J.H., 2016. Consecutive Inhibition of ISG15 Expression and ISGylation by Cytomegalovirus Regulators. PLoS Pathog 12(8), e1005850.

Krauss, S., Kaps, J., Czech, N., Paulus, C., Nevels, M., 2009. Physical Requirements and Functional Consequences of Complex Formation between the Cytomegalovirus IE1 Protein and Human STAT2. Journal of Virology 83(24), 12854.

Langley, C., Goodwin, O., Dzimianski, J.V., Daczkowski, C.M., Pegan, S.D., 2019. Structure of interferon-stimulated gene product 15 (ISG15) from the bat species Myotis davidii and the impact of interdomain ISG15 interactions on viral protein engagement. Acta Crystallogr D Struct Biol 75(Pt 1), 21-31.

Lee, J.-Y., Koga, H., Kawaguchi, Y., Tang, W., Wong, E., Gao, Y.-S., Pandey, U.B., Kaushik, S., Tresse, E., Lu, J., Taylor, J.P., Cuervo, A.M., Yao, T.-P., 2010. HDAC6 controls autophagosome maturation essential for ubiquitin-selective quality-control autophagy. EMBO J 29(5), 969-980.

Lee, M.K., Kim, Y.J., Kim, Y.-E., Han, T.-H., Milbradt, J., Marschall, M., Ahn, J.-H., 2018. Transmembrane Protein pUL50 of Human Cytomegalovirus Inhibits ISGylation by Downregulating UBE1L. Journal of Virology 92(15), e00462-00418.

Lenschow, D.J., Lai, C., Frias-Staheli, N., Giannakopoulos, N.V., Lutz, A., Wolff, T., Osiak, A., Levine, B., Schmidt, R.E., Garcia-Sastre, A., Leib, D.A., Pekosz, A., Knobeloch, K.P., Horak, I., Virgin, H.W.t., 2007. IFN-stimulated gene 15 functions as a critical antiviral molecule against influenza, herpes, and Sindbis viruses. Proc Natl Acad Sci U S A 104(4), 1371-1376.

Lindner, H.A., Fotouhi-Ardakani, N., Lytvyn, V., Lachance, P., Sulea, T., Menard, R., 2005. The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme. J Virol 79(24), 15199-15208.

Liu, W.J., Ye, L., Huang, W.F., Guo, L.J., Xu, Z.G., Wu, H.L., Yang, C., Liu, H.F., 2016. p62 links the autophagy pathway and the ubiqutin-proteasome system upon ubiquitinated protein degradation. Cell Mol Biol Lett 21, 29-29.

Loeb, K.R., Haas, A.L., 1992. The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. J Biol Chem 267(11), 7806-7813.

Malakhova, O.A., Zhang, D.E., 2008. ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. J Biol Chem 283(14), 8783-8787.

Meraro, D., Gleit-Kielmanowicz, M., Hauser, H., Levi, B.Z., 2002. IFN-stimulated gene 15 is synergistically activated through interactions between the myelocyte/lymphocyte-specific transcription factors, PU.1, IFN regulatory factor-8/IFN consensus sequence binding protein, and IFN regulatory factor-4: characterization of a new subtype of IFN-stimulated response element. J Immunol 168(12), 6224-6231.

Mielech, A.M., Deng, X., Chen, Y., Kindler, E., Wheeler, D.L., Mesecar, A.D., Thiel, V., Perlman, S., Baker, S.C., 2015. Murine coronavirus ubiquitin-like domain is important for papain-like protease stability and viral pathogenesis. J Virol 89(9), 4907-4917.

Mielech, A.M., Kilianski, A., Baez-Santos, Y.M., Mesecar, A.D., Baker, S.C., 2014. MERS-CoV papain-like protease has deISGylating and deubiquitinating activities. Virology 450-451, 64-70.

Min, J.-Y., Krug, R.M., 2006. The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2′-5′ oligo (A) synthetase/RNase L pathway. Proceedings of the National Academy of Sciences 103(18), 7100-7105.

Nakashima, H., Nguyen, T., Goins, W.F., Chiocca, E.A., 2015. Interferon-stimulated gene 15 (ISG15) and ISG15-linked proteins can associate with members of the selective autophagic process, histone deacetylase 6 (HDAC6) and SQSTM1/p62. The Journal of biological chemistry 290(3), 1485-1495.

Napolitano, A., van der Veen, A.G., Bunyan, M., Borg, A., Frith, D., Howell, S., Kjaer, S., Beling, A., Snijders, A.P., Knobeloch, K.-P., Frickel, E.-M., 2018. Cysteine-Reactive Free ISG15 Generates IL-1β-Producing CD8α(+) Dendritic Cells at the Site of Infection. Journal of immunology (Baltimore, Md. : 1950) 201(2), 604-614.

Narasimhan, J., Wang, M., Fu, Z., Klein, J.M., Haas, A.L., Kim, J.J., 2005. Crystal structure of the interferon-induced ubiquitin-like protein ISG15. J Biol Chem 280(29), 27356-27365.

Okumura, A., Pitha, P.M., Harty, R.N., 2008. ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. Proc Natl Acad Sci U S A 105(10), 3974-3979.

Okumura, F., Zou, W., Zhang, D.E., 2007. ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. Genes Dev 21(3), 255-260.

Osei Kuffour, E., König, R., Häussinger, D., Schulz, W.A., Münk, C., 2019. ISG15 Deficiency Enhances HIV-1 Infection by Accumulating Misfolded p53. mBio 10(4), e01342-01319.

Owhashi, M., Taoka, Y., Ishii, K., Nakazawa, S., Uemura, H., Kambara, H., 2003. Identification of a ubiquitin family protein as a novel neutrophil chemotactic factor. Biochem Biophys Res Commun 309(3), 533-539.

Padovan, E., Terracciano, L., Certa, U., Jacobs, B., Reschner, A., Bolli, M., Spagnoli, G.C., Borden, E.C., Heberer, M., 2002. Interferon stimulated gene 15 constitutively produced by melanoma cells induces e-cadherin expression on human dendritic cells. Cancer Res 62(12), 3453-3458.

Park, J.H., Yang, S.W., Park, J.M., Ka, S.H., Kim, J.-H., Kong, Y.-Y., Jeon, Y.J., Seol, J.H., Chung, C.H., 2016. Positive feedback regulation of p53 transactivity by DNA damage-induced ISG15 modification. Nature Communications 7(1), 12513.

Paulus, C., Krauss, S., Nevels, M., 2006. A human cytomegalovirus antagonist of type I IFN-dependent signal transducer and activator of transcription signaling. Proc Natl Acad Sci U S A 103(10), 3840-3845.

Perng, Y.-C., Lenschow, D.J., 2018. ISG15 in antiviral immunity and beyond. Nature Reviews Microbiology 16(7), 423-439.

Pincetic, A., Kuang, Z., Seo, E.J., Leis, J., 2010. The interferon-induced gene ISG15 blocks retrovirus release from cells late in the budding process. Journal of virology 84(9), 4725-4736.

Potter, J.L., Narasimhan, J., Mende-Mueller, L., Haas, A.L., 1999. Precursor processing of pro-ISG15/UCRP, an interferon-beta-induced ubiquitin-like protein. J Biol Chem 274(35), 25061-25068.

Rahnefeld, A., Klingel, K., Schuermann, A., Diny, N.L., Althof, N., Lindner, A., Bleienheuft, P., Savvatis, K., Respondek, D., Opitz, E., Ketscher, L., Sauter, M., Seifert, U., Tschope, C., Poller, W., Knobeloch, K.P., Voigt, A., 2014. Ubiquitin-like protein ISG15 (interferon-stimulated gene of 15 kDa) in host defense against heart failure in a mouse model of virus-induced cardiomyopathy. Circulation 130(18), 1589-1600.

Ratia, K., Kilianski, A., Baez-Santos, Y.M., Baker, S.C., Mesecar, A., 2014. Structural Basis for the Ubiquitin-Linkage Specificity and deISGylating activity of SARS-CoV papain-like protease. PLoS Pathog 10(5), e1004113.

Recht, M., Borden, E.C., Knight, E., Jr., 1991. A human 15-kDa IFN-induced protein induces the secretion of IFN-gamma. J Immunol 147(8), 2617-2623.

Rodriguez, M.R., Monte, K., Thackray, L.B., Lenschow, D.J., 2014. ISG15 functions as an interferon-mediated antiviral effector early in the murine norovirus life cycle. Journal of virology 88(16), 9277-9286.

Scholte, F.E.M., Zivcec, M., Dzimianski, J.V., Deaton, M.K., Spengler, J.R., Welch, S.R., Nichol, S.T., Pegan, S.D., Spiropoulou, C.F., Bergeron, E., 2017. Crimean-Congo Hemorrhagic Fever Virus Suppresses Innate Immune Responses via a Ubiquitin and ISG15 Specific Protease. Cell Rep 20(10), 2396-2407.

Seo, E.J., Leis, J., 2012. Budding of Enveloped Viruses: Interferon-Induced ISG15-Antivirus Mechanisms Targeting the Release Process. Adv Virol 2012, 532723-532723.
Singh, P.K., Singh, S., Farr, D., Kumar, A., 2019. Interferon-stimulated gene 15 (ISG15) restricts Zika virus replication in primary human corneal epithelial cells. The Ocular Surface 17(3), 551-559.

Smith, G.L., Benfield, C.T., Maluquer de Motes, C., Mazzon, M., Ember, S.W., Ferguson, B.J., Sumner, R.P., 2013. Vaccinia virus immune evasion: mechanisms, virulence and immunogenicity. J Gen Virol 94(Pt 11), 2367-2392.

Speer, S.D., Li, Z., Buta, S., Payelle-Brogard, B., Qian, L., Vigant, F., Rubino, E., Gardner, T.J., Wedeking, T., Hermann, M., Duehr, J., Sanal, O., Tezcan, I., Mansouri, N., Tabarsi, P., Mansouri, D., Francois-Newton, V., Daussy, C.F., Rodriguez, M.R., Lenschow, D.J., Freiberg, A.N., Tortorella, D., Piehler, J., Lee, B., Garcia-Sastre, A., Pellegrini, S., Bogunovic, D., 2016. ISG15 deficiency and increased viral resistance in humans but not mice. Nat Commun 7, 11496.

Sridharan, H., Zhao, C., Krug, R.M., 2010. Species specificity of the NS1 protein of influenza B virus: NS1 binds only human and non-human primate ubiquitin-like ISG15 proteins. J Biol Chem 285(11), 7852-7856.

Sulea, T., Lindner, H.A., Purisima, E.O., Menard, R., 2005. Deubiquitination, a new function of the severe acute respiratory syndrome coronavirus papain-like protease? J Virol 79(7), 4550-4551.

Sun, L., Wang, X., Zhou, Y., Zhou, R.H., Ho, W.Z., Li, J.L., 2016. Exosomes contribute to the transmission of anti-HIV activity from TLR3-activated brain microvascular endothelial cells to macrophages. Antiviral Res 134, 167-171.

Swaim, C.D., Scott, A.F., Canadeo, L.A., Huibregtse, J.M., 2017. Extracellular ISG15 Signals Cytokine Secretion through the LFA-1 Integrin Receptor. Mol Cell 68(3), 581-590.e585.

Tang, Y., Zhong, G., Zhu, L., Liu, X., Shan, Y., Feng, H., Bu, Z., Chen, H., Wang, C., 2010. Herc5 Attenuates Influenza A Virus by Catalyzing ISGylation of Viral NS1 Protein. The Journal of Immunology 184(10), 5777-5790.

Tokarz, S., Berset, C., La Rue, J., Friedman, K., Nakayama, K., Nakayama, K., Zhang, D.E., Lanker, S., 2004. The ISG15 isopeptidase UBP43 is regulated by proteolysis via the SCFSkp2 ubiquitin ligase. J Biol Chem 279(45), 46424-46430.

Villarroya-Beltri, C., Baixauli, F., Mittelbrunn, M., Fernandez-Delgado, I., Torralba, D., Moreno-Gonzalo, O., Baldanta, S., Enrich, C., Guerra, S., Sanchez-Madrid, F., 2016. ISGylation controls exosome secretion by promoting lysosomal degradation of MVB proteins. Nat Commun 7, 13588.

Vuillier, F., Li, Z., Commere, P.H., Dynesen, L.T., Pellegrini, S., 2019. USP18 and ISG15 coordinately impact on SKP2 and cell cycle progression. Sci Rep 9(1), 4066.

Wong, J.J.Y., Pung, Y.F., Sze, N.S.-K., Chin, K.-C., 2006. HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. Proceedings of the National Academy of Sciences 103(28), 10735-10740.

Yeung, T.L., Tsai, C.C., Leung, C.S., Au Yeung, C.L., Thompson, M.S., Lu, K.H., Freedman, R.S., Birrer, M.J., Wong, K.K., Mok, S.C., 2018. ISG15 Promotes ERK1 ISGylation, CD8+ T Cell Activation and Suppresses Ovarian Cancer Progression. Cancers (Basel) 10(12).

Zhang, D., Zhang, D.-E., 2011. Interferon-stimulated gene 15 and the protein ISGylation system. J Interferon Cytokine Res 31(1), 119-130.

Zhang, H.M., Yang, J., Sun, H.R., Xin, X., Wang, H.D., Chen, J.P., Adams, M.J., 2007. Genomic analysis of rice stripe virus Zhejiang isolate shows the presence of an OTU-like domain in the RNA1 protein and a novel sequence motif conserved within the intergenic regions of ambisense segments of tenuiviruses. Arch Virol 152(10), 1917-1923.

Zhang, X., Bogunovic, D., Payelle-Brogard, B., Francois-Newton, V., Speer, S.D., Yuan, C., Volpi, S., Li, Z., Sanal, O., Mansouri, D., Tezcan, I., Rice, G.I., Chen, C., Mansouri, N., Mahdaviani, S.A., Itan, Y., Boisson, B., Okada, S., Zeng, L., Wang, X., Jiang, H., Liu, W., Han, T., Liu, D., Ma, T., Wang, B., Liu, M., Liu, J.Y., Wang, Q.K., Yalnizoglu, D., Radoshevich, L., Uze, G., Gros, P., Rozenberg, F., Zhang, S.Y., Jouanguy, E., Bustamante, J., Garcia-Sastre, A., Abel, L., Lebon, P., Notarangelo, L.D., Crow, Y.J., Boisson-Dupuis, S., Casanova, J.L., Pellegrini, S., 2015. Human intracellular ISG15 prevents interferon-alpha/beta over-amplification and auto-inflammation. Nature 517(7532), 89-93.

Zhang, Y., Thery, F., Wu, N.C., Luhmann, E.K., Dussurget, O., Foecke, M., Bredow, C., Jimenez-Fernandez, D., Leandro, K., Beling, A., Knobeloch, K.P., Impens, F., Cossart, P., Radoshevich, L., 2019. The in vivo ISGylome links ISG15 to metabolic pathways and autophagy upon Listeria monocytogenes infection. Nat Commun 10(1), 5383.

Zhao, C., Hsiang, T.Y., Kuo, R.L., Krug, R.M., 2010. ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. Proc Natl Acad Sci U S A 107(5), 2253-2258.

Zhao, C., Sridharan, H., Chen, R., Baker, D.P., Wang, S., Krug, R.M., 2016. Influenza B virus non-structural protein 1 counteracts ISG15 antiviral activity by sequestering ISGylated viral proteins. Nature Communications 7(1), 12754.

CHAPTER 3

CHARACTERIZATION AND NONCOVALENT INHIBITION OF THE DEUBIQUITINASE AND DEISGYLASE ACTIVITY OF SARS-COV-2 PAPAIN-LIKE PROTEASE[2]

---

[2] Freitas, B.T., I.A. Durie, J. Murray, J.E. Longo, H.C. Miller, D. Crich, R.J. Hogan, R.A. Tripp, and S.D. Pegan. 2020. *ACS Infectious Diseases*. 6 (8), 2099-2109. doi: 10.1021/acsinfecdis.0c0018. https://pubs.acs.org/doi/10.1021/acsinfecdis.0c00168, Reprinted here with permission of the publisher.

Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent for COVID-19, is a novel human betacoronavirus that is rapidly spreading worldwide. The outbreak currently includes over 3.7 million cases and 260,000 fatalities. As a betacoronavirus, SARS-CoV-2 encodes for a papain-like protease (PLpro) that is likely responsible for cleavage of the CoV viral poly-peptide. The PLpro is also responsible for suppression of host innate immune responses by virtue of its ability to reverse host ubiquitination and ISGylation events. Here, the biochemical activity of SARS-CoV-2 PLpro against ubiquitin and ISG15 substrates is evaluated revealing that the protease has a marked reduction in its ability to process K48 linked Ub substrates compared to its counterpart in SARS-CoV. Additionally, its substrate activity more closely mirrors that of the PLpro from the Middle East respiratory syndrome coronavirus and prefers ISG15s from certain species including humans. Additionally, naphthalene based PLpro inhibitors are shown to be effective at halting SARS-CoV-2 PLpro activity as well as SARS-CoV-2 replication.

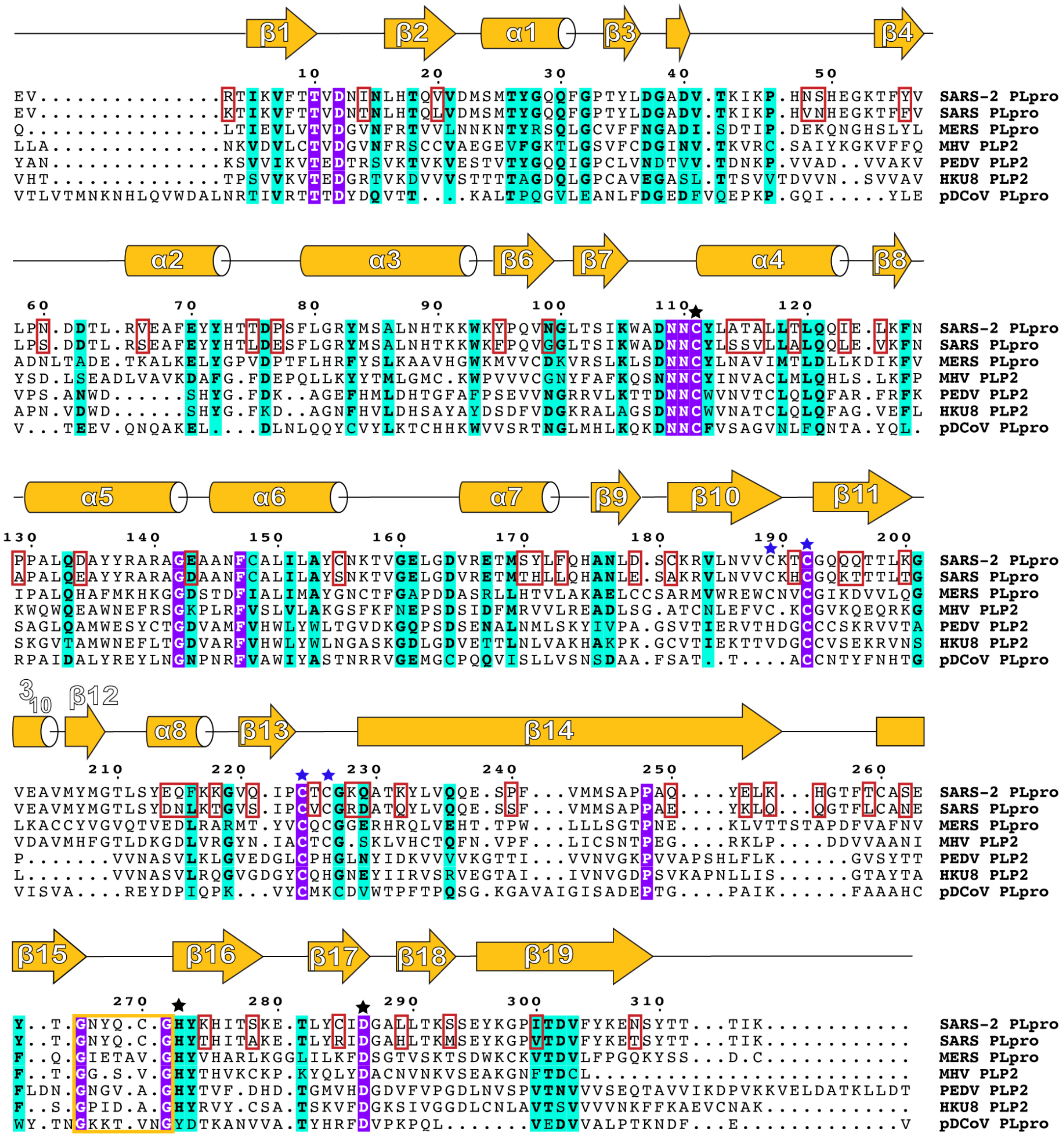
Introduction

COVID-19 disease is caused by SARS-CoV-2, which was identified in Wuhan, China[1-2]. SARS-CoV-2 is classified as a betacoronavirus from the same species as the severe acute respiratory syndrome coronavirus (SARS-CoV), which was responsible for a pandemic in 2002-2003[1, 3]. SARS-CoV-2 has rapidly spread worldwide to over 184 countries with at least 3.6 million cases and >260,000, fatalities according to the latest World Health Organization situation report as of May 6, 2020. The rapid spread of SARS-CoV-2 and its ability to cause death disproportionately in older individuals, or individuals with underlying conditions, has created an urgent need for antiviral therapeutics and vaccines for use against the virus[4].

Upon entry into the cell, SARS-CoV-2 and other betacoronaviruses initially translate two polypeptides pp1a and pp1ab that encode up to 16 nonstructural proteins (Nsp1 to Nsp16). Included within this polypeptide are proteins necessary to form the viruses replicase complex. Once formed, this complex then transcribes the viruses RNA genome before translation of the viruses nucleocapsid protein and structural proteins S, E, and M. Lastly, these components are formed into mature virions within the endoplasmic reticulum–Golgi intermediate compartment[5]. One of the essential steps for successful viral replication is the formation of the viral replicase complex through the cleavage of the pp1a and pp1ab polypeptides by two viral proteases[6-7]. One of the main proteases from CoV, the 3C-like protease, is known for its ability to cleave Nsp4-Nsp16. In addition to the 3C-like protease, CoVs can also encode for up to two papain-like proteases (PLPs) of which one cleaves Nsp1-3. For example, CoVs such as the mouse hepatitis virus (MHV) and other human coronaviruses include NL63, OC43, HKU1, and 229E encode for a PLP1 and PLP2[8]. The genome of SARS-CoV-2 mirrors that of the Middle East respiratory syndrome CoV (MERS-CoV) and SARS-CoV by coding for a single PLpro[8].

Beyond the role of PLpros to cleave the viral polypeptide, PLpros, and their PLP2 counterparts in some CoVs, have also been observed to suppress host innate immune responses through the reversal of post-translational modification of proteins by ubiquitin (Ub) and interferon-stimulated gene product 15 (ISG15)[7, 9]. Up to eight different linkage forms of ubiquitination, as well as ISGylation, events have been observed to regulate facets of the innate immune defense, which a virus must outpace before the infection is cleared by the adaptive immune system[10-12]. Specifically, modification of host proteins by Ub and ISG15 has been shown to facilitate the NFκB inflammation and IFN-I responses[13]. Also, ubiquitination and ISGylation can upregulate the production of cytokines, chemokines, and other IFN-stimulated gene products with antiviral properties during infection[13-14]. Apart from the modification of host proteins by Ub and ISG15, degradation or sequestration of viral proteins via ISGylation has also been found to play a role in host immunity[15-17]. For PLpros and their PLP2 equivalents, their direct overall impact on CoV pathogenesis has been previously shown to be substantial[7, 13-14, 18]. Although the exact role of the deISGylating activity of these proteases remains unclear[19-20], a recent study using an altered MHV PLP2 with ablated deubiquitinase functionality was shown to attenuate pathogenesis in mice. At least part of this reduced pathogenesis for the virus encoding for the altered MHV PLP2 was linked to the IFN response being triggered earlier than under infection by wild-type MHV[21].

Although the role of these proteases in suppressing the innate immune response is clear, their Ub and ISG15 substrate specificities can vary widely[7]. Differences in substrate specificity for these proteases also extend to the eight different linkage forms of polyubiquitin (poly-Ub)[7, 22]. PLpros and PLP2s have also been shown to be sensitive to these species-species variations within ISG15 with their preference including ISG15s from species that they productively infect[8]. This has given rise to the suggestion that viral deubiquitinases and deISGylases may differ in effectiveness toward certain innate immune pathways[7, 15, 23-24]. Recent studies have shown that specificity among PLpros for Ub and ISG15 substrates can be altered with as little as a single amino acid change[8, 19-20]. For instance, the ablation in the deubiquitinating activity of MHV PLP2 that leads to a significant change in pathogenesis came through a change of aspartate to alanine at a single location[21]. Overall, the pp1ab from SARS-CoV-2 Wuhan-Hu-1 isolate (accession number MN908947.3) has an 80% amino acid identity with SARS-CoV-1 (accession number P0C6U8) when determined by NCBI p-blast. Focusing on PLpro the two viruses share an 83% sequence identity (Figure 3.1). This raises the prospect that SARS-CoV-2 PLpro may not possess the same deubiquitinating and deISGylating activities as its SARS-CoV counterpart.



```
                      10        20        30        40         50
1   EV...............RTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADV.TKIKP.HNSHEGKTFYV  SARS-2 PLpro
1   EV...............KTIKVFTTVDNTNLHTQLVDMSMTYGQQFGPTYLDGADV.TKIKP.HVNHEGKTFFV  SARS PLpro
1   Q................LTIEVLVTVDGVNFRTVVLNNKNTYRSQLGCVFFNGADI.SDTIP.DEKQNGHSLYL  MERS PLpro
1   LLA..............NKVDVLCTVDGVNFRSCCVAEGEVFGKTLGSVFCDGINV.TKVRC.SAIYKGKVFFQ  MHV PLP2
1   YAN..............KSVVIKVTEDTRSVKTVKVESTVTYGQQIGPCLVNDTVV.TDNKP.VVAD..VVAKV  PEDV PLP2
1   VHT..............TPSVVKVTEDGRTVKDVVVSTTTTAGDQLGPCAVEGASL.TTSVVTDVVN..SVVAV  HKU8 PLP2
1   VTLVTMNKNHLQVWDALNRTIVRTTTDYDQVTT....KALTPQGVLEANLFDGEDFVQEPKP.GQI.....YLE  pDCoV PLpro

     60        70        80        90        100       110       120
58  LPN.DDTL.RVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIKWADNNCYLATALLTLQQIE.LKFN  SARS-2 PLpro
58  LPS.DDTL.RSEAFEYYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLE.VKFN  SARS PLpro
57  ADNLTADE.TKALKELYGPVDPTFLHRFYSLKAAVHGWKMVVCDKVRSLKLSDNNCYLNAVIMTLDLLKDIKFV  MERS PLpro
59  YSD.LSEADLVAVKDAFG.FDEPQLLKYYTMLGMC.KWPVVVCGNYFAFKQSNNNCYINVACLMLQHLS.LKFP  MHV PLP2
57  VPS.ANWD......SHYG.FDK..AGEFHMLDHTGFAFPSEVVNGRRVLKTTDNNCWVNVTCLQLQFAR.FRFK  PEDV PLP2
58  APN.VDWD......SHYG.FKD..AGNFHVLDHSAYAYDSDFVDGKRALAGSDNNCWVNATCLQLQFAG.VEFL  HKU8 PLP2
65  V...TEEV.QNQAKEL....DLNLQQYCVYLKTCHHKWVVSRTNGLMHLKQKDNNCFVSAGVNLFQNTA.YQL.  pDCoV PLpro

    130       140       150       160       170       180       190       200
129 PPALQDAYYRARAGEAANFCALILAYCNKTVGELGDVRETMSYLFQHANLD.SCKRVLNVVCKTCGQQQTTLKG  SARS-2 PLpro
129 APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLE.SAKRVLNVVCKHCGQKTTTLTG  SARS PLpro
130 IPALQHAFMKHKGGDSTDFIALIMAYGNCTFGAPDDASRLLHTVLAKAELCCSARMVWREWCNVCGIKDVVLQG  MERS PLpro
129 KWQWQEAWNEFRSGKPLRFVSLVLAKGSFKFNEPSDSIDFMRVVLREADLSG.ATCNLEFVC.KCGVKQEQRKG  MHV PLP2
120 SAGLQAMWESYCTGDVAMFVHWLYWLTGVDKGQPSDSENALNMLSKYIVPA.GSVTIERVTHDGCCCSKRVVTA  PEDV PLP2
121 SKGVTAMWNEFLTGDVARFVHWLYWLNGASKGDLGDVETTLNLVAKHAKPK.GCVTIEKTTVDGCCVSEKRVNS  HKU8 PLP2
129 RPAIDALYREYLNGNPNRFVAWIYASTNRRVGEMGCPQQVISLLVSNSDAA.FSAT..T....ACCNTYFNHTG  pDCoV PLpro

              210       220       230       240        250           260
202 VEAVMYMGTLSYEQFKKGVQ.IPCTCGKQATKYLVQQE.SPF...VMMSAPPAQ....YELK...HGTFTCASE  SARS-2 PLpro
202 VEAVMYMGTLSYDNLKTGVS.IPCVCGRDATQYLVQQE.SSF...VMMSAPPAE....YKLQ...QGTFLCANE  SARS PLpro
204 LKACCYVGVQTVEDLRARMT.YVCQCGGERHRQLVEHT.TPW...LLLSGTPNE....KLVTTSTAPDFVAFNV  MERS PLpro
201 VDAVMHFGTLDKGDLVRGYN.IACTCG.SKLVHCTQFN.VPF...LICSNTPEG....RKLP....DDVVAANI  MHV PLP2
193 P.......VVNASVLKLGVEDGLCPHGLNYIDKVVVVKGTTI...VVNVGKPVVAPSHLFLK......GVSYTT  PEDV PLP2
194 L.......VVNASVLRQGVGDGYCQHGNEYIIRVSRVEGTAI...IVNVGDPSVKAPNLLIS......GTAYTA  HKU8 PLP2
196 VISVA....REYDPIQPK...VYCMKCDVWTPFTPQSG.KGAVAIGISADEPTG....PAIK......FAAAHC  pDCoV PLpro

          270         280         290        300        310
264 Y..T.GNYQ.C.GHYKHITSKE.TLYCIDGALLTKSSEYKGPITDVFYKENSYTT...TIK............  SARS-2 PLpro
264 Y..T.GNYQ.C.GHYTHITAKE.TLYRIDGAHLTKMSEYKGPVTDVFYKETSYTT...TIK............  SARS PLpro
269 F..Q.GIETAV.GHYVHARLKGGLILKFDSGTVSKTSDWKCKVTDVLFPGQKYSS...D.C............  MERS PLpro
261 F..T.GG.S.V.GHYTHVKCKP.KYQLYDACNVNKVSEAKGNFTDCL..........................  MHV PLP2
251 FLDN.GNGV.A.GHYTVF.DHD.TGMVHDGDVFVPGDLNVSPVTNVVVSEQTAVVIKDPVKKVELDATKLLDT  PEDV PLP2
252 F..S.GPID.A.GHYRVY.CSA.TSKVFDGKSIVGGDLCNLAVTSVVVVNKFFKAEVCNAK............  HKU8 PLP2
252 WY.TNGKKT.VNGYDTKANVVA.TYHRFDVPKPQL.......VEDVVALPTKNDF...E.............V  pDCoV PLpro
```

Figure 3.1. Sequence alignment of PLPs from coronaviruses. The PLpro or PLP from SARS-CoV-2 (accession number MN908947.3), SARS-CoV-1 (accession number P0C6U8), MERS-CoV (accession number AFS88944), HCoV-OC43 (accession number AMK59674), HCoV-229E (accession number APT69896), and HCoV-HKU1 (accession number ARB07606). The secondary structure shown is the predicted by DSSP for SARS-CoV PLpro (5E6J). Similarity and alignment calculations were performed using ClustalW. Residue positions that are fully conserved are marked in purple, with those being highly conserved are marked in cyan. Residues that form the catalytic triad are marked with black stars, while residues forming the zinc finger motif are marked with blue stars. The BL2 loop is boxed in gold. Sites of amino acid difference between PLpros of SARS-CoV-1 and SARS-CoV-2 are boxed in red.

The dual viral polypeptide cleavage and immune suppression roles of PLpros have previously made them a sought-after target for small molecule antiviral development[25-27]. In 2008, the first classes of non-covalent drug-like inhibitors, now known as naphthalene PLpro inhibitors, were discovered. Certain members of this class of PLpro inhibitors exhibited nanomolar inhibition against SARS-CoV PLpro and could stymie viral replication in the low micromolar range[3, 28]. These and other SARS-CoV PLpro based naphthalene inhibitors are promising for their potency and high selectivity for SARS-CoV PLpro over host proteases[3, 26-27]. They also demonstrated no cellular toxicity in Vero E6 cells or A549 cells, with some analogs considered to be metabolically stable[3, 26-27]. However, they showed no appreciable ability to inhibit PLpros from other circulating CoVs[3, 25-27]. With the SARS-CoV outbreak effectively contained in 2003 with no reemergence, the interest in these potential CoV therapeutics had waned. Given the urgent need for SARS-CoV-2 therapies, whether these naphthalene PLpro inhibitors can now serve as a jumping-off point for SARS-CoV-2 antiviral development is an open question.

Here, we show the first biochemical characterization of the deubiquitinating and deISGylating activities of the SARS-CoV-2 PLpro using 7-amino-4-methyl coumarin (AMC) conjugated Ub and ISG15. These studies reveal marked differences in SARS-CoV-2 PLpro's kinetic values for these substrates compared to its SARS-CoV-1 counterpart, and explore the protease's ability to cleave the eight different poly-Ub linkages. The preference of SARS-CoV-2 PLpro for certain species' ISG15s is also examined. Lastly, we show that naphthalene PLpro inhibitors designed for SARS-CoV can inhibit SARS-CoV-2 PLpro as well as impede SARS-CoV-2 replication.

Results

Differences between the PLpro from SARS-CoV and SARS-CoV-2 within the UIM

To explore the potential impact of the 54 differences between the PLpros of SARS-CoV and SARS-CoV2 on enzymatic activity, a homology model was constructed of SARS-CoV-2 PLpro encoded by the severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1 (accession number MN908947.3; Figure 3.1). The PDB entry 5E6J of SARS-PLpro bound to K-48 di-Ub was chosen as a template to provide the best representation of a SARS-CoV-2 PLpro in a holo open conformation receptive to substrate binding[18, 29]. From the surface perspective of the SARS-CoV-2 PLpro homology model, 40 of the 54 difference sites were spread out relatively equally over the fingers, palm, thumb and UbL domain of the protease (Figure 3.2, S3.1). On closer examination of the SARS-CoV-2 PLpro's ubiquitin interacting motif (UIM) that is known to accommodate both Ub and ISG15, six sites were found to differ in amino acids from those found in its SARS-CoV counterpart. Specifically, the differences on the SARS-CoV-2 surface within the UIM were S170(T), Y171(H), Y216(L), Q195(K), T225(V) and K232(Q) where the equivalent SARS-CoV residues are marked in parenthesis. Intriguingly, one of these sites in SARS-CoV-2 PLpro, K232, is equivalent to Q233 in SARS-CoV PLpro. Previously, a mutation Q233E notably diminished the deubiquitinase activity of that PLpro in favor of more robust deISGylase activity[8]. This further suggests that the enzymatic activities of SARS-CoV-2 PLpro may indeed differ from those of the SARS-CoV PLpro.

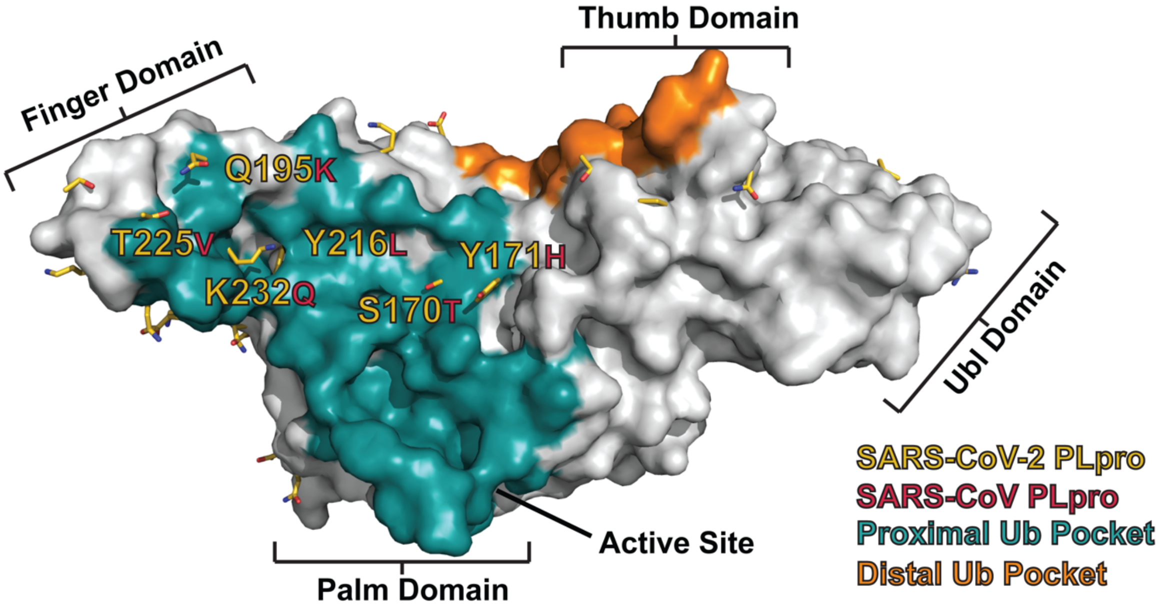


Figure 3.2. Surface rending of a SARS-CoV-2 PLpro homology model highlighting its differences with SARS-CoV-1 PLpro. The SARS-CoV-2 PLpro is shown in grey, with the proximal ubiquitin binding site in teal and the distal ubiquitin binding site in orange. Amino acid sites were PLpro differs between SARS-CoV-2 and SARS-CoV are colored in yellow.

Deubiquitinase and DeISGylase activity of SARS-CoV-2 PLpro

To ascertain whether the amino acid differences between SARS-CoV-2 PLpro and its SARS-CoV-1 counterpart translate into differences in enzyme kinetics, SARS-CoV-2 PLpro K_M and k_{cat} values for Ub-AMC and ISG15-AMC as well as the last five consensus amino acids between them (RLRGG; Peptide-AMC) were determined (Table 3.1, Figure S3.2). The enzymatic efficiency of SARS-CoV-2 PLpro for Ub-AMC was 1.3 ± 0.1 $\mu M^{-1}min^{-1}$ with K_M and k_{cat} values 7.9 ± 1.4 and 10.1 ± 0.6 respectively. For the ISG15-AMC substrate, the enzymatic efficiency of SARS-CoV-2 PLpro was 10.3 ± 0.5 $\mu M^{-1}min^{-1}$ with K_M and k_{cat} values 3.9 ± 0.5 and 40.0 ± 1.8. SARS-CoV-2 PLpro's enzymatic efficiency towards peptide-AMC is 0.0051 $\mu M^{-1}min^{-1}$ when assessed using first-order kinetics. Compared to published kinetic values of PLpros, or its PLP2 equivalent, from MHV, SARS-CoV, and MERS-CoV, the PLpro of SARS-CoV-2 surprisingly mirrors closely that of PLpro from MERS-CoV. Both have low enzymatic efficiency towards the peptide substrate and are 2500-3500 times more efficient towards the ISG15 substrate. This is in contrast to SARS-CoV-1 PLpro that is only ~100 times more efficient towards ISG15 substrates. SARS-CoV-2 is ~10 times more efficient as an deISGylase than as a deubiquitinase. However, SARS-CoV-1 PLpro is still a more robust deISGylase with 3 times better enzymatic efficiency towards ISG15-AMC than SARS-CoV-2. As a deubiquitinase, SARS-CoV-2 appears to have the highest substrate affinity among PLpros with the lowest turnover, which is orthogonal to its SARS-CoV-1 counterpart.

Table 3.1. Kinetic analysis of SARS-CoV-2 PLpro with PEP-AMC, Ub-AMC, and ISG15-AMC.

	Substrate		
	Peptide-AMC	Ub-AMC	ISG15-AMC
SARS-CoV-2 PLpro			
k_{cat}/K_m (μM^{-1}min^{-1})	0.0051[a]	1.3±0.1	10.3±0.5
k_{cat} (min^{-1})		10.0±0.8	40.0±1.8
K_m (μM^{-1})		7.9±1.4	3.9±0.5
SARS-CoV-1 PLpro[b]			
k_{cat}/K_m (μM^{-1}min^{-1})	0.3[a]	1.5±0.3	28.9±5.3
k_{cat} (min^{-1})		75.9±8.1	436±40
K_m (μM^{-1})		50.6±7.4	15.1±2.4
MERS-CoV PLpro[b]			
k_{cat}/K_m (μM^{-1}min^{-1})	0.003[a]	1.3±0.2	9.9±1.6
k_{cat} (min^{-1})		18.8±1.2	32.6±1.8
K_m (μM^{-1})		14.3±2.0	3.3±0.5
MHV PLP[c]			
k_{cat}/K_m (μM^{-1}min^{-1})	0.0016[a]	38.3±6.3	2.3±0.1[a]
k_{cat} (min^{-1})		49.8±2.9	
K_m (μM^{-1})		1.3±0.2	

[a] For non-saturating substrates, *k*app is calculated to approximate k_{cat}/K_m

[b] The kinetic parameters of SARS-CoV PLpro (pp1ab ; 1-315) and MERS-CoV PLpro (pp1ab 1484-1802; 3-322) are from Baez-Santos *et al.*[38]

[c] The kinetic parameters of MHV PLP are from Chen, Y. *et al.*[32]

Poly-Ub linkage preferences for SARS-CoV-2 PLpro

In some cases, like the viral ovarian domain proteases encoded by Nairoviruses, viral deubiquitinases have demonstrated different levels of activity towards the more natural ubiquitin chains than for monomeric Ub-AMC[30-31]. To examine if this is the case for SARS-CoV-2 PLpro, its ability to cleave the eight different linkage types of poly-Ub, K6, K11, K27, K33, K48, K63, and linear was assessed. Similar to previous studies with MHV PLP2, 20 nM of SARS-CoV-2 PLpro was incubated with 10 µM of each di-Ub linkage[32]. No cleavage of any di-Ub moiety by SARS-CoV-2 PLpro was detected after 60 min (Figure S3.2). Even use of 10 fold of the enzyme over 120 minutes failed to result in a detectable di-Ub cleavage event (data not shown). The same was observed for tetrameric K63 linked polyubiquitin. Of the polymeric ubiquitin chains tested, only tetrameric K48 polymeric ubiquitin chains appeared to be cleaved by the protease (Figure 3.3).

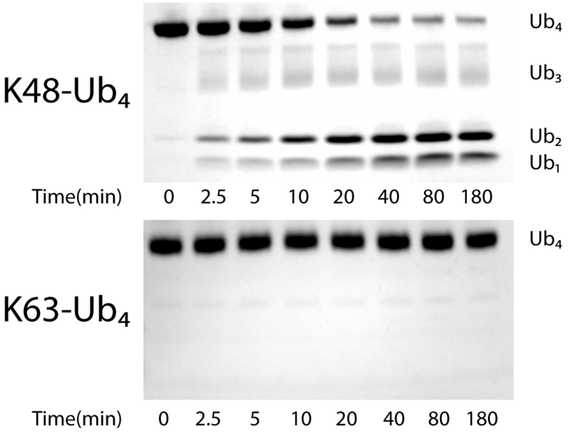


Figure 3.3. SARS-CoV-2 PLpro preferences for K63 and K48 Ub$_4$ linkages. Gel cleavage assay of unlabeled K48 and K63 linked tetra-Ub, visualized by Coomassie Blue staining. At 37°C, 13.7 µM of each Ub moiety was incubated with 23 nM SARS-CoV-2 PLpro for at least 180 minutes with samples taken at time points indicated.

SARS-CoV-2 PLpro exhibits ISG15 species preferences

While Ub is almost completely conserved among animals, sequence similarity for ISG15s within the Mammalia class alone can dip below 60%[33]. Only at the genus level do ISG15s from different species appear to have a higher level of similarity[33] (Figure 3.4a). As this species-species variance in ISG15 has been shown to impact the deISGylase activity of viral deISGylases including PLpros, whether SARS-CoV PLpro followed this phenomenon was explored. Taking advantage of the ability of viral deISGylases to cleave the precursor of ISG15 (proISG15) into mature ISG15 [8, 34-36], the ability of SARS-CoV-2 PLpro to cleave the proISG15s from human, vesper bat, pig, mouse, camel, sheep, cow, Egyptian fruit bat, hedgehog, northern tree shrew, and fish were examined. As with other viral deISGylases, SARS-CoV-2 showed a range of ability to successfully engage and process these ISG15 substrates from different species (Figure 3.4b). Among the 11, the protease appears to prefer ISG15s from sheep and the Vesper bat. This is followed by moderate activity for ISG15s from human, pig, camel, and mouse. Weak SARS-CoV-2 PLpro deISGylase activity was observed for the Egyptian fruit bat, hedgehog, and northern tree shrew. No protease activity was observed for the fish pro-ISG15 substrate. Overall, like other viral deISGylases[8, 37], SARS-CoV-2 PLpro appears to be species-specific for certain subsets of ISG15s including at least one ISG15 from species they are known to infect.

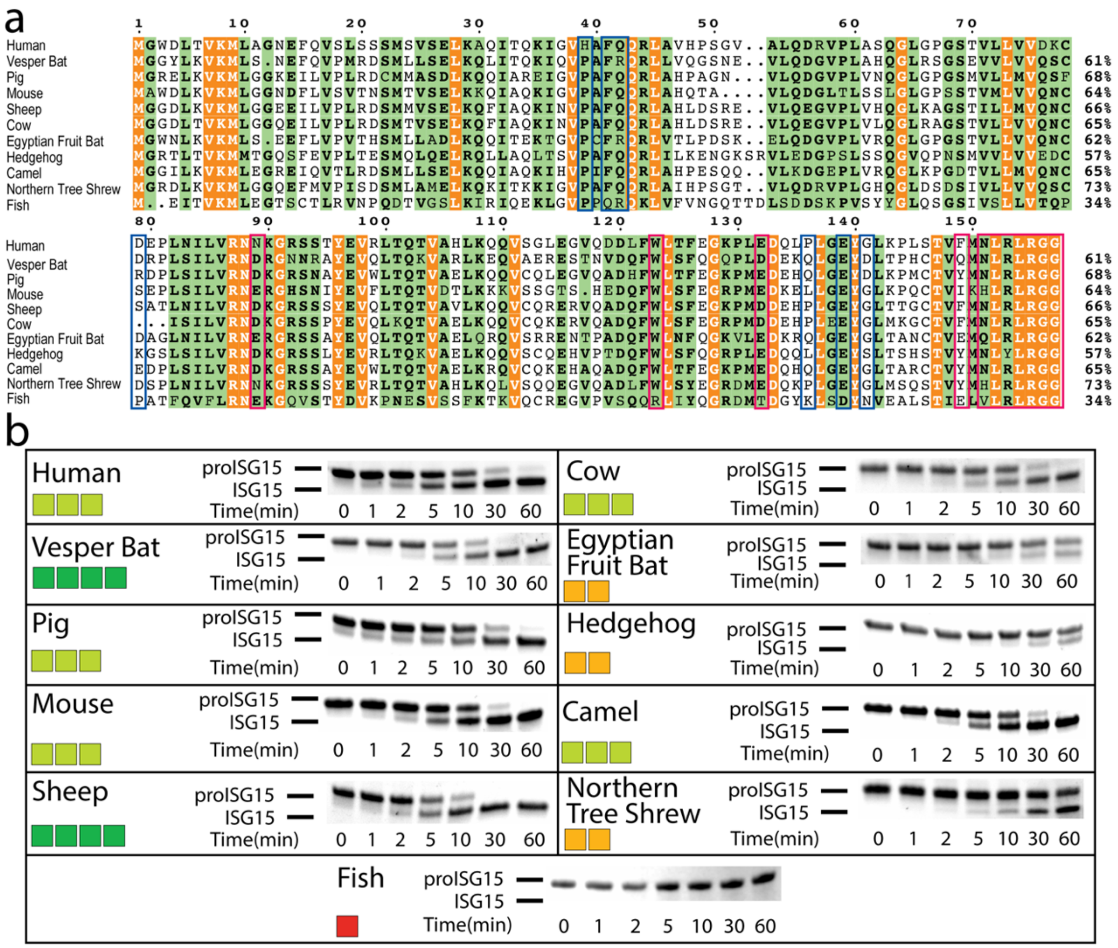


Figure 3.4. The activity of SARS-CoV-2 PLpro for proISG15 from multiple species. (a) Sequence alignment of ISG15s from human (*Homo sapiens*, accession number: AAH09507.1), vesper bat (*Myotis davidii*, accession number: ELK23605.1), pig (*Sus scrofa*; Accession ACB87600.1), mouse (*Mus musculus*, accession number: AAB02697.1), dromedary camel (*Camelus dromedarius*, accession number: XP_010997700.1), sheep (*Ovis aries*, accession number: AF152103.1), cow (*Bos taurus*; NP_776791.1), Egyptian fruit bat (*Rousettus aegyptiacus*; XP_015999857.1), hedgehog (*Erinaceus europaeus*; XP_007525810.2), northern tree shrew (*Tupaia belangeri*, accession number: AFH66859.1), and jackknife fish (*Oplegnathus fasciatus*, accession number: BAJ16365.1). Sequence ruler is based on human ISG15. Similarity and alignment calculations were performed using ClustalW. Residue positions that are fully conserved are marked in yellow, with those being highly conserved are marked in green. Human ISG15 similarity to other ISG15s indicated to the right of the alignment. Red boxes indicate ISG15 amino acid sites known to directly interact with PLpro from SARS-CoV-1 and MERS-CoV[8, 19-20]. Blue boxes indicate amino acid sites that have been implicated in impacting the ISG15 interdomain orientation that can play a role in ISG15-SARS-CoV-1 PLpro engagement[33]. (b) SARS-CoV-2 PLpro was evaluated for cleavage of proISG15s from species in (a), At 37°C, 10 µM of each ISG15 was incubated with 20 nM SARS-CoV-2 PLpro for at least one hour with samples taken at time points indicated. Summary of proISG15 cleavage assays for different CoV PLPs presented as a heat map. Colors range from dark red (no cleavage) to green (relatively robust cleavage).

Inhibition of SARS-CoV-2 PLpro by SARS-CoV PLpro inhibitors

Although SARS-CoV-2 PLpro and SARS-CoV PLpro differ by 54 residues, those lining the active site and nearby P3 and P4 sites are identical (Figure 3.5a). This includes residues in the BL2 loop that were previously shown to be key in binding naphthalene based PLpro inhibitors[3, 27]. To examine whether previously developed SARS-CoV PLpro naphthalene based inhibitors might be effective at inhibiting the SARS-CoV-2, five compounds that either had been previously shown to be efficacious or were close analogs, were chosen for testing (Figure 3.5b)[3]. Emulating Ratia et al., who detailed the potency of these four compounds against SARS-CoV-1, we utilized the peptide-AMC substrate concentration of 50 μM[3]. The most potent of these four proved to be GRL-0617 with an IC_{50} of 2.4 μM, followed by compound 6 with a low micromolar IC_{50} of 5.0 μM towards SARS-CoV-2. These activities were relatively in line with the 600 nM and 2.6 μM IC_{50} values for GRL-0617 and compound 6 respectively reported against SARS-CoV PLpro[3]. The original high-throughput screen lead compounds, 7724772[3] and 6577871[26], that had 20 μM and 59 μM IC_{50} values against SARS-CoV PLpro presented a similar trend of results with SARS-CoV-2 PLpro. A fifth compound (9247873) was also tested but no inhibitory effect was observed at 200 μM.

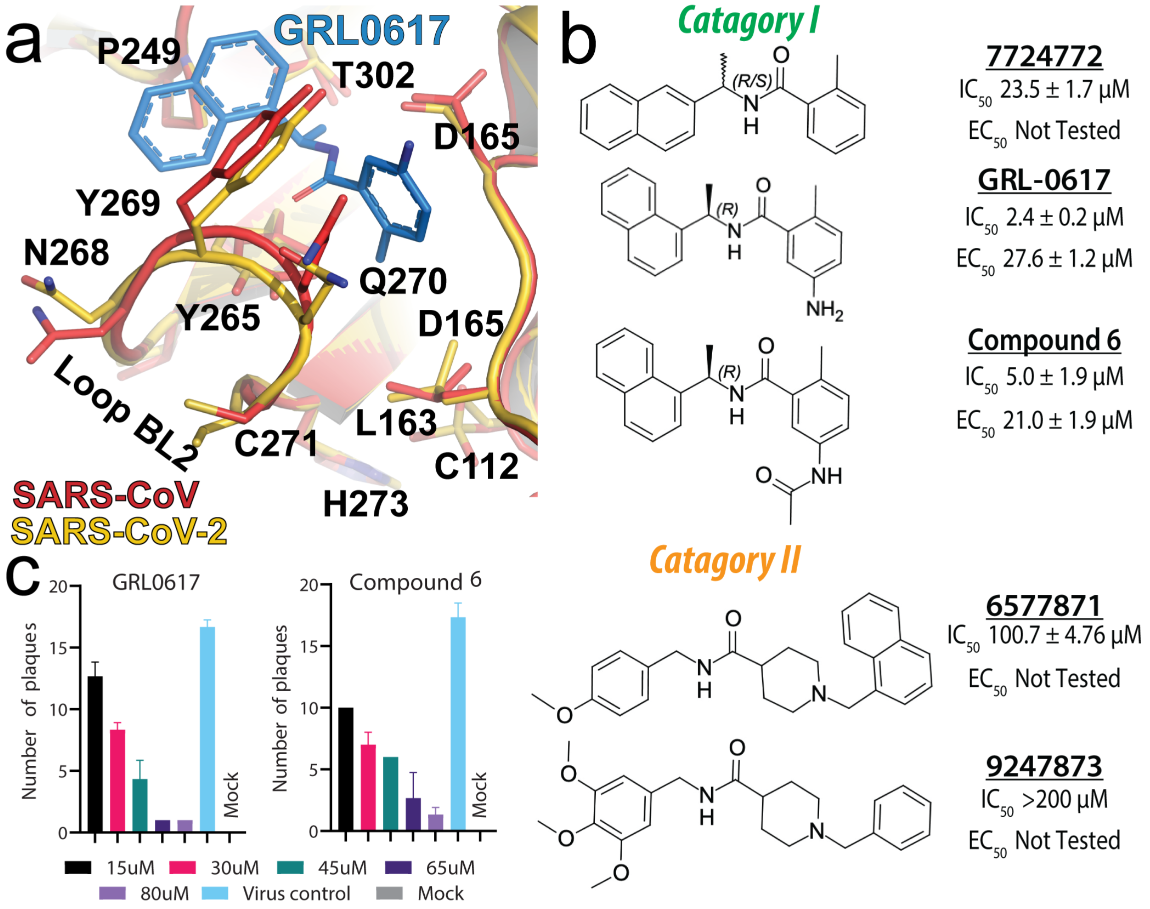


Figure 3.5. SARS-CoV-2 PLpro model with GRL-0617 as well as enzymatic and antiviral data for PLpro inhibitors against SARS-CoV-2 PLpro and SARS-CoV-2 a) Comparison of the P3, P4 substrate binding site of SARS-CoV-2 PLpro homology model and SARS-CoV counterpart (PDB 3E9S). b) IC_{50} and EC_{50} values related to the inhibition of SARS-CoV-2 PLpro and SARS-CoV-2 replication by SARS-CoV PLpro inhibitors. (c) SARS-CoV-2 plaque reduction assay data for GRL-0617 and compound 6. SARS-CoV-2 was incubated with the compounds and assessed 66 hours post-infection to determine if the compounds neutralized the virus infection. Plaques were quantified by visual inspection and compared to non-treated virus control.

Antiviral Activity of SARS-CoV-2 PLpro inhibition

To examine whether the non-covalent naphthalene-based SARS-CoV-2 PLpro inhibitors also possessed antiviral activity for SARS-CoV-2, GRL-0617 and compound 6 were selected for examination against the virus. Plaque reduction assays were performed using Vero E6 cells and the SARS-CoV-2 USA-WA1/2020 isolate to determine efficacy of inhibiting SARS-CoV-2 PLpro. SARS-CoV-2 USA-WA1/2020 isolate was readily available and like other SARS-CoV-2 isolates, its PLpro was fully conserved with the Wuhan-Hu-1 isolate. Excitingly, GRL-0617 and compound 6 exhibited EC_{50} values of 27.6 μM and 21.0 μM respectively (Figure 3.5c). In line with previous studies[3, 27], no cytotoxicity was observed when treating the GRL-0617 and compound 6 at the concentrations utilized in this study.

Discussion

Deubiquitinating activities of SARS-CoV-2 PLpro

The recent revelation that MHV PLP2 deubiquitinase activity is tied to the pathogenesis of this coronavirus by its downregulation of the IFN response mirrors in many ways what was previously observed in the Crimean Congo hemorrhagic fever virus (CCHFV)[15, 21]. When the CCHFV encoded protease that possessed both deubiquitinase and deISGylase functionality had its deubiquitinase functionality ablated, a more robust IFN response was observed than with the wild-type virus[15].

Comparison of SARS-CoV-2 PLpro kinetic parameters to these other two proteases taken from other studies[32, 38] highlights that SARS-CoV-2 PLpro appears to perform enzymatically more like MERS-CoV PLpro than its SARS-CoV counterpart (Table 3.1). Like MERS-CoV PLpro, SARS-CoV-2 PLpro can be readily saturated with Ub-AMC but turns the substrate over substantially more slowly. Although the catalytic efficiency of SARS-CoV-2 PLpro and MERS-CoV PLpro is similar to that of SARS-PLpro for Ub-AMC, their kinetic parameters reveal that they are more susceptible to product inhibition by Ub than their SARS-CoV PLpro equivalent. Given that cellular pools of free ubiquitin in mammalian cells have been found to range from 10-23 μM[39], the overall performance of MERS-CoV and SARS-CoV-2 towards monomeric ubiquitinated substrates in a cellular context could be more divergent than that of SARS-CoV PLpro. In other words, while the PLpro of MERS-CoV and SARS-CoV-2 is at, or near saturating conditions, the PLpro from SARS-CoV-1 is not even at its K_M concentration. The more complex environment of polyubiquitin chain cleavage seems to attest to this difference between the PLpros from MERS-CoV and SARS-CoV-2 versus that of SARS-CoV PLpro. Like MERS-CoV PLpro, SARS-CoV-2 PLpro cleaves K48-linked tetra-Ub substantially slower than SARS-CoV PLpro in previous studies[38]. However, SARS-CoV-2 PLpro is not entirely similar to MERS-CoV in its deubiquitinase activity. Unlike the MERS-CoV protease, SARS-CoV-2 PLpro is similar to its SARS-CoV counterpart in that it shows no appreciable activity for K63 linked polyubiquitin chains[38].

Given the 83% identity of the SARS-CoV and SARS-CoV-2 at the amino acid level, the appearance of such different deubiquitinating enzymatic profiles between the PLpros encoded by these viruses further highlights that even proteins from viruses within the same species can perform in notably divergent ways. It also highlights some potentially interesting insights from an evolutionary point of view. Among the seven amino acid differences within the UIM, the natural appearance of lysine at SARS-CoV-2 PLpro amino acid site 232 was particularly surprising. The mutation of this equivalent site in SARS-CoV PLpro to glutamate creates an electrostatic repulsion with ubiquitin that diminished that protease's deubiquitinase activities. In SARS-CoV-2, nature appears to have selected for lysine at this position that should logically, and appears to have, increased the protease's affinity for Ub at the expense of overall deubiquitinating functionality. Coronavirus PLpro's activities towards K48-linked ubiquitin have been suggested to counter NF-κB translocation to the nucleus with activity toward K63-linked ubiquitin stymieing others[27].

With viral deubiquitinase activity being a factor in pathogenesis, nature selecting for a viral deubiquitinase with weaker K48 cleavage capability, relatively little K63 cleavage activity, and slower activity towards mono-Ub than its SARS-CoV-1 counterpart at concentrations resembling cellular concentrations may seem counter-intuitive at first. However, increased lethality of the host is not necessarily the primary driver of viral evolution as this can lead to the virus eradicating itself. If nature values successful viral propagation more as a driver, having a weaker viral deubiquitinase may be a better fit. Whether or not this is the case with SARS-CoV-2, which has been able to evade quarantine efforts because of the more often than not initial mild disease symptoms it causes, is an open question. A question whose answer might include virulence factors like PLpros.

SARS-CoV-2 PLpro deISGylating activities

As has been found with other PLpros and PLP2s from prominent coronaviruses, SARS-CoV-2 PLpro has a pronounced preference for ISG15 over Ub[8]. Despite the recent insight into the role of these coronavirus protease deubiquitinations, the exact role of viral deISGylase in coronaviruses is a mystery. However, the relatively consistent dominant presence of this type of protease activity among PLpros and PLP2s, as well as its demonstrated viral evasion role in other viruses,[15] highlights how important coronavirus deISGylase activity maybe for the virus.

Species-species variations in ISG15 have been shown to impact viral replication of influenza B highlighting the role ISG15 can have on the zoonotic range of influenza B[40-41]. Although it is not known in detail how PLpro deISGylation activity plays a role in coronavirus infection, this activity has been observed to be sensitive to species-species variances within ISG15[8, 33]. For instance, PLP2 from MHV can readily process mouse ISG15 substrates but not human ISG15 ones[8]. In the case of MER-CoV

PLpro, camel ISG15 is among those species ISG15s toward which it has the most activity. The PLpro from SARS-CoV-2 also exhibits this species-specificity phenomenon for its deISGylase activity (Figure 3.4). It also keeps with the trend set by other viral deISGylases in that it can engage ISG15s from at least the species the virus is known to productively infect: humans[20, 37].

The vesper bat ISG15 was one of the fastest cleaved substrates by SARS-CoV-2; this bat circulates within the Hubei province, lending credence to SARS-CoV-2 originating from bats. However, Egyptian fruit bat ISG15 was cleaved very slowly and did not appear to be a suitable substrate. This is not necessarily surprising as bat ISG15s can have as low as 60% sequence identity and coronaviruses have been seen as being specific to certain bat species[33, 42-43]. Whether the vesper bat's presence in the general region from which SARS-CoV-2 originates, or if by happenstance vesper bat ISG15 has similarity at amino acid positions to those of host species for SARS-CoV-2, has a role in the preference requires further investigation[8, 20, 37, 44].

When compared with SARS-CoV PLpro, the SARS-CoV-2 protease had similar species preferences, particularly in regards to humans, vespers bats, and mice. However, it had slightly higher activity towards sheep, and less towards camel[8]. SARS-CoV PLpro also had no activity for fish ISG15 in contrast to its counterparts in MER-CoV, MHV, and SARS-CoV. Although without experimentation all 54 amino acid sites where the two viruses differ could be responsible for this divergence in species-specific deISGylase activity, the 7 divergent sites located within the two viruses' known PLpro-ISG15 interfaces likely hold the most promise. Recently, selective removal of general deISGylase activities from the PLpros of MERS-CoV and SARS-CoV has illuminated the path to molecular tools that can reveal the role of

PLpro deISGylase activity has on coronavirus pathogenesis[8, 19-20]. This new information that species-specificity of SARS-CoV-2 PLpro is not fully conserved with that of its SARS-CoV-1 counterpart may provide a new tool for unraveling the role of viral deISGylation in coronavirus replication among different hosts. Specifically, a SARS-CoV-2 encoding for a PLpro that has been modified to selectively ablate the deISGylase activities of one species in favor of another.

Use of CoV PLpro inhibitors as a starting point for SARS-CoV-2 therapeutics

Protease inhibitors have a long history of being used as a basis for antiviral therapy, the most salient examples being HIV and Hepatitis C[45]. Within coronaviruses themselves, main protease inhibitors have been shown to reverse the progression of fatal coronavirus infection[46]. With no therapeutics or vaccines available for the treatment of those infected by SARS-CoV-2, there is an overwhelming need to identify lead compounds that are effective against proposed viral drug targets with SARS-CoV-2. The low micromolar efficacy of GRL-0617 and compound 6 toward SARS-CoV-2 PLpro and the virus itself suggests that previously designed SARS-CoV PLpro inhibitors would be a good place to start. These two compounds, as well as the two scaffolds of the five compounds tested, represent known non-covalent inhibitor classes of compounds. Additionally, these compounds have been shown to have low cellular toxicity in multiple cell lines[3, 27], and some have displayed the potential to be metabolically stable[27]. Future experiments will need to be performed in additional cell lines as will experiments to tease out some of the pharmacological nuances with the compounds. For now, these scaffolds, or others similarly targeting PLpro, highlight a viable path to antiviral development and potential use.

Conclusions

The biochemical characterization of the deubiquitinating and deISGylating activities of SARS-CoV-2 revealed that it more closely resembles that of its counterpart in MERS-CoV than in SARS-CoV. This includes a marked reduction in deubiquitinating activities to include that of cleaving K48-linked tetra-Ub. As with other coronaviruses PLpros and PLP2s, the deISGylating activity of SAR-CoV-2 PLpro appeared the more dominant of its various proteolytic functions. This activity also appeared to be species-specific only cleaving ISG15 substrates from select species including humans. Although the 54 differences between the PLpros from SARS-CoV and SAR-CoV-2 impacted the proteases functionality, they did not appreciably affect the activity of naphthalene based PLpro inhibitors designed for SARS-CoV efficacy against SARS-CoV-2 from a drug discovery perspective. This revelation offers a potential rapid development path to generating PLpro targeted therapeutics for use against SARS-CoV-2.

Methods

Chemicals and reagents

5-Amino-2-methyl-N-[(R)-1-(1-naphthyl)ethyl]benzamide (GRL-0617) was purchased from Raystar, CN, 5-(acetylamino)-2-methyl-N-[(1R)-1-(1-napthanlenyl)ethyl]-benzamide (compound 6) was purchased from MedChem Express, 2-methyl-N-[1-(2-naphthyl)ethyl]benzamide (7724772) was purchased from Chembridge, N-(4-methoxybenzyl)-1-(1-naphthylmethyl)-4-piperidinecarboxamide oxalate (6577871) was purchased from Chembridge, 1-benzyl-N-(3,4,5-trimethoxybenzyl)-4-piperidinecarboxamide (9247873) from Chembridge, Z-RLRGG-7-amino-4-methyl-courmarin (Peptide-AMC) was purchased from Bachem, ubiquitin-7-amino-4-methylcourmarin (Ub-AMC) was purchased from Boston Biochem, human ISG15-7-

amino-4-methylcourmarin (ISG15-AMC) was purchased from Boston Biochem, Lys6, Lys11, Lys29, Lys33, Lys48, Lys63, and linear linked di-Ub were obtained from Boston Biochem, DL-dithiothroitol (DTT) was purchased from GoldBio, Isopropyl-beta-D-thiogalactopyranoside (IPTG) was purchased from GoldBio, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was purchased from Fisher BioReagents, Imidazole was purchased from Acros Organics, tris(hydroxymethyl)aminomethane (Tris) was purchased from Fisher Scientific, Sodium Chloride (NaCl) was purchased from Fisher Chemical, Bovine Serum Albumin (BSA) was purchased from Sigma Life Science.

Homology modeling of SARS-CoV-2 PLpro

SARS-CoV-2 PLpro homology models were generated using the MODELLER software suite, version 9.19[47]. For all models, the PLpro from SARS-CoV-2 (accession number MN908947.3) was used as the unknown. The homology model of SARS-CoV-2 PLpro in its holo open form used PDB entry 5E6J as a template, while PDB entry 3E9S was used as a template for the SARS-CoV-2 PLpro homology model used in the docking of GRL-0617. The X-ray structure of 6W9C is now available.

Construction, Expression and Purification of viral deubiquitinases

The ubiquitin-like domain (UbL) and the catalytic core of SARS-CoV-2 PLpro (orf1ab 1564-1876; 1-315) were cloned into pET-15b by Genscript and transformed into T7 express *E. coli*. Cells were cultured in 4.5 L LB Broth containing 100 µg/mL Ampicillin at 37^{o}C until OD600 reached 0.6. Once reached, expression was induced by the addition of 0.5 mM Isopropyl β- D-thiogalactopyranoside (IPTG) and the culture was incubated at 18^{o}C overnight. The culture was centrifuged at 12,000 x g for 10 minutes and the pellet was collected and stored in -80^{o}C freezer. The cell pellet was dissolved into lysis buffer

(500 mM NaCl and 50 mM Tris-HCl [pH = 7.0]) and then sonicated in Fisher Scientific series 150 on ice at 50% power with 5-second pulses for 6 minutes. The lysate was centrifuged at 26,000 x g for 45 min to remove all insoluble products. The supernatant was then filtered and placed onto Ni-nitrilotriacetic agarose resin (Qiagen). The resin was washed using five column volumes of lysis buffer containing 10 mM Imidazole. The protein was eluted using 5 column volumes of lysis buffer containing 300 mM Imidazole. Thrombin was added to the elution to remove the 6X His-tag and the combined solution was dialyzed in size exclusion buffer (100 mM NaCl, 5 mM HEPES and 2 mM dithiothreitol (DTT) [pH = 7.4]) and run over Size Exclusion Superdex 200 column (GE Healthcare, Pittsburgh PA). Purity was confirmed by gel electrophoresis. The Oman strain of Crimean Congo Hemorrhagic Fever viral ovarian tumor domain protease (1-169) used as a di-Ub control was expressed and purified as previously described[15].

SARS-CoV-2 PLpro deubiquitinase and deISGylating assays

All assays were run using Corning Costar half-volume 96 well plates containing AMC buffer (100 mM NaCl, 50 mM HEPES [pH = 7.5], 0.01 mg/mL bovine serum albumin (BSA), and 5 mM DTT) to a final volume of 50 µL performed in triplicate. CLAIROstar plate reader (BMG Lab Tech, inc.) was used to measure the fluorescence of the AMC cleavage and the data was analyzed using MARS (BMG Lab Tech, inc.). The AMC fluorescence was observed from the cleavage of Ub-AMC and ISG15-AMC obtained from Boston Biochem, MA. ISG15-AMC concentrations of substrate ranged from 1 µM to 15 µM and Ub-AMC ranged from 0.5 µM to 30 µM. Protease concentrations used for the Ub-AMC and ISG15-AMC assays were 5 nM and 0.5 nM respectively. To calculate K_M and V_{max} values, the initial rates were fitted to the Michalis-Menten equation, υ =

$V_{max}/(1+(K_M/[S]))$, using the Enzyme Kinetics (v. 1.3) module of SigmaPlot (v. 10.0, SPSS Inc.). V_{max} was translated into k_{cat} using $k_{cat} = V_{max}/[E]$.

SARS-CoV-2 PLpro poly-Ub cleavage assays

Lys6, Lys11, Lys29, Lys33, Lys48, Lys63, and linear linked di-Ub obtain from Boston Biochem were incubated at 10 μM with 20 nM SARS-CoV-2 PLpro. Reactions were performed in AMC buffer at a volume of 75 μL and a temperature of 37°C. 10 μL samples were taken at the indicated time points and heat-shocked at 98°C for five minutes. Lys48 and Lys63 linked tetra-Ub obtained from Boston Biochem were incubated at 13.65 μM with 23nM SARS-CoV-2 PLpro. Reactions were performed in AMC buffer at a volume of 80 μL and a temperature of 37°C. 10 μL samples were taken at the indicated time points and heat-shocked at 98°C for five minutes. SDS-PAGE analysis was performed using Mini-PROTEAN® TGX™ and coomassie blue.

Protease activity assay with proISG15 substrates

At 37°C, 20 nM SARS-CoV-2 PLpro was run against 10 μM of each ISG15. Reaction mixtures were 100 μL in PLpro buffer (100 mM NaCl, 5 mM HEPES [pH = 7.4]). 10 μL samples were taken at the indicated time points and the reaction quenched in 2x Laemmli sample buffer followed by boiling at 98°C for five minutes. SDS-PAGE analysis was performed using Mini-PROTEAN® TGX Stain-Free™.

SARS-CoV-2 PLpro inhibition IC_{50} value determination

IC_{50} assays were performed using similar methods to Peptide-AMC, Ub-AMC and ISG15-AMC cleavage experiments and those described previously[3]. SARS-CoV-2 PLpro was run at 100 nM against 50 μM Peptide-AMC in 98% AMC buffer 2% DMSO. Reactions were performed in duplicate with inhibitor concentrations ranging from 1.25 μM to 20 μM, or 100 μM depending on compound tested. IC_{50} calculations were performed using Prism8 from GraphPad. For 7724772, compound 6, and GRL-0617, maximum inhibition of 100% was reached. For 6577871, maximum inhibition of 61% was reached.

SARS-CoV-2 antiviral activity assays

SARS-CoV-2 (2019-nCoV/USA-WA1/2020; accession number MN985325.1) was received from BEI resources and propagated in Vero clone E6, Vero E6, CRL-1586. Infections were done at a multiplicity of infection (MOI) of 0.1 in serum-free in Dulbecco's minimal essential medium (DMEM) for 1h after which the virus-containing media was decanted and replaced with DMEM supplemented with 10% heat-inactivated fetal bovine serum[48]. The virus was propagated for 72 hours before it was harvested and the titer determined by plaque assay on Vero E6 cells[49]. The viral plaques were counted and the titer was determined as PFU/ml. The Vero cells were plated at 3 X 10^5 cells/well in 12-well plates and incubated overnight at 37ºC. The following day the GRL-0617 and compound 6 were prepared into the following concentrations/well in a separate plate; 15 μM, 30 μM, 45 μM, 65 μM and 80 μM. The cells were washed once with PBS 1X and then infected with 8,000 PFU/well with the GRL-0617 and compound 6 and incubated for 66 hours at 37ºC at 5% CO_2. The cells were then fixed and stained with crystal violet to determine plaque numbers. These were all done in triplicate and the calculations were performed using Prism8 from GraphPad. Cytotoxicity assessment of compound 6 and GRL-0617 was performed using Lonza Toxilight bioassay. Vero E6 cells were seeded at 10,000 cells per well and incubated overnight at 37C. The plates were washed with 1X PBS and then the compounds were added at the specific concentrations and incubated for 72 hours. The bioassay was completed following the instructions of the assay and direct luminometer light output (RLUs) are measured.

References

1. Gorbalenya, A. E.; Baker, S. C.; Baric, R. S.; de Groot, R. J.; Drosten, C.; Gulyaeva, A. A.; Haagmans, B. L.; Lauber, C.; Leontovich, A. M.; Neuman, B. W.; Penzar, D.; Perlman, S.; Poon, L. L. M.; Samborskiy, D. V.; Sidorov, I. A.; Sola, I.; Ziebuhr, J.; Coronaviridae Study Group of the International Committee on Taxonomy of, V., The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology* **2020,** *5*, 536-544. DOI: 10.1038/s41564-020-0695-z.

2. WHO Pneumonia of unknown cause – China. https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/ (accessed May 6th, 2020).

3. Ratia, K.; Pegan, S.; Takayama, J.; Sleeman, K.; Coughlin, M.; Baliji, S.; Chaudhuri, R.; Fu, W.; Prabhakar, B. S.; Johnson, M. E.; Baker, S. C.; Ghosh, A. K.; Mesecar, A. D., A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proc Natl Acad Sci U S A* **2008,** *105*, 16119-16124. DOI: 10.1073/pnas.0805240105.

4. CDC People at Risk for Serious Illness from COVID-19. https://www.cdc.gov/coronavirus/2019-ncov/specific-groups/high-risk-complications.html (accessed May 6th, 2020).

5. Fehr, A. R.; Perlman, S., Coronaviruses: an overview of their replication and pathogenesis. *Methods in molecular biology (Clifton, N.J.)* **2015,** *1282*, 1-23. DOI: 10.1007/978-1-4939-2438-7_1.

6. Barretto, N.; Jukneliene, D.; Ratia, K.; Chen, Z.; Mesecar, A. D.; Baker, S. C., The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. *J Virol* **2005,** *79*, 15189-15198. DOI: 10.1128/JVI.79.24.15189-15198.2005.

7. Mielech, A. M.; Deng, X.; Chen, Y.; Kindler, E.; Wheeler, D. L.; Mesecar, A. D.; Thiel, V.; Perlman, S.; Baker, S. C., Murine Coronavirus Ubiquitin-Like Domain Is Important for Papain-Like Protease Stability and Viral Pathogenesis. *Journal of Virology* **2015,** *89*, 4907-4917. DOI: 10.1128/jvi.00338-15.

8. Daczkowski, C. M.; Dzimianski, J. V.; Clasman, J. R.; Goodwin, O.; Mesecar, A. D.; Pegan, S. D., Structural Insights into the Interaction of Coronavirus Papain-Like Proteases and Interferon-Stimulated Gene Product 15 from Different Species. *J Mol Biol* **2017,** *429*, 1661-1683. DOI: 10.1016/j.jmb.2017.04.011.

9. Ratia, K.; Saikatendu, K. S.; Santarsiero, B. D.; Barretto, N.; Baker, S. C.; Stevens, R. C.; Mesecar, A. D., Severe acute respiratory syndrome coronavirus papain-like protease: Structure of a viral deubiquitinating enzyme. *Proceedings of the National Academy of Sciences* **2006,** *103*, 5717-5722. DOI: 10.1073/pnas.0510851103.

10. Zhao, C.; Denison, C.; Huibregtse, J. M.; Gygi, S.; Krug, R. M., Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc Natl Acad Sci U S A* **2005,** *102*, 10200-10205.

11. Davis, M. E.; Gack, M. U., Ubiquitination in the antiviral immune response. *Virology* **2015,** *479-480*, 52-65. DOI: 10.1016/j.virol.2015.02.033.

12. Sadler, A. J.; Williams, B. R., Interferon-inducible antiviral effectors. *Nat Rev Immunol* **2008,** *8*, 559-568. DOI: 10.1038/nri2314.

13. Frieman, M.; Ratia, K.; Johnston, R. E.; Mesecar, A. D.; Baric, R. S., Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling. *J Virol* **2009,** *83*, 6689-6705. DOI: 10.1128/JVI.02220-08.

14. Devaraj, S. G.; Wang, N.; Chen, Z.; Chen, Z.; Tseng, M.; Barretto, N.; Lin, R.; Peters, C. J.; Tseng, C.-T. K.; Baker, S. C.; Li, K., Regulation of IRF-3-dependent Innate Immunity by the Papain-like Protease Domain of the Severe Acute Respiratory Syndrome Coronavirus. *Journal of Biological Chemistry* **2007,** *282*, 32208-32221. DOI: 10.1074/jbc.M704870200.

15. Scholte, F. E. M.; Zivcec, M.; Dzimianski, J. V.; Deaton, M. K.; Spengler, J. R.; Welch, S. R.; Nichol, S. T.; Pegan, S. D.; Spiropoulou, C. F.; Bergeron, E., Crimean-Congo Hemorrhagic Fever Virus Suppresses Innate Immune Responses via a Ubiquitin and ISG15 Specific Protease. *Cell Rep* **2017,** *20*, 2396-2407. DOI: 10.1016/j.celrep.2017.08.040.

16. Zhao, C.; Sridharan, H.; Chen, R.; Baker, D. P.; Wang, S.; Krug, R. M., Influenza B virus non-structural protein 1 counteracts ISG15 antiviral activity by sequestering ISGylated viral proteins. *Nat Commun* **2016,** *7*, 12754. DOI: 10.1038/ncomms12754.

17. Dzimianski, J. V.; Scholte, F. E. M.; Bergeron, E.; Pegan, S. D., ISG15: It's Complicated. *J Mol Biol* **2019,** *431*, 4203-4216. DOI: 10.1016/j.jmb.2019.03.013.

18. Bailey-Elkin, B. A.; Knaap, R. C.; Johnson, G. G.; Dalebout, T. J.; Ninaber, D. K.; van Kasteren, P. B.; Bredenbeek, P. J.; Snijder, E. J.; Kikkert, M.; Mark, B. L., Crystal structure of the Middle East respiratory syndrome coronavirus (MERS-CoV) papain-like protease bound to ubiquitin facilitates targeted disruption of deubiquitinating activity to demonstrate its role in innate immune suppression. *J Biol Chem* **2014,** *289*, 34667-34682. DOI: 10.1074/jbc.M114.609644.

19. Clasman, J. R.; Everett, R. K.; Srinivasan, K.; Mesecar, A. D., Decoupling deISGylating and deubiquitinating activities of the MERS virus papain-like protease. *Antiviral Res* **2020,** *174*, 104661. DOI: 10.1016/j.antiviral.2019.104661.

20. Daczkowski, C. M.; Goodwin, O. Y.; Dzimianski, J. V.; Farhat, J. J.; Pegan, S. D., Structurally Guided Removal of DeISGylase Biochemical Activity from Papain-Like Protease Originating from Middle East Respiratory Syndrome Coronavirus. *J Virol* **2017,** *91*. DOI: 10.1128/JVI.01067-17.

21. Deng, X.; Chen, Y.; Mielech, A. M.; Hackbart, M.; Kesely, K. R.; Mettelman, R. C.; O'Brien, A.; Chapman, M. E.; Mesecar, A. D.; Baker, S. C., Structure-Guided Mutagenesis Alters Deubiquitinating Activity and Attenuates Pathogenesis of a Murine Coronavirus. *J Virol* **2020**. DOI: 10.1128/JVI.01734-19.

22. Ratia, K.; Kilianski, A.; Baez-Santos, Y. M.; Baker, S. C.; Mesecar, A., Structural Basis for the Ubiquitin-Linkage Specificity and deISGylating Activity of SARS-CoV Papain-Like Protease. *PLoS Pathog* **2014,** *10*, e1004113. DOI: 10.1371/journal.ppat.1004113.

23. Frias-Staheli, N.; Giannakopoulos, N. V.; Kikkert, M.; Taylor, S. L.; Bridgen, A.; Paragas, J.; Richt, J. A.; Rowland, R. R.; Schmaljohn, C. S.; Lenschow, D. J.; Snijder, E. J.; Garcia-Sastre, A.; Virgin, H. W. t., Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. *Cell Host Microbe* **2007,** *2*, 404-416. DOI: 10.1016/j.chom.2007.09.014.

24. van Kasteren, P. B.; Bailey-Elkin, B. A.; James, T. W.; Ninaber, D. K.; Beugeling, C.; Khajehpour, M.; Snijder, E. J.; Mark, B. L.; Kikkert, M., Deubiquitinase function of arterivirus papain-like protease 2 suppresses the innate immune response in infected host cells. *Proc Natl Acad Sci U S A* **2013,** *110*, E838-847. DOI: 10.1073/pnas.1218464110.

25. Lee, H.; Lei, H.; Santarsiero, B. D.; Gatuz, J. L.; Cao, S.; Rice, A. J.; Patel, K.; Szypulinski, M. Z.; Ojeda, I.; Ghosh, A. K.; Johnson, M. E., Inhibitor recognition specificity of MERS-CoV papain-like protease may differ from that of SARS-CoV. *ACS Chem Biol* **2015,** *10*, 1456-1465. DOI: 10.1021/cb500917m.

26. Baez-Santos, Y. M.; Barraza, S. J.; Wilson, M. W.; Agius, M. P.; Mielech, A. M.; Davis, N. M.; Baker, S. C.; Larsen, S. D.; Mesecar, A. D., X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *J Med Chem* **2014,** *57*, 2393-2412. DOI: 10.1021/jm401712t.

27. Baez-Santos, Y. M.; St John, S. E.; Mesecar, A. D., The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral Res* **2015,** *115*, 21-38. DOI: 10.1016/j.antiviral.2014.12.015.

28. Ghosh, A. K.; Takayama, J.; Aubin, Y.; Ratia, K.; Chaudhuri, R.; Baez, Y.; Sleeman, K.; Coughlin, M.; Nichols, D. B.; Mulhearn, D. C.; Prabhakar, B. S.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D., Structure-based design, synthesis, and biological evaluation of a series of novel and reversible inhibitors for the severe acute respiratory syndrome-coronavirus papain-like protease. *J Med Chem* **2009,** *52*, 5228-5240. DOI: 10.1021/jm900611t.

29. Békés, M.; van der Heden van Noort, Gerbrand J.; Ekkebus, R.; Ovaa, H.; Huang, Tony T.; Lima, Christopher D., Recognition of Lys48-Linked Di-ubiquitin and Deubiquitinating Activities of the SARS Coronavirus Papain-like Protease. *Molecular Cell* **2016,** *62*, 572-585. DOI: 10.1016/j.molcel.2016.04.016.

30. Capodagli, G. C.; Deaton, M. K.; Baker, E. A.; Lumpkin, R. J.; Pegan, S. D., Diversity of Ubiquitin and ISG15 Specificity among Nairoviruses' Viral Ovarian Tumor Domain Proteases. *Journal of Virology* **2013,** *87*, 3815-3827 %! Diversity of Ubiquitin and ISG3815 Specificity among Nairoviruses' Viral Ovarian Tumor Domain Proteases.

31. Dzimianski, J. V.; Beldon, B. S.; Daczkowski, C. M.; Goodwin, O. Y.; Scholte, F. E. M.; Bergeron, E.; Pegan, S. D., Probing the impact of nairovirus genomic diversity on viral ovarian tumor domain protease (vOTU) structure and deubiquitinase activity. *PLoS Pathog* **2019,** *15*, e1007515. DOI: 10.1371/journal.ppat.1007515.

32. Chen, Y.; Savinov, S. N.; Mielech, A. M.; Cao, T.; Baker, S. C.; Mesecar, A. D., X-ray Structural and Functional Studies of the Three Tandemly Linked Domains of Non-structural Protein 3 (nsp3) from Murine Hepatitis Virus Reveal Conserved Functions. *J Biol Chem* **2015,** *290*, 25293-25306. DOI: 10.1074/jbc.M115.662130.

33. Langley, C.; Goodwin, O.; Dzimianski, J. V.; Daczkowski, C. M.; Pegan, S. D., Structure of interferon-stimulated gene product 15 (ISG15) from the bat species Myotis davidii and the impact of interdomain ISG15 interactions on viral protein engagement. *Acta Crystallogr D Struct Biol* **2019,** *75*, 21-31. DOI: 10.1107/S2059798318015322.

34. Daczkowski, C. M.; Goodwin, O. Y.; Dzimianski, J. V.; Farhat, J. J.; Pegan, S. D., Structurally Guided Removal of DeISGylase Biochemical Activity from Papain-Like Protease Originating from Middle East Respiratory Syndrome Coronavirus. *J Virol* **2017,** *91*. DOI: 10.1128/JVI.01067-17.

35. Deaton, M. K.; Dzimianski, J. V.; Daczkowski, C. M.; Whitney, G. K.; Mank, N. J.; Parham, M. M.; Bergeron, E.; Pegan, S. D., Biochemical and Structural Insights into the Preference of Nairoviral DeISGylases for Interferon-Stimulated Gene Product 15 Originating from Certain Species. *J Virol* **2016,** *90*, 8314-8327. DOI: 10.1128/JVI.00975-16.

36. Deaton, M. K.; Spear, A.; Faaberg, K. S.; Pegan, S. D., The vOTU domain of highly-pathogenic porcine reproductive and respiratory syndrome virus displays a differential substrate preference. *Virology* **2014,** *454-455*, 247-253. DOI: 10.1016/j.virol.2014.02.026.

37. Dzimianski, J. V.; Scholte, F. E. M.; Williams, I. L.; Langley, C.; Freitas, B. T.; Spengler, J. R.; Bergeron, E.; Pegan, S. D., Determining the molecular drivers of species-specific interferon-stimulated gene product 15 interactions with nairovirus ovarian tumor domain proteases. *PLoS One* **2019,** *14*, e0226415. DOI: 10.1371/journal.pone.0226415.

38. Baez-Santos, Y. M.; Mielech, A. M.; Deng, X.; Baker, S.; Mesecar, A. D., Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus. *J Virol* **2014,** *88*, 12511-12527. DOI: 10.1128/JVI.01294-14.

39. Kaiser, S. E.; Riley, B. E.; Shaler, T. A.; Trevino, R. S.; Becker, C. H.; Schulman, H.; Kopito, R. R., Protein standard absolute quantification (PSAQ) method for the

measurement of cellular ubiquitin pools. *Nat Methods* **2011,** *8*, 691-696. DOI: 10.1038/nmeth.1649.

40. Sridharan, H.; Zhao, C.; Krug, R. M., Species specificity of the NS1 protein of influenza B virus: NS1 binds only human and non-human primate ubiquitin-like ISG15 proteins. *J Biol Chem* **2010,** *285*, 7852-7856. DOI: 10.1074/jbc.C109.095703.

41. Versteeg, G. A.; Hale, B. G.; van Boheemen, S.; Wolff, T.; Lenschow, D. J.; García-Sastre, A., Species-specific antagonism of host ISGylation by the influenza B virus NS1 protein. *Journal of Virology* **2010,** *84*, 5423-5430. DOI: 10.1128/JVI.02395-09.

42. van Doremalen, N.; Schafer, A.; Menachery, V. D.; Letko, M.; Bushmaker, T.; Fischer, R. J.; Figueroa, D. M.; Hanley, P. W.; Saturday, G.; Baric, R. S.; Munster, V. J., SARS-Like Coronavirus WIV1-CoV Does Not Replicate in Egyptian Fruit Bats (Rousettus aegyptiacus). *Viruses* **2018,** *10*. DOI: 10.3390/v10120727.

43. Wang, P. F.; Li, Y.; Qian, Z. H.; Li, J. X.; Ge, X. J., Isolation and characterization of microsatellite loci from Pterocarya stenoptera (Juglandaceae). *Appl Plant Sci* **2018,** *6*, e01205. DOI: 10.1002/aps3.1205.

44. Jiang, T.; Feng, J. Myotis davidii. The IUCN Red List of Threatened Species 2019: e.T136250A22003049. https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T136250A22003049.en (accessed May 6th, 2020).

45. Agbowuro, A. A.; Huston, W. M.; Gamble, A. B.; Tyndall, J. D. A., Proteases and protease inhibitors in infectious diseases. *Med Res Rev* **2018,** *38*, 1295-1331. DOI: 10.1002/med.21475.

46. Kim, Y.; Liu, H.; Galasiti Kankanamalage, A. C.; Weerasekara, S.; Hua, D. H.; Groutas, W. C.; Chang, K. O.; Pedersen, N. C., Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor. *PLoS Pathog* **2016,** *12*, e1005531. DOI: 10.1371/journal.ppat.1005531.

47. Webb, B.; Sali, A., Protein Structure Modeling with MODELLER. *Methods Mol Biol* **2017,** *1654*, 39-54. DOI: 10.1007/978-1-4939-7231-9_4.

48. Harcourt, J.; Tamin, A.; Lu, X.; Kamili, S.; Sakthivel, S. K.; Murray, J.; Queen, K.; Tao, Y.; Paden, C. R.; Zhang, J.; Li, Y.; Uehara, A.; Wang, H.; Goldsmith, C.; Bullock, H., A.; Wang, L.; Whitaker, B.; Lynch, B.; Gautam, R.; Schindewolf, C.; Lokugamage, K. G.; Scharton, D.; Plante, J. A.; Mirchandani, D.; Widen, S. G.; Narayanan, K.; Makino, S.; Ksiazek, T. G.; Plante, K. S.; Weaver, S. C.; Lindstrom, S.; Tong, S.; Menachery, V. D.; Thornburg, N. J., Severe Acute Respiratory Syndrome Coronavirus 2 from Patient with 2019 Novel Coronavirus Disease, United States. *Emerging Infectious Disease journal* **2020,** *26*. DOI: 10.3201/eid2606.200516.

49. Zhang, Y.; Wei, Y.; Li, J.; Li, J., Development and optimization of a direct plaque assay for human and avian metapneumoviruses. *J Virol Methods* **2012,** *185*, 61-68. DOI: 10.1016/j.jviromet.2012.05.030.

CHAPTER 4

DEVELOPING NON-COVALENT PROTEASE INHIBITORS FOR THE TREATMENT OF SARS AND SARS-LIKE CORONAVIRUSES[3]

[4]Freitas, B.T., D.A. Ahiadorme, R.S. Bagul, I.A. Durie, J. Hill. N.E. Kramer, J. Murray, B.M. O'Boyle, E. Onobun, M.G. Pirrone, J.D. Shepard, Y. Subedi, K. Upadhyaya, R.J. Tripp, B.S. Cummings, D.C. Crich, and S.D. Pegan. To be submitted to *Nature Communications*

Abstract

Over the last 20 years, both severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1) and SARS-CoV-2 spread from animal hosts to humans causing outbreaks of severe disease. Both viruses originate from a subgroup of betacoronaviruses known as subgroup 2b. The emergence of two dangerous human pathogens from this subgroup along with previous studies illustrating the potential of other subgroup 2b members to cross over to humans has underscored the need for antiviral development against them. Coronaviruses suppress their host's innate immune response in part through the reversal of ubiquitination and ISGylation with their papain-like protease (PLpro). To identify unique or overarching subgroup 2b structural features or enzymatic tendencies, the PLpro from a subgroup 2b bat coronavirus, BtSCoV-Rf1.2004 was biochemically and structurally evaluated. This revealed PLpros from subgroup 2b coronaviruses tend to have narrow substrate specificity for K48 polyubiquitin and ISG15 originating from certain species. The PLpro of BtSCoV-Rf1.2004 was utilized as a tool alongside those of SARS-CoV-1 and SARS-CoV-2 to design 28 novel non-covalent drug-like pan subgroup 2b PLpro inhibitors, which included determining the effects of using previously unexplored core linkers within these compounds. Two crystal structures of BtSCoV-Rf1.2004 PLpro bound to these inhibitors aided in recognizing shared structural features among subgroup 2b proteases and compound design. By screening these three subgroup 2b PLpros against this novel set of inhibitors and performing cytotoxicity studies, new directions for pan-coronavirus subgroup 2b antiviral development through PLpro inhibition can be envisioned.

Introduction

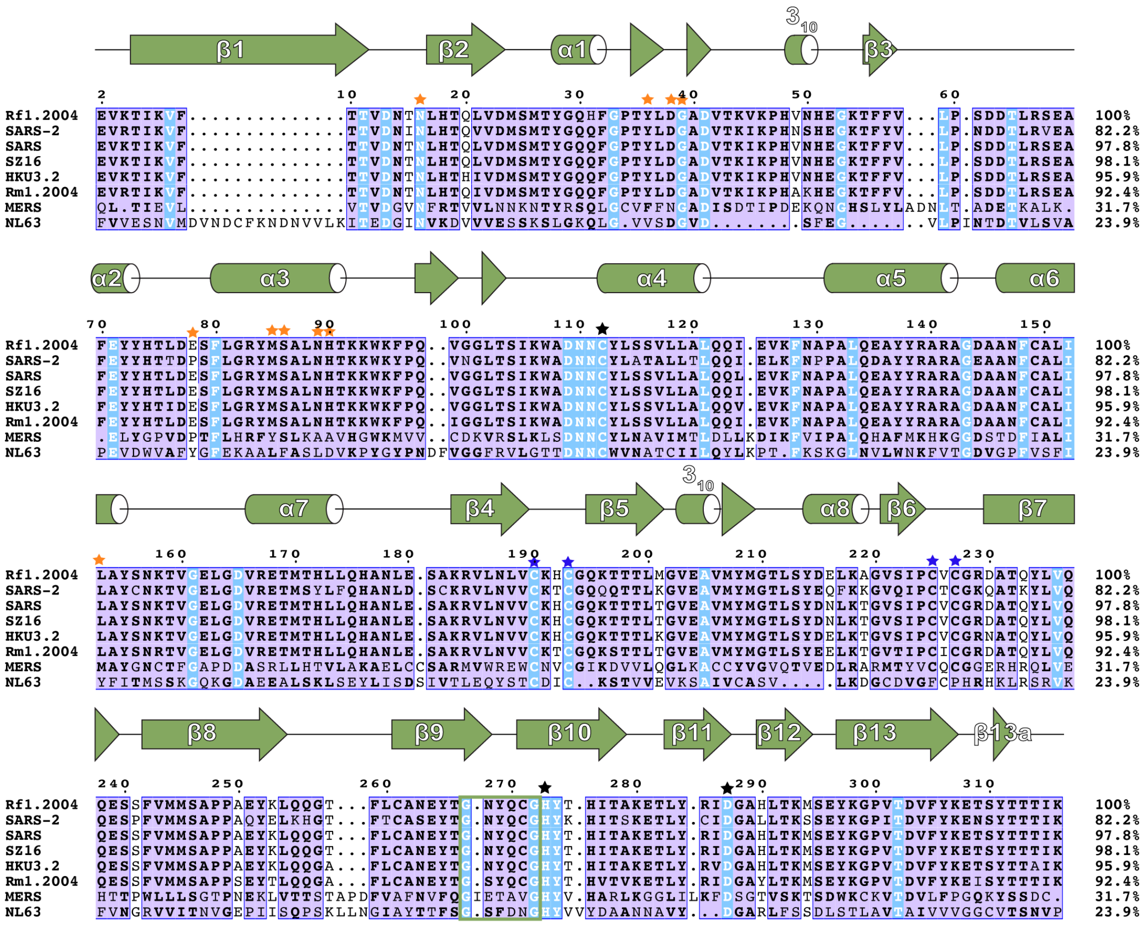
In the last two decades, betacoronaviruses have spilled over from animals to cause disease outbreaks in humans on multiple occasions with often deleterious outcomes[1-3]. Coronaviruses can be divided into four branches, and betacoronaviruses can be further broken down into four subgroups, 2a-2d. In 2015, a study highlighted subgroup 2b contained not only the 2003 pandemic causing SARS-CoV-1 but a cluster of 15 SARS-like coronaviruses[4]. Many of these SARS-like viruses are prevalent among Chinese horseshoe bats but also have been shown to infect a range of diverse hosts[2]. Additionally, these viruses were found to be similar enough that after replacing the spike protein of SARS-CoV-1 with that of another subgroup 2b virus from bats it was still capable of productively infecting primary human airway cells. They even could reach *in vitro* titers equivalent to epidemic strains of SARS-CoV-1[4]. Since then, this subgroup has been found to include over 30 human and zoonotic viruses. This list now includes SARS-CoV-2 that caused a worldwide pandemic with over 122 million cases and 2.7 million deaths[5]. As was foreshadowed in 2015, the emergence of SARS-CoV-2 further stresses that SARS-CoV-2 and SARS-CoV-1 are not likely to be the last coronaviruses to cross over to humans. The 2003 and 2019 pandemics highlight the urgent need to develop effective coronavirus therapeutics that can thwart current and future coronavirus subgroup 2b health threats.

As evidenced by the COVID-19 pandemic, the development and large-scale distribution of an effective vaccine targeting coronavirus spike proteins is possible but can take up to a year to field. Additionally, the two-shot regimens take more than 30 days to build full immunity and cannot be utilized by individuals with certain immune conditions, reducing their ability to quickly contain outbreaks[6]. Moreover, some of the most vulnerable populations such as the elderly, newborns, and the immunocompromised are typically unable to receive vaccines[7]. The spike protein has also shown a propensity for mutation, further stressing the need for alternative therapeutic strategies[8]. Hence, there is a strong need to add to the public's arsenal of small molecules that can serve as therapeutics for SARS-CoV-2 as well as the next subgroup 2b coronavirus outbreak.

As with other ssRNA(+) coronaviruses, betacoronaviruses upon infection translate two polypeptides pp1a and pp1ab, which are cleaved into 16 nonstructural proteins (Nsp1 to Nsp16). Nsp1-16 form the virus's membrane-bound replicase complex, which is necessary to transcribe the viral RNA genome prior to translation[9]. CoVs encode two proteases that process the polyproteins into their functional units[10]. The 3C-like protease, also known as the main protease, cleaves Nsp4-Nsp16. The papain-like protease (PLpro) cleaves Nsp1-3[11, 12]. The genomes of coronaviruses can contain up to two PLPs. However, the genomes of subgroup 2b viruses mirror that of Middle East respiratory syndrome (MERS) CoV, a subgroup 2c virus by encoding only one PLpro[10].

In addition to serving a critical function in cleaving the viral polypeptide, PLpros as well as papain-like protease 2 (PLP2) in two papain-like proteases, have demonstrated immunosuppressive effects on host organisms by reversing post-translational modification by ubiquitin (Ub) and interferon-stimulated gene product 15 (ISG15)[12, 13]. Several forms of ubiquitination and ISGylation have been shown to regulate various aspects of the innate immune system[14-16]. Post-translational modification with Ub and ISG15 has a profound impact on host type-I IFN and NF-κB inflammatory responses, as well as upregulating the production of cytokines, chemokines, and other ISGs[17, 18]. Additionally, some Ub-like protein (Ubl) modifiers have been shown to facilitate inhibition, sequestration, or proteasomal degradation of marked proteins[19-22]. By reversing these modifications viral encoded PLpros can allow critical viral proteins to remain intact, active, and suppress an anti-viral immune state until the virus is able to replicate. To determine if there are identifiable trends in enzymatic activity within subgroup 2b PLpros, we examined the PLpro from BtSCoV-Rf1.2004 for comparison. BtSCoV-Rf1.2004 is a subgroup 2b virus isolated from greater horseshoe bats and of the currently identified 2b viruses it is the most distantly related to SARS-CoV-1, other than SARS-CoV-2 (Figure 4.1)[4].

Figure 4.1. Sequence alignment of PLPs from SARS family coronaviruses. The PLpro or PLP from BtSCoV-Rf1.2004 (accession number ABD75321.1), SARS-CoV-2 (accession number MN908947.3), SARS-CoV-1 (accession number P0C6U8), SZ16 (accession number AY304488.1), HKU3.2 (accession number AAZ41328.1), Rm1.2004 (accession number ABD75330.1), MERS-CoV (accession number AFS88944), and NL63 (accession number P0C6U6.1). The secondary structure shown is predicted by DSSP for BtSCoV-Rf1.2004 PLpro. Similarity and alignment calculations were performed using ClustalW. Residue positions that are fully conserved are marked in blue, with those being highly conserved marked in purple. Residues that form the catalytic triad are marked with black stars, while residues forming the zinc finger motif are marked with blue stars. Residues forming interactions between the Ubl domain and thumb domain of PLpros are marked in orange. The BL2 loop is boxed in green.



The dual activity of PLpro, aiding viral replication and suppressing the innate immune system, makes it a highly sought-after drug target[23-26]. Its conserved nature among subgroup 2b coronaviruses also presents an opportunity to develop inhibitors that have broad specificity for future subgroup 2b viral threats. Several approaches have been taken to inhibit SARS-CoV-1 PLpro, including covalent modification of the catalytic cysteine by electrophilic warheads, as well as competitive inhibition by several non-covalently binding compounds[27]. This includes two sets of non-covalent drug-like naphthalene-containing compounds that bind to a pocket adjacent to the active site[28]. Excitingly, these naphthalene-based scaffolds originally identified to inhibit SARS-CoV-1 PLpro have been discovered to possess similar efficacy against SARS-CoV-2[29]. This subsequently has drawn significant attention to the potential of these scaffolds to serve as a basis for therapeutic development for SARS-CoV-2[24, 30, 31]. These inhibitors typically consist of a naphthyl group and an arene group with various backbones linking them. Thus far, the effective inhibitors can be categorized into two sets: series I compounds such as **GRL0617** (Figure 4.2A), which typically utilize an amide backbone, and series II compounds such as **6577871**, which commonly have a piperidine scaffold (Figure 4.2B)[27, 29, 32]. Both categories have proven effective in inhibiting PLpros from SARS-CoV-1 and SARS-CoV-2 but are not fully optimized in backbone structure or arene ring decoration. Additionally, it remains unknown whether these inhibitors would be effective against a wider array of coronavirus subgroup 2b PLpros.

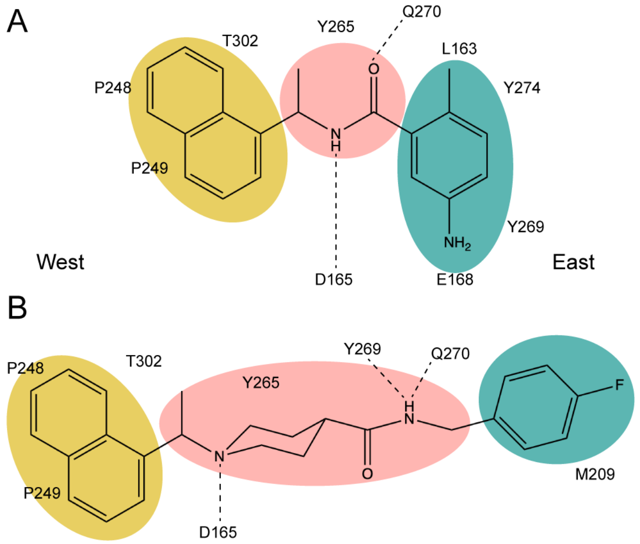


Figure 4.2. PLpro interactions with series I and II naphthalene-based inhibitors. Western naphthyl groups are highlighted in yellow, eastern arene groups are highlighted in teal, and core linker groups are highlighted in pink. (A) Key hydrophobic and hydrogen bonding interactions of series I inhibitor **GRL0617**; (B) Key hydrophobic and hydrogen bonding interactions of the series II initial hit **6577871**.

Here, we evaluated the enzymatic activity of the PLpro of BtSCoV-Rf1.2004 to reveal shared substrate preferences among subgroup 2b coronavirus PLpros. Two structures of BtSCoV-Rf12004 PLpro were solved in complex with non-covalent drug-like naphthalene inhibitors. From a global structural perspective, this new information revealed signature structural motifs that separate subgroup 2b PLpros from those of other coronaviruses. Additionally, these structures, in combination with structural activity relationship data obtained with a library composed of compounds with unexplored chemical cores, illustrate the potential of developing a novel inhibitor with pan subgroup 2b potency. Evaluation of these compounds' cytotoxicity against multiple cell lines further accentuates the potential of non-covalent drug-like naphthalene scaffolds for future therapeutic development against subgroup 2b members.

Results

PLpros are Highly Conserved among SARS-like Coronaviruses

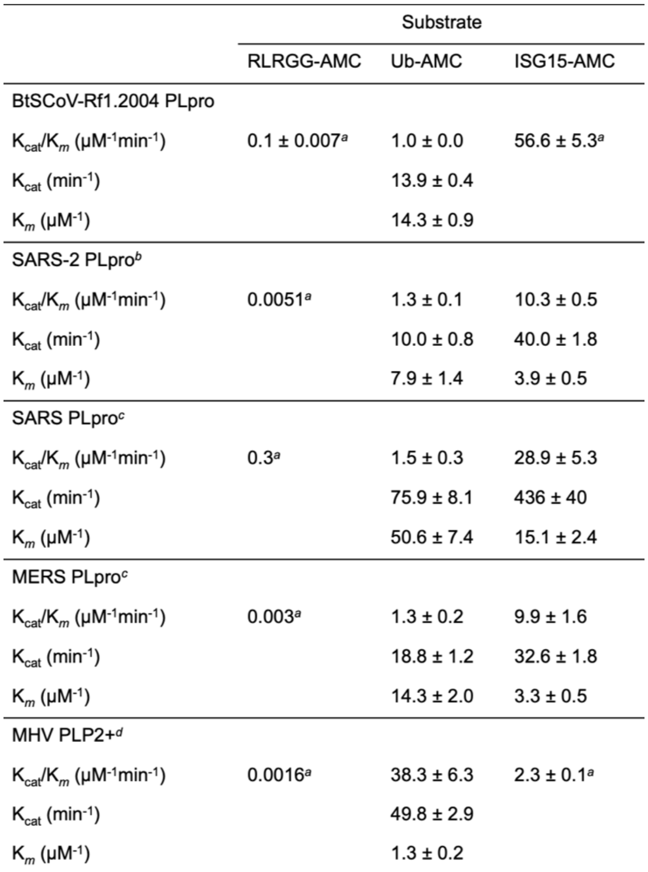
Prior to SARS-CoV-2, subgroup 2b consisted of fifteen SARS-like viruses originating from several species of old-world bats, primarily Horseshoe bats[2, 4]. Sequence analysis of these viruses along with that of SARS-CoV-2 revealed that each encoded one of six unique PLpro sequences (Figure 4.1). The recent SARS-CoV-2 strains originating from the United Kingdom and Brazil bring the total to eight, having A146D and K233Q PLpro mutations, respectively. As of this writing, the SARS-CoV-2 strain originating from South Africa has not been reported to have mutations within the PLpro[33]. Both the genomes and PLpros of subgroup 2b viruses share sequence identity greater than 80 percent. PLpros and PLP2 have been shown to vary in terms of substrate preference and activity. To gain a better understanding of subgroup 2b substrate preferences, a third subgroup 2b PLpro originating from BtSCoV-Rf1.2004 was selected to identify subgroup 2b DUB and deISGylase activity, as well as to assist in subgroup 2b broad-spectrum inhibitor development. BtSCoV-Rf1.2004 was isolated from the greater horseshoe bat (*R. ferrumequinum*) in Yichang, China, just 320 km from the location of the initial outbreak of the SARS-CoV-2 pandemic[2]. BtSCoV-Rf1.2004 was chosen because at 93.9% conservation it is the most distantly related subgroup 2b virus from SARS-CoV-1, other than SARS-CoV-2. Highlighting how conserved PLpro is among subgroup 2b viruses, despite the distant relation of BtSCoV-Rf1.2004 to SARS-CoV-1 and SARS-CoV-2, the PLpro of BtSCoV-Rf1.2004 retains 97.8% and 82.2% sequence homology, respectively. Initial examination of key functional areas, particularly the active site, naphthalene inhibitor binding pocket, and the Ub interacting motif (UIM) reveal that these areas are highly conserved among subgroup 2b PLpros (Figure 4.1). Despite three major emergent strains now observed, only two mutations within SARS-CoV-2 PLpro have been observed. Only one, K233Q of the Brazilian strain, is located in a functionally relevant location. However, this mutation can be viewed as conservative in nature as it converts the lysine residue to the subgroup 2b consensus glutamine residue.

Deubiquitinase and DeISGylase Activity of BtSCoV-Rf1.2004 PLpro

To identify differences in enzyme kinetics between BtSCoV-Rf1.2004 PLpro and those of other subgroup 2b viruses, the kinetic parameters of BtSCoV-Rf1.2004 PLpro for various substrates were determined. The kinetic values for PLpro substrates Ub-AMC and human ISG15-AMC, as well as the last five consensus amino acids between them (RLRGG; peptide-AMC), were determined by monitoring the release of the fluorogenic reporter group 7-amino-4-methylcoumarin (AMC) moiety (Table 4.1). The catalytic efficiency of BtSCoV-Rf1.2004 PLpro for peptide-AMC is

0.1±0.007 $\mu M^{-1}min^{-1}$, which is within the range of other subgroup 2b PLpros. The efficiency of BtSCoV-Rf1.2004 PLpro for Ub-AMC is 1.0±0.0$\mu M^{-1}min^{-1}$ with K_M and k_{cat} values of 14.3±0.9 and 13.9±0.4, respectively. Similar to other subgroup 2b viruses, BtSCoV-Rf1.2004 PLpro demonstrated a clear preference for ISG15 over Ub. BtSCoV-Rf1.2004 PLpro was not able to be saturated with ISG15-AMC at the concentrations used, however using first-order enzyme kinetics the catalytic efficiency was determined to be 56.6±5.3$\mu M^{-1}min^{-1}$. The difficulty in saturating the PLpro appears to be due to its robust deISGylase activity, which is twice as efficient as any other SARS-CoV PLpro. Despite not being able to calculate a maximum turnover number, a rate of 595 min^{-1} was observed. The calculated maximum turnover number for ISG15-AMC by SARS-CoV-1 is 436 min^{-1}, making BtSCoV-Rf1.2004 PLpro the most robust deISGylase among coronavirus PLpros evaluated so far[29, 34, 35].

Table 4.1. Kinetic analysis of BtSCoV-Rf1.2004 PLpro with PEP-AMC, Ub-AMC, and ISG15-AMC



	Substrate		
	RLRGG-AMC	Ub-AMC	ISG15-AMC
BtSCoV-Rf1.2004 PLpro			
K_{cat}/K_m ($\mu M^{-1} min^{-1}$)	0.1 ± 0.007[a]	1.0 ± 0.0	56.6 ± 5.3[a]
K_{cat} (min^{-1})		13.9 ± 0.4	
K_m (μM^{-1})		14.3 ± 0.9	
SARS-2 PLpro[b]			
K_{cat}/K_m ($\mu M^{-1} min^{-1}$)	0.0051[a]	1.3 ± 0.1	10.3 ± 0.5
K_{cat} (min^{-1})		10.0 ± 0.8	40.0 ± 1.8
K_m (μM^{-1})		7.9 ± 1.4	3.9 ± 0.5
SARS PLpro[c]			
K_{cat}/K_m ($\mu M^{-1} min^{-1}$)	0.3[a]	1.5 ± 0.3	28.9 ± 5.3
K_{cat} (min^{-1})		75.9 ± 8.1	436 ± 40
K_m (μM^{-1})		50.6 ± 7.4	15.1 ± 2.4
MERS PLpro[c]			
K_{cat}/K_m ($\mu M^{-1} min^{-1}$)	0.003[a]	1.3 ± 0.2	9.9 ± 1.6
K_{cat} (min^{-1})		18.8 ± 1.2	32.6 ± 1.8
K_m (μM^{-1})		14.3 ± 2.0	3.3 ± 0.5
MHV PLP2+[d]			
K_{cat}/K_m ($\mu M^{-1} min^{-1}$)	0.0016[a]	38.3 ± 6.3	2.3 ± 0.1[a]
K_{cat} (min^{-1})		49.8 ± 2.9	
K_m (μM^{-1})		1.3 ± 0.2	

Poly-Ub Linkage Preferences for SARS-CoV PLpro

PLpros have often shown a greater ability to process poly-Ub chains than those of mono-Ub conjugates[24]. To determine if this is the case with BtSCoV-Rf1.2004, PLpro cleavage activity was tested against the eight different linkage types of di-Ub M1, K6, K11, K27, K29, K33, K48, and K63. Utilizing similar experimental conditions to those used to evaluate other PLpros[29, 36], 10 μM of each di-Ub was incubated with 20 nM PLpro from BtSCoV-Rf1.2004. Divergent from PLpros and related PLP2s originating out of subgroup 2a and 2c, BtSCoV-Rf1.2004 demonstrated little ability to cleave di-Ub of any linkage type. Only a small amount of K48 was cleaved after 60 minutes by BtSCoV-Rf1.2004 PLpro (Figure 4.3A). This is similar to what was seen in assays using other SARS-CoV PLpros, which showed little to no activity toward di-Ub[29, 36]. However, when tested against tetra-Ub chains BtSCoV-Rf1.2004 PLpro demonstrated significantly higher DUB activity (Figure 4.3B). The K48 linked tetra-Ub was entirely converted to di-Ub in less than five minutes and well-defined bands of tri and di-Ub appear in the K63 tetra-Ub cleavage experiment after approximately forty minutes.

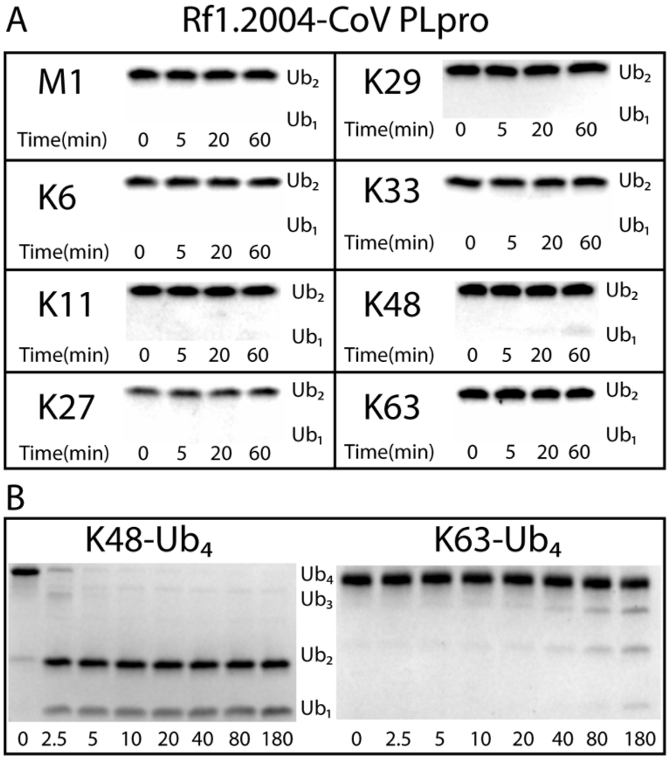
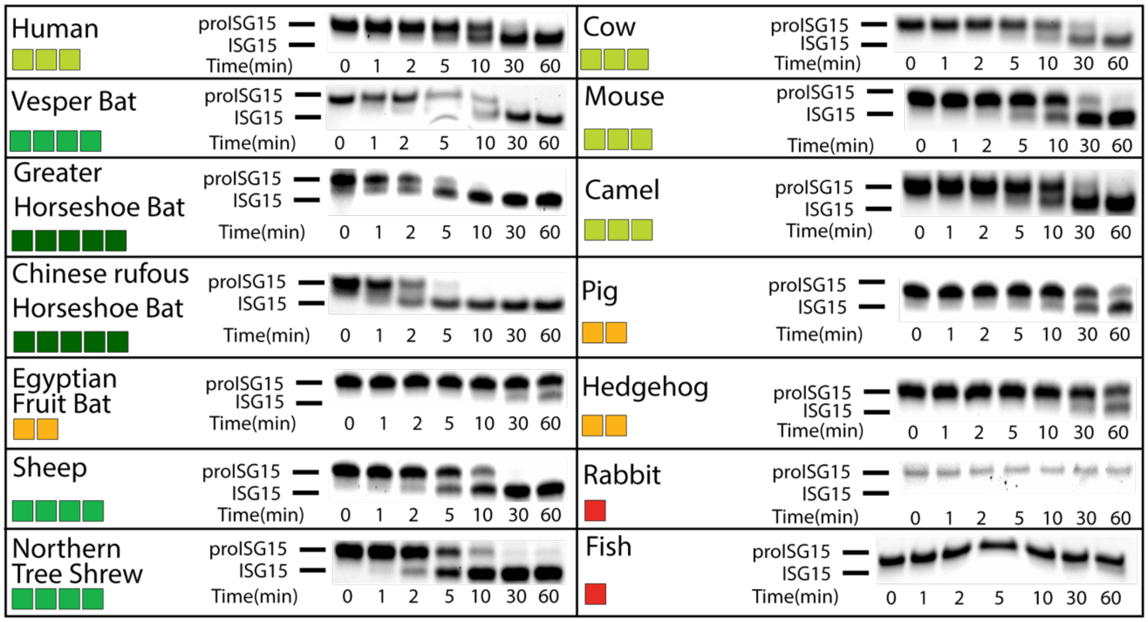


Figure 4.3. Activity of BtSCoV-Rf1.2004 PLpro for different linkages of poly-Ub. A) At 37°C 10µM each of M1, K6, K11, K27, K29, K33, K48, and K63 linked di-Ub were incubated with 20nM BtSCoV-Rf1.2004 PLpro. Samples were taken from the reaction tube at indicated time points. B) Under similar reaction conditions 13.65µM each of K48 and K63 linked tetra-Ub was incubated with 23nM PLpro for 3 hours with samples taken at given time points.

Species Specific DeISGylase Activity of BtSCoV-Rf1.2004 PLpro

DeISGylases have been shown to be selective for ISG15s from species which they have been found to productively infect[10, 29, 37-39]. Unlike Ub that is almost completely conserved between species, ISG15 can vary with sequence similarity as low as 60% within class Mammalia[37]. *In vivo* ISG15 is translated as a pro-form that consists of the mature ISG15 with several amino acids following the LRLRGG cleavage site (proISG15). Given BtSCoV-Rf1.2004 PLpro's similarity with other subgroup 2b PLpros in substrate preference for ISG15 over mono-Ub, molecular weight shift experiments using proISG15s from various species were used to assess ISG15 species preference among a collection of fourteen species[10, 29, 37]. The experimental parameters that were employed were consistent with previous studies focusing on other PLpros incubating 20 nM PLpro with 10 μM proISG15 originating from the various species (Figure 4.4)[10, 29].

Figure 4.4. Activity of BtSCoV-Rf1.2004 PLpro for proISG15 from multiple species. BtSCoV-Rf1.2004 PLpro was evaluated for the cleavage of proISG15s from the following species: human (*Homo sapiens*; AAH09507.1), cow (*Bos taurus*; NP_776791.1), vesper bat (*Myotis davidii*; ELK23605.1), Egyptian fruit bat (*Rousettus aegyptiacus*; XP_015999857.1), pig (*Sus scrofa*; ACB87600.1), hedgehog (*Erinaceus europaeus*; XP_007525810.2), mouse (*Mus musculus*; AAB02697.1), dromedary camel (*Camelus dromedarius*; XP_010997700.1), sheep (*Ovis aries*; AF152103.1), northern tree shrew (*Tupaia belangeri*; AFH66859.1), greater horseshoe bat (*Rhinolophus ferrumequinum*; XP_032969719.1), Chinese rufous horseshoe bat (*Rhinolophus sinicus*; XP_019567580.1), rabbit (*Oryctolagus cuniculus*; XP_017195918), and jackknife fish (*Oplegnathus fasciatus*; BAJ16365.1). At 37°C, 10μM of each ISG15 was incubated with 20nM of BtSCoV-Rf1.2004 PLpro for at least 1hr with samples taken at the time points indicated. The summary of the proISG15 cleavage assays for BtSCoV-Rf1.2004 PLpro is presented as a heat map. Colors range from dark red (no cleavage) to green (relatively robust cleavage).



In line with previous findings of other PLpros preferring ISG15s from species that serve as reservoirs for their parent virus, BtSCoV-Rf1.2004 PLpro distinctly possesses the highest activity towards ISG15s encoded by the greater horseshoe bat and the Chinese rufous horseshoe bat. Horseshoe bats are known reservoirs for subgroup 2b viruses and are present throughout much of Southeast Asia, sub-Saharan Africa, and Southern Europe[2]. BtSCoV-Rf1.2004 PLpro also has varying levels of proficiency for cleaving vesper bat, sheep, and northern tree shrew proISG15s with more moderate activity toward human, cow, camel, and mouse. Relatively weak activity towards pig, hedgehog, and Egyptian fruit bat was observed with no detectable activity against rabbit or fish ISG15. The cleavage rate of several important ISG15s such as human, mouse, and vesper bat by BtSCoV-Rf1.2004 PLpro closely matched what was seen with SARS-CoV-1 and SARS-CoV-2[10, 29]. While some species' ISG15s are cleaved more slowly than long-chain K48-Ub, those of BtSCoV-Rf1.2004's natural hosts *R. ferrumequinum* and *R. sinicus* appear to be the preferred substrates.

X-ray crystal structure of betacoronavirus subgroup 2b BtSCoV-Rf1.2004 PLpro

To evaluate any differences between subgroup 2b PLpros an X-ray crystal structure of the BtSCoV-Rf1.2004 PLpro was obtained and compared to structures of SARS-CoV-1 and SARS-CoV-2 PLpros. The structure was determined to a resolution of 3.16Å in space group $P2_12_12$ (Table S4.1). A homology model of BtSCoV-Rf1.2004 PLpro based on a SARS-CoV-1 PLpro (PDB 3E9S) catalytic core was used as a search model. Upon finding two monomers of the catalytic core in the asymmetric unit, the Ubl domain for the monomers was subsequently located using the Ubl of 3E9S as a search model.

Globally, the BtSCoV-Rf1.2004 PLpro resembles other subgroup 2b PLpros in secondary and tertiary structure (Figure 4.5A/B). It consists of a catalytic core made up of a palm, thumb, and zinc finger domain. The core was found in a holo, open conformation[32]. Additionally, the PLpros contain an N-terminal Ubl domain. The Ubl domain of BtSCoV-Rf1.2004 is shifted approximately 90 degrees from previously seen elongated forms[10] and is tucked against the catalytic domain when compared to the typical extended conformation (Figure S4.1). Until now this tucked conformation had only been seen in SARS-CoV-1 PLpro in complex with mouse ISG15 (PDB 5TL7). When in this conformation a seam of electrostatic interactions is formed between residues on the Ubl domain and α-helix 3 of the thumb domain. These interactions appear to stabilize the association of the two domains.

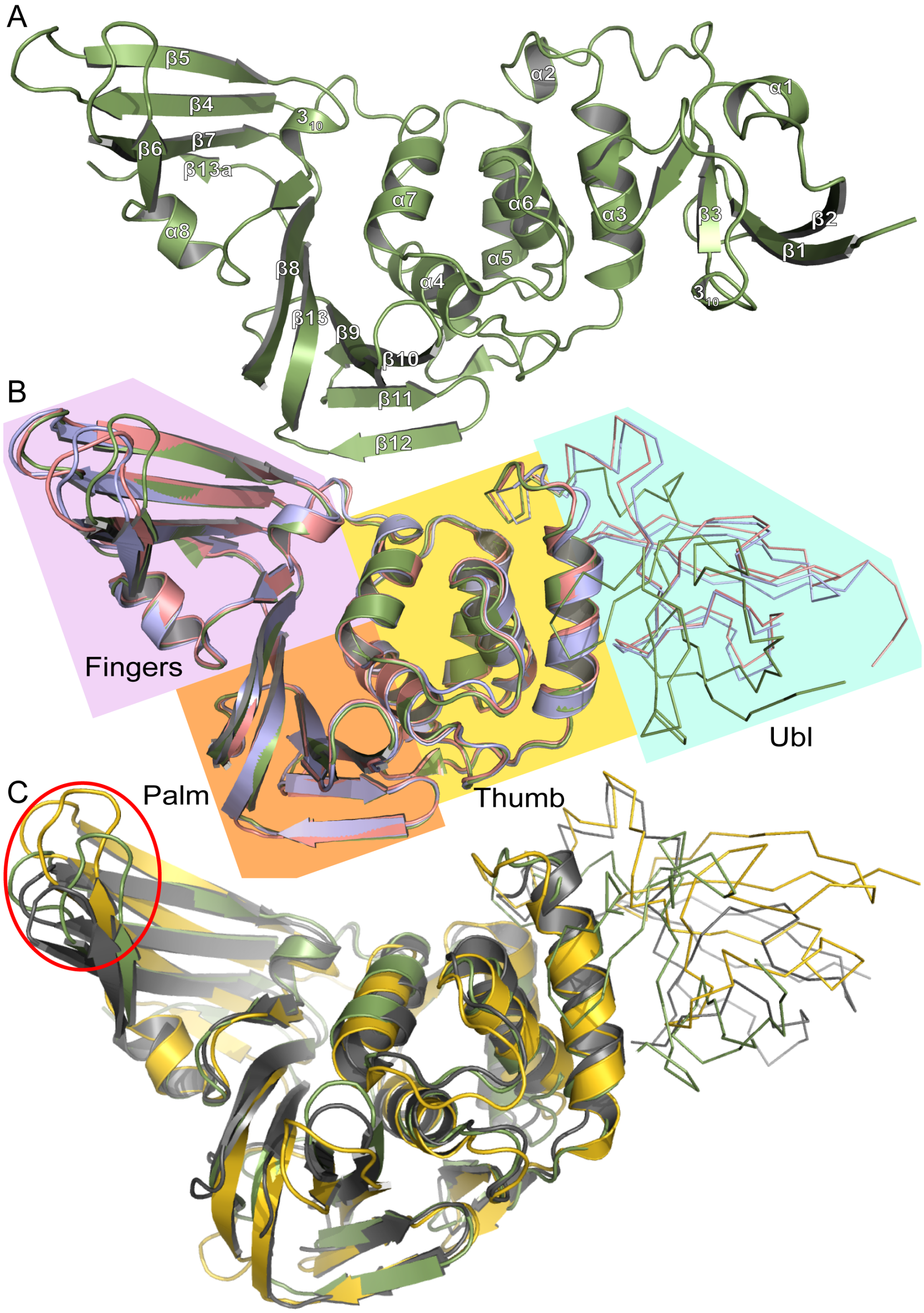


Figure 4.5. Tertiary structure of subgroup 2b PLpros compared to subgroups 2a and 2c A) Cartoon representation of BtSCoV-Rf1.2004 PLpro (Green) secondary structure with helix and sheet labels corresponding to Figure 4.2 DSSP calculations B) Overlaid cartoon representations of BtSCoV-Rf1.2004 PLpro (Green), SARS-CoV-1 PLpro (PDB 3E9S) (Pink), and SARS-CoV-2 PLpro (PDB 7JIR) (Blue). The four PLpro domains are labeled and color coded: Fingers (Purple), Palm (Orange), Thumb (Yellow), Ubl (Cyan). C) Overlaid cartoon representations of BtSCoV-Rf1.2004 PLpro (Green), MERS-CoV PLpro (PDB 5W8T)(Yellow), and MHV PLP (5WFI)(Grey) with their Ubl domains represented by ribbons.

When compared with the PLP of the subgroup 2a MHV (PDB 5WFI) and the PLpro of subgroup 2c MERS-CoV (PDB 5W8T), it appears that most secondary structures are conserved, however, there is variation in the overall tertiary structure between the PLpros of coronavirus subgroups (Figure 4.5C). The finger domains vary the most between subgroups, with not only loops shifting, but β-sheets orienting differently. This is most pronounced in the MERS-CoV PLpro, where the finger domain is shifted approximately 6 Å away from the P3/P4 pocket relative to MHV and BtSCoV-Rf1.2004. This variance seems to go beyond the open and closed conformations of the finger domain induced by substrate binding previously observed[40, 41].

Interactions between GRL0617 and the P3-P4 pocket of BtSCoV-Rf1.2004 PLpro

When examining the BL2 loop F_o - F_c density was readily observed for **GRL0617** in both monomers (Figure 4.6A). Within the pocket **GRL0617** is oriented similarly in all subgroup 2b PLpros, forming hydrogen bonds with D165 and Q270, as well as hydrophobic interactions with P248 and P249 (Figure 4.6B/C). However, between the monomers of BtSCoV-Rf1.2004, there are differences in the positions of the Q270 side chain, the eastern arene ring, and most notably the western naphthyl group (Figure 4.7C). The differences appear to be due to a crystal contact present on the B chain that is absent on the A chain in which V226 of the B chain zinc finger forms a hydrophobic interaction with the western naphthyl group of B chain **GRL0617** (Figure 4.7A/B).

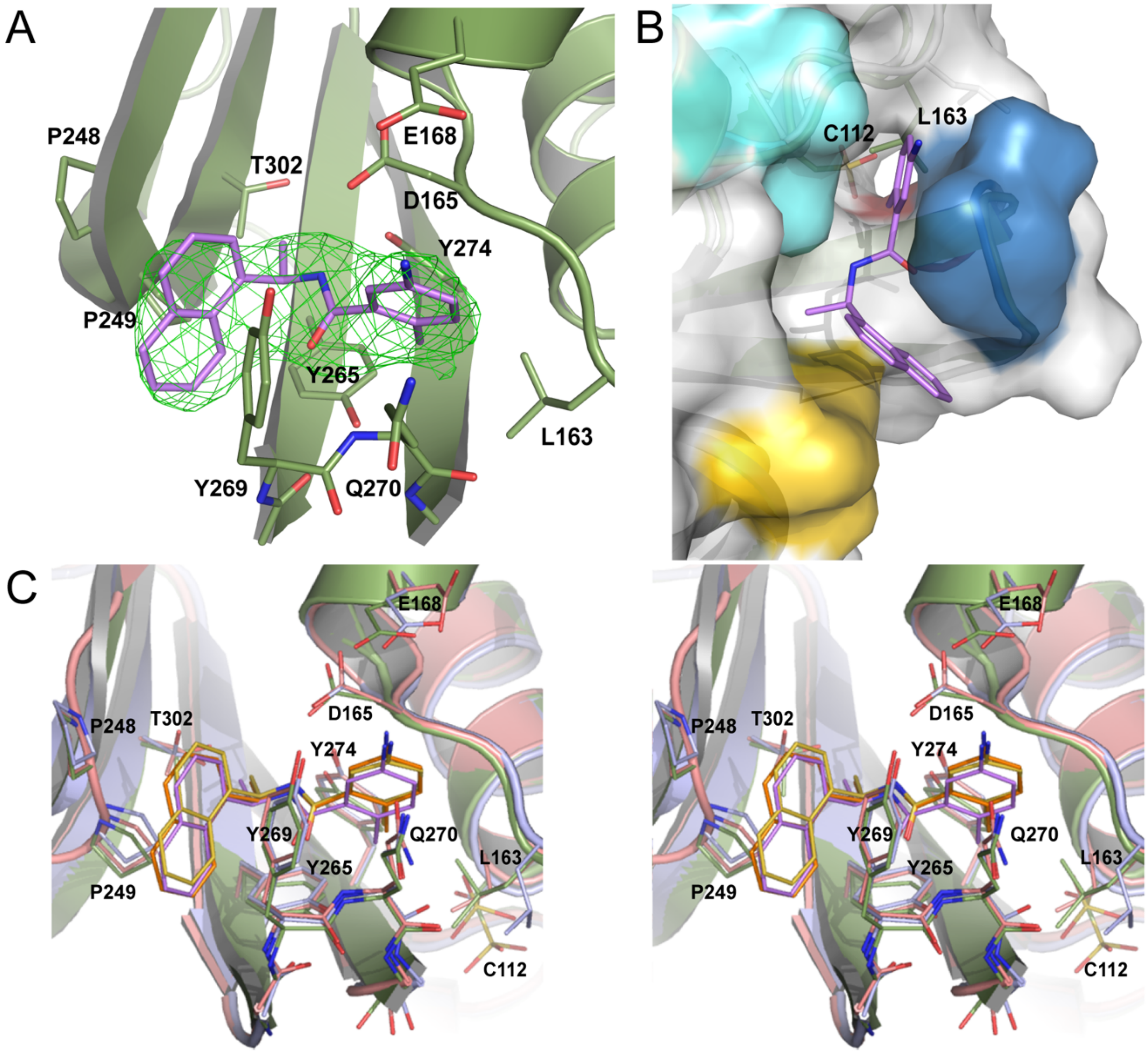


Figure 4.6. Inhibitor binding pocket of three subgroup 2b viruses. A) A F_o-F_c electron density map is shown contoured at 1σ (green mesh). With **GRL0617** shown in purple and BtSCoV-Rf1 PLpro shown in green. B) **GRL0617** (purple) bound to BtSCoV-Rf1 PLpro (green cartoon) overlaid with SARS-CoV-2 (white surface and cartoon) showing a possible path to active site for future inhibitors. The surfaces of the P2 site, P3 site, P4 site, and BL2 loop are shown in red, cyan, yellow, and blue respectively. C) Stereoview overlay of **GRL0617** bound to three different SARS-CoV PLpros: BtSCoV-Rf1 (green and purple), SARS-CoV-1 (pink and orange), and SARS-CoV-2 (blue and yellow).

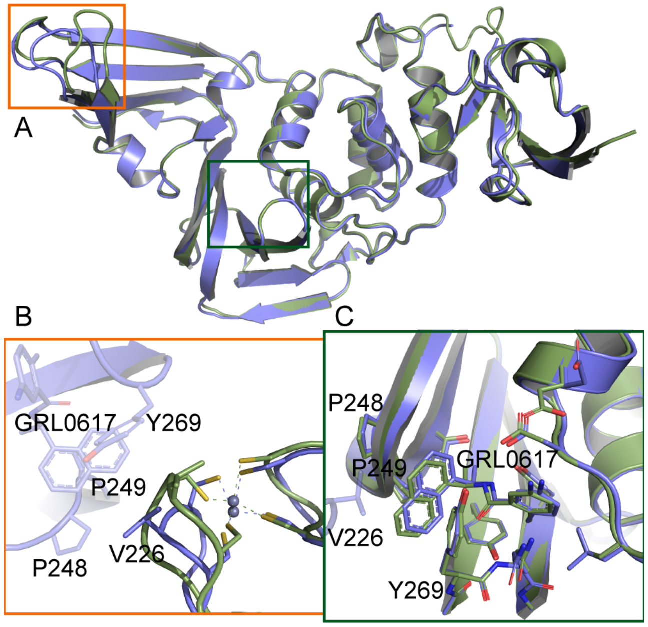


Figure 4.7. Crystal contacts affecting PLpro conformation A) Overlaid cartoon representations of BtSCoV-Rf1.2004 PLpro chain A (green) and chain B (blue). B) Zinc finger loops of the two domains differ due to a crystal contact being made by chain B. C) **GRL0617** of chain B is shifted in the binding pocket due to a crystal contact with chain B V226.

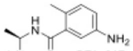
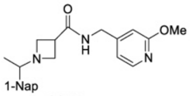
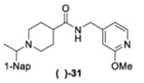
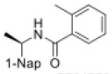
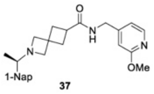
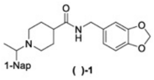
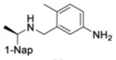
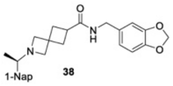
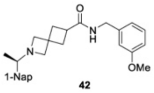
Crystal contacts on the BL2 loops of SARS-CoV-1 and SARS-CoV-2 PLpros influence the positioning of the loops and **GRL0617** in these structures as well. The absence of a BL2 loop contact on the A chain of BtSCoV-Rf1.2004 PLpro likely places **GRL0617** in its most natural orientation within the pocket. The chain B V226 contact is also responsible for a shift in the zinc finger loop (Figure 4.7). The zinc finger loop appears to be the most variable region within the catalytic core of the PLpro (Figure 4.5B). SARS-CoV-1 and SARS-CoV-2 PLpro structures form different crystal contacts at similar sites to those of BtSCoV-Rf1.2004 chain B that cause shifts in the zinc finger and BL2 loops. This suggests the influence of crystal should be considered when using these structures in a structure-based drug design approach.

Beyond the crystal lattice induced changes to the BL2 loop among the subgroup 2b PLpro structures, the only other notable difference is a rotamer in residue L163, which in SARS-CoV-1 and BtSCoV-Rf1.2004 is angled towards the 2-methyl group of **GRL0617**, but in SARS-CoV-2 angled away from the pocket (Figure 4.6B). This altered conformation is seen in structures of SARS-CoV-1 PLpro bound to human and mouse ISG15, along with a shifted BL2 loop to accommodate the C-terminal GG cleavage motif of Ubl substrates (PDB 5TL6 and 5TL7). These changes provide insight into the flexibility of the BL2 loop when accommodating BL2 inhibitors as well as how the influence of crystal contacts on the BL2 loop in some structures may need to be accounted for when using these structures for drug discovery purposes.

Inhibition of SARS-CoV Family PLpros with Novel Inhibitors

With **GRL0617** being readily accommodated within the BtSCoV-Rf1.2004 PLpro P3/P4 site, the practicality of targeting PLpro broadly appears feasible. As a result, 31 compounds were synthesized based on series I and series II compounds designed for the inhibition of SARS-CoV-1 PLpro. While variation in PLpro sequences exists among subgroup 2b viruses, the residues lining the active site and the adjacent P3 and P4 sites are fully conserved (Figure 4.1). These residues bind RLRGG residues of Ubl substrates to facilitate cleavage[13]. They have also been shown to be critical in binding naphthalene-based compounds that successfully inhibit PLpro cleavage activity in SARS-CoV-1 and SARS-CoV-2[32]. Preliminary studies show that inhibitors designed to bind at this site in SARS-CoV-1 exhibit similar IC_{50} and EC_{50} values when tested against SARS-CoV-2[29] (Table 4.2).

Table 4.2. IC$_{50}$ values of series I and II compounds when tested against PLpros of SARS-CoV-2, SARS-CoV-1, and BtSCoV-Rf1.2004 PLpros.

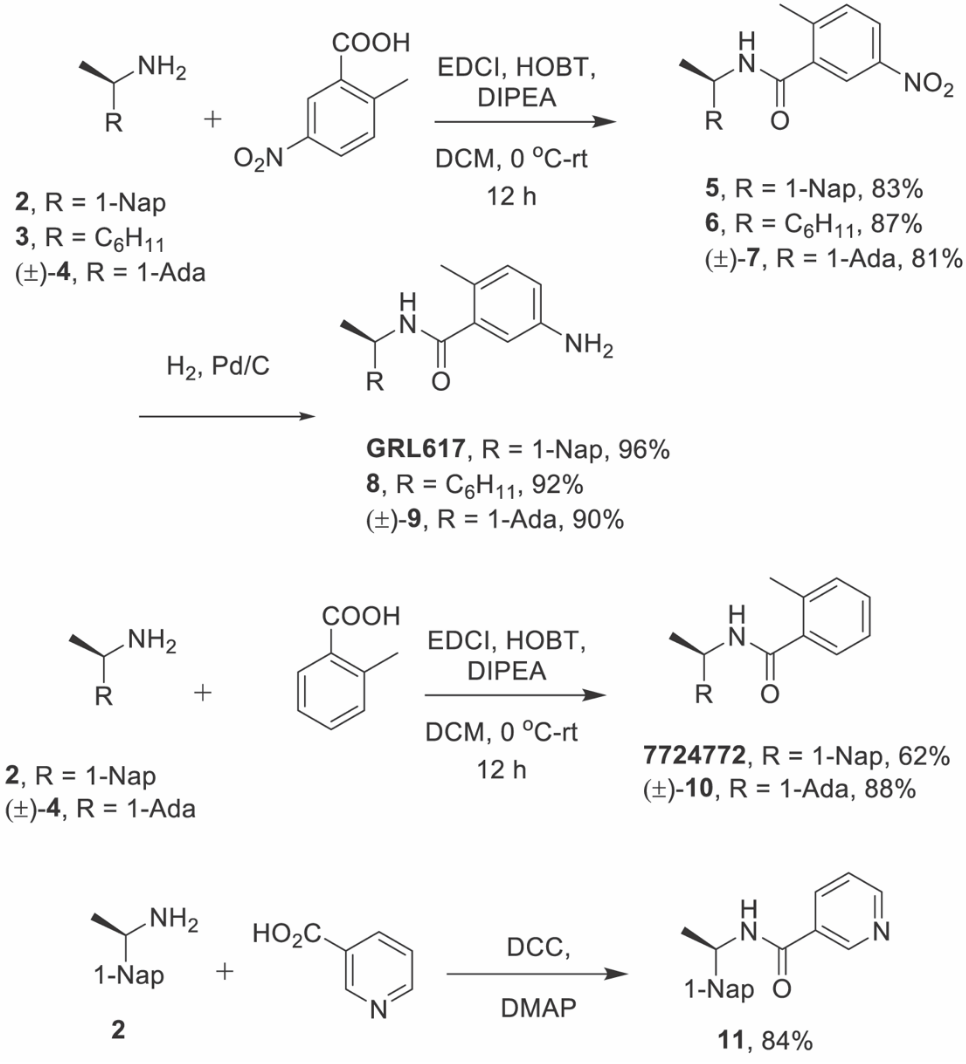
Compound	IC$_{50}$ (μM)	Compound	IC$_{50}$ (μM)	Compound	IC$_{50}$ (μM)
 GRL-617	2.4 ± 0.2 1.4 ± 0.0 1.2 ± 0.0	 ( )-34	27.3 ± 3.6 11.4 ± 0.5 15.8 ± 1.7	 ( )-31	1.1 ± 0.1 1.0 ± 0.0 8.4 ± 0.8
 7724772	23.5 ± 1.7 14.5 ± 0.6 26.2 ± 1.4	 37	5.4 ± 0.2 10.7 ± 0.7 7.5 ± 0.5	 ( )-1	1.8 ± 0.1 1.6 ± 0.0 2.4 ± 0.1
 26	41.1 ± 3.2 36.8 ± 2.3 26.2 ± 1.4	 38	19.6 ± 3.0 15.2 ± 1.3 21.0 ± 3.1	 42	28.6 ± 2.0 9.1 ± 1.1 22.5 ± 1.7

The naphthalene-based inhibitors can be categorized into two series by their backbone structure. The series I and II hits share a common binding site for their western hydrophobic moieties. The two series differ in the eastern arene moieties, which occupy proximal but different binding pockets, and by the central core – a simple amide unit in series I and a piperidine ring in series II (Figure 4.2). This suggested that there exists considerable scope for innovation in the central core, in terms of both the spacing it provides between the eastern and western arenes and the basic structure employed. With this in mind, we set out to explore alternative central cores based on piperidine analogs such as azetidines, and spiroazetidines[42], and other bicyclic systems that would allow for variation of the separation between the eastern and western arene while retaining the basic geometry of a 4-substituted piperidine ring.

The X-ray crystallographically-derived essential interactions between the series I compound **GRL0617** and the SARS-CoV-1 PLpro (PDB 3E9S) are summarized in Figure 4.2A. Early X-ray structures of a series II inhibitor **1** with SARS-CoV-1 PLpro (PDB 3MJ5) revealed the occupancy of the same hydrophobic pocket by the naphthalene ring but showed the piperidine ring in an unlikely high-energy twist-boat conformation with no interaction of either the basic ring nitrogen or the adjacent amide NH with D165.[43] A subsequent reevaluation of this X-ray structure, however, represented the piperidine ring in a more standard chair conformation with an H-bond from the basic piperidine nitrogen to D165.[44] Adopting the latter structure as the more likely, the strong parallels between the interactions of the western halves of the series I and II compounds with the protein are evident and form a solid base on which to engineer next generation compounds for the inhibition of subgroup 2b PLpros, including SARS-CoV-2. The aromatic rings of the eastern halves of the series I and II compounds occupy different sites because of the differing dimensions of their central portions. In series I, the backbone amide of Q270 H-bonds to the amide carbonyl of the inhibitor, whereas in series II the hydrocarbon part of the Q270 side chain forms one face of hydrophobic pocket encapsulating the eastern arene, which is completed by the side chain of M209 and other residues. Y269 interacts with the amino group in the eastern arene of the series I compound, but with the amide carbonyl of the series II compound.

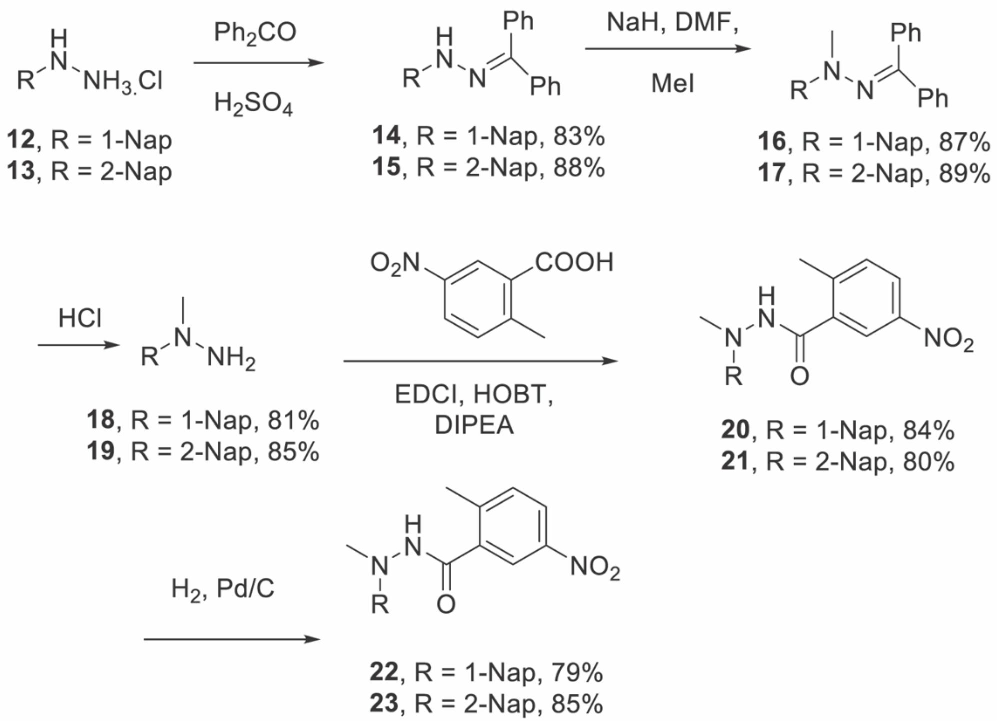
We began with a brief exploration of the hydrophobic eastern part of the binding domain by the synthesis of analogs of series I compounds in which the naphthalene unit was exchanged for a cyclohexyl or 1-adamantanyl moiety. We first prepared an authentic sample of **GRL0617** coupling *R*-1-(1-naphthyl)ethylamine **2** with 2-methyl-5-nitrobenzoic acid followed by hydrogenolytic reduction of the intermediate nitroarene **5**. Carbodiimide coupling of *R*-1-cyclohexylethylamine **3** with 2-methyl-5-nitrobenzoic acid then gave the amide **6**, which was reduced with hydrogen over palladium on charcoal to afford **8**. Similarly, (±)-1-(1-adamantanyl)ethylamine **4** was condensed with 2-methyl-5-nitrobenzoic acid to give **7**, hydrogenolysis of which afforded **9**. 2-Methylbenzoic acid also was coupled *R*-1-(1-naphthyl)ethylamine **2** and (±)-1-(1-adamantanyl)ethylamine **4** to give an authentic sample of (±)-**7724772**[45] and **10**. A further analog, **11** of **GRL0617** in which the 5-amino-2-methylbenzamide moiety was replaced by a nicotinamide group was also prepared at this time through carbodiimide coupling (Scheme 4.1).

Scheme 4.1. Synthesis of Compounds **GRL0617**, **8**, **9**, **7724772**, **10** and **11**.



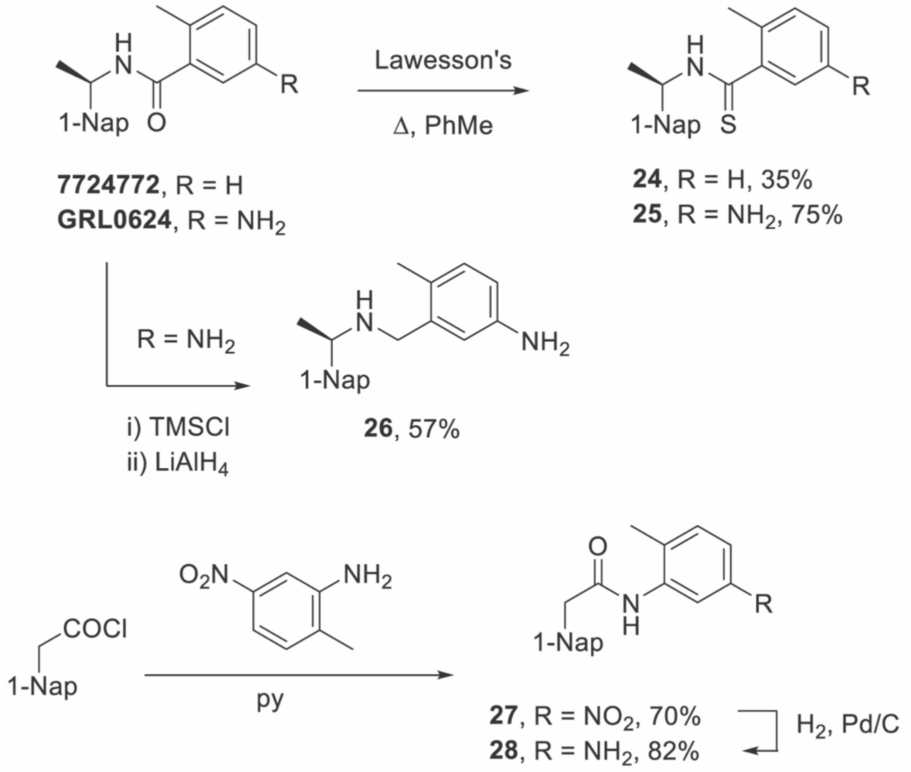
Focusing on series I, we next investigated the influence of absolute configuration at the stereogenic center in the amide portion of **GRL0617** by replacing the *R*-configured CHMe by an NMe moiety in the form of the hydrazide **22** (Scheme 4.2). For good measure we also accessed the regioisomer **23** by analogous means. The substitution of the *R*-configured CHMe by an NMe moiety in both **22** and **23** draws on analogy with our recent work with trisubstituted hydroxylamines,[46, 47] according to which we consider the *N*-methyl hydrazide moiety as an isostere of either enantiomer of the parent compound owing to the pyramidal yet rapidly inverting nature of the hydrazide nitrogen atom. The 1- and 2-naphthylhydrazines **12** and **13** required for these syntheses were accessed by a literature method.[48]

Scheme 4.2. Preparation of the Regioisomeric Hydrazides **22** and **23**.



Remaining with series I, we also prepared thioamides **24** and **25** by treatment of the parents with Lawesson's reagent, and reduced the amide in **GRL0617** by silylation and subsequent treatment with lithium aluminum hydride,[49] to the corresponding amine **26**. Finally, in this series, we prepared the inverted amide **28** by condensation of 1-naphthylacetyl chloride with 2-methyl-5-nitroaniline giving **27**, followed by hydrogenolysis (Scheme 4.3).

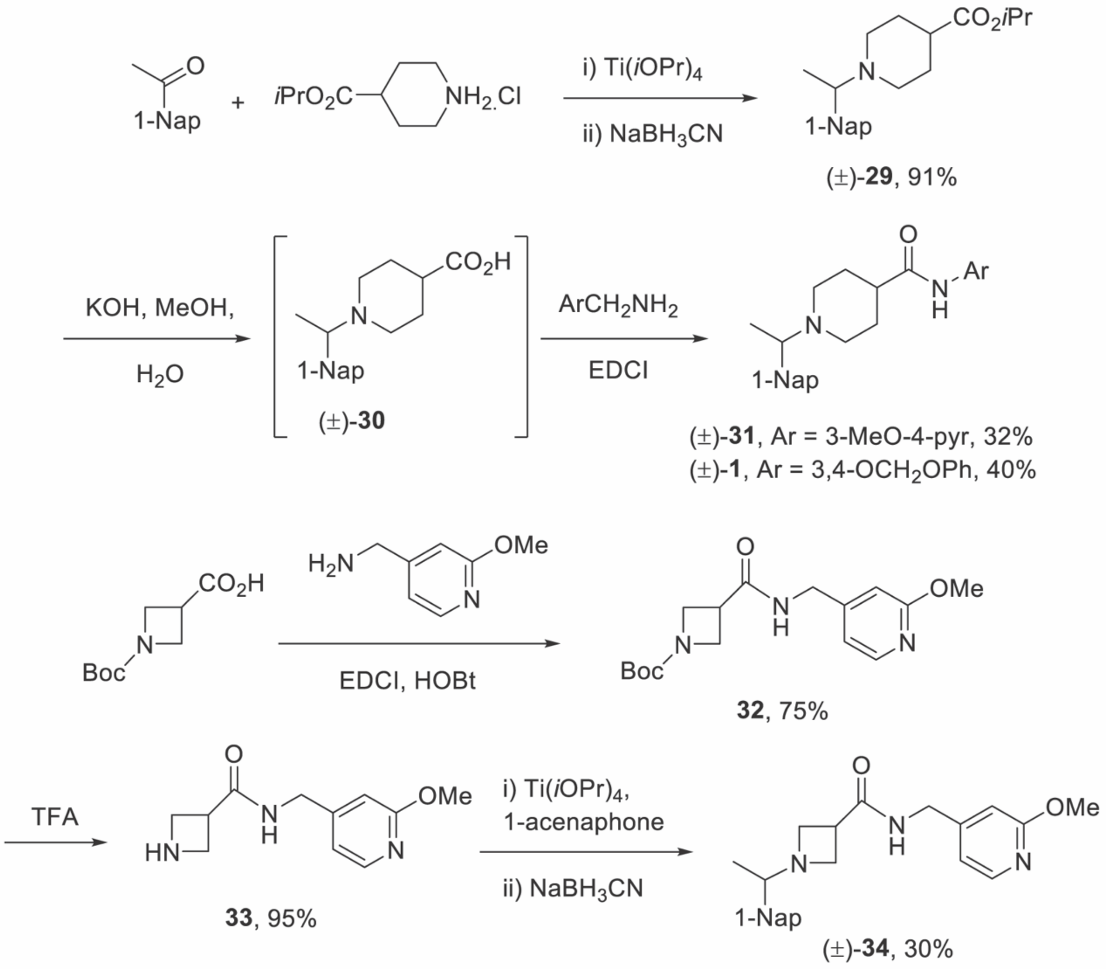
Scheme 4.3. Preparation of **24**, **25**, **26**, **27**, and **28**.



Turning to the series II compounds, reductive amination of 1-acetonaphthone isopropyl piperidine-4-carboxylate[50] by condensation in the presence of titanium tetraisopropoxide followed by sodium cyanoborohydride reduction gave the racemic amine

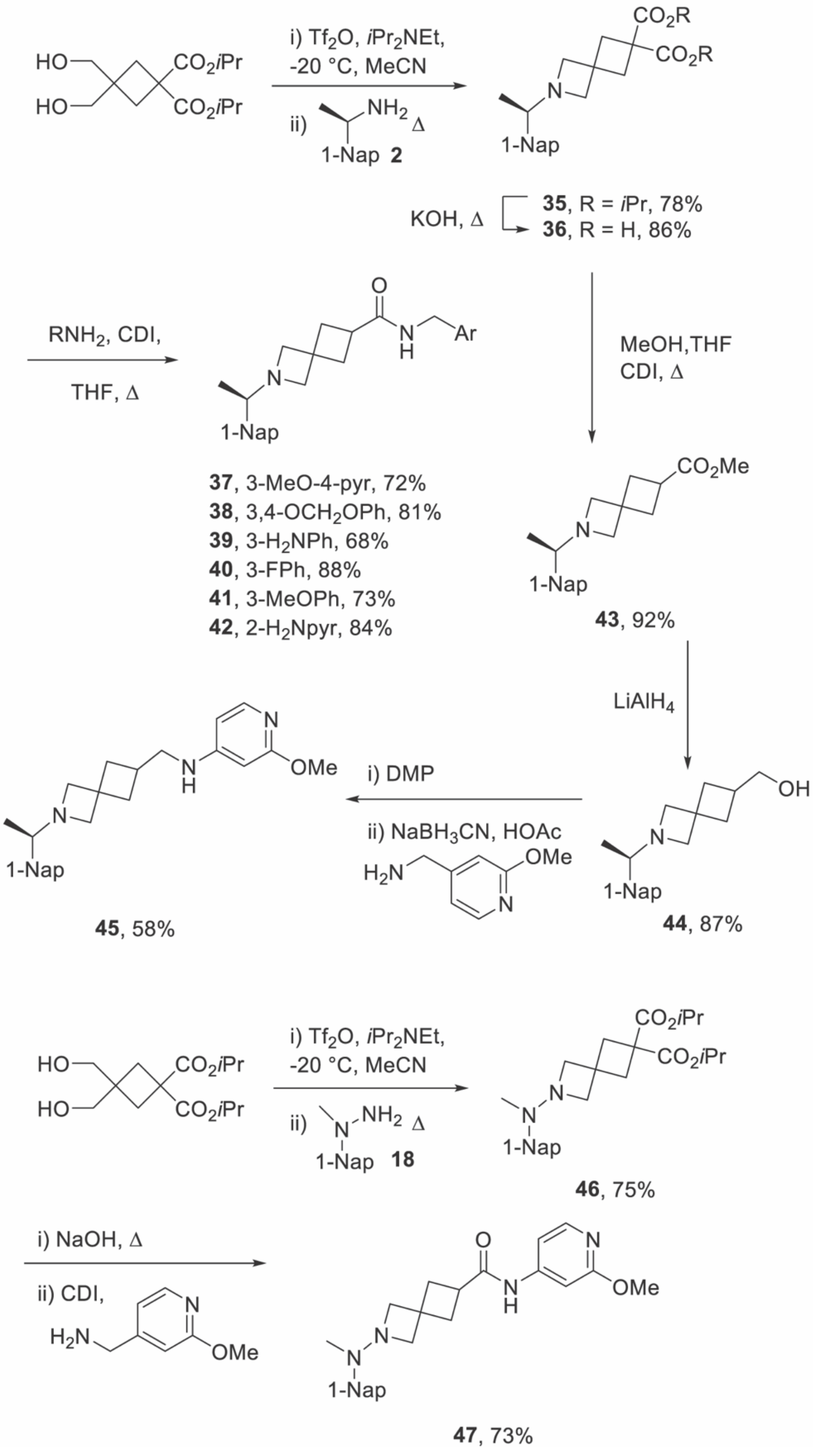
29.[51] Saponification then afforded the corresponding acid **30**, which was condensed with (2-methoxy-4-pyridyl)methylamine and 3,4-methylenedioxybenzylamine to give **31** and **2**, respectively (Scheme 4.4). An analog of **31** in which the 4-piperidine carboxamide moiety was replaced by a 3-azetidinecarboxamide was prepared by related approach in which *N*-(tert-butyloxycabonyl)azetidine-3-carboxylic acid[52] was condensed with (2-methoxy-4-pyridyl)methylamine to give the amide **32**, followed by removal of the carbamate and ultimate reductive amination with 1-acetonaphthone giving **34** (Scheme 4.4).

Scheme 4.4. Preparation of **1**, **29**, **31**, and **34**.



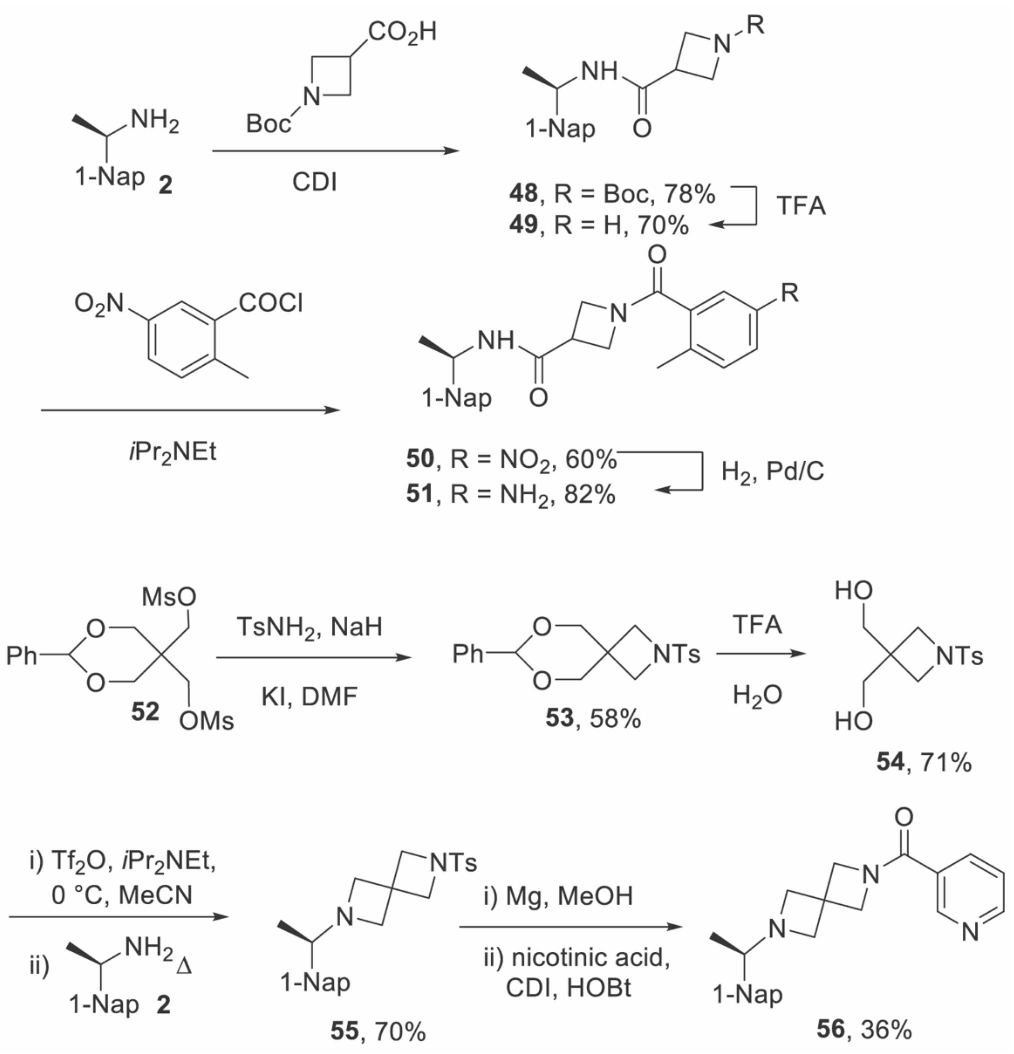
We next turned to the preparation of series II analogs in which the central piperidine ring was replaced by a spiroazetidine moiety. To this end diisopropyl 3,3-di(hydroxymethyl)cyclobutane-1,1-dicarboxylate[53] was activated with triflic anhydride in the presence of Hunig's base before addition of 1-(1*R*-naphthyl)ethylamine **2** and heating to 70°C to afford the spiroazetidine **35**. Saponification then gave the dicarboxylic acid **36**, which was coupled to an assortment of benzylic amines with carbonyl diimidazole in THF at reflux[54] to afford **37**, **38**, **39**, **40**, **41**, and **42**. Heating of the dicarboxylic acid **36** with carbonyl diimidazole in a mixture of methanol and THF at reflux afforded the ester **43**, which was reduced with lithium aluminum hydride to afford alcohol **44**. Oxidation of this alcohol with the Dess-Martin periodinane then gave an aldehyde, which was coupled to 2-methyoxy-4-pyridylmethylamine under standard reductive amination conditions affording the diamine **45**. Finally, we returned to the *N*-methyl hydrazide concept and prepared an analog **47**, of **37** in which the chiral CHMe unit was replaced by the NMe moiety, using a minor variation on the general route to the spiroazetidines (Scheme 4.5).[55, 56]

Scheme 4.5. Preparation of Spiroazetidine Derivatives **37**-**42**, **45**, and **47**.



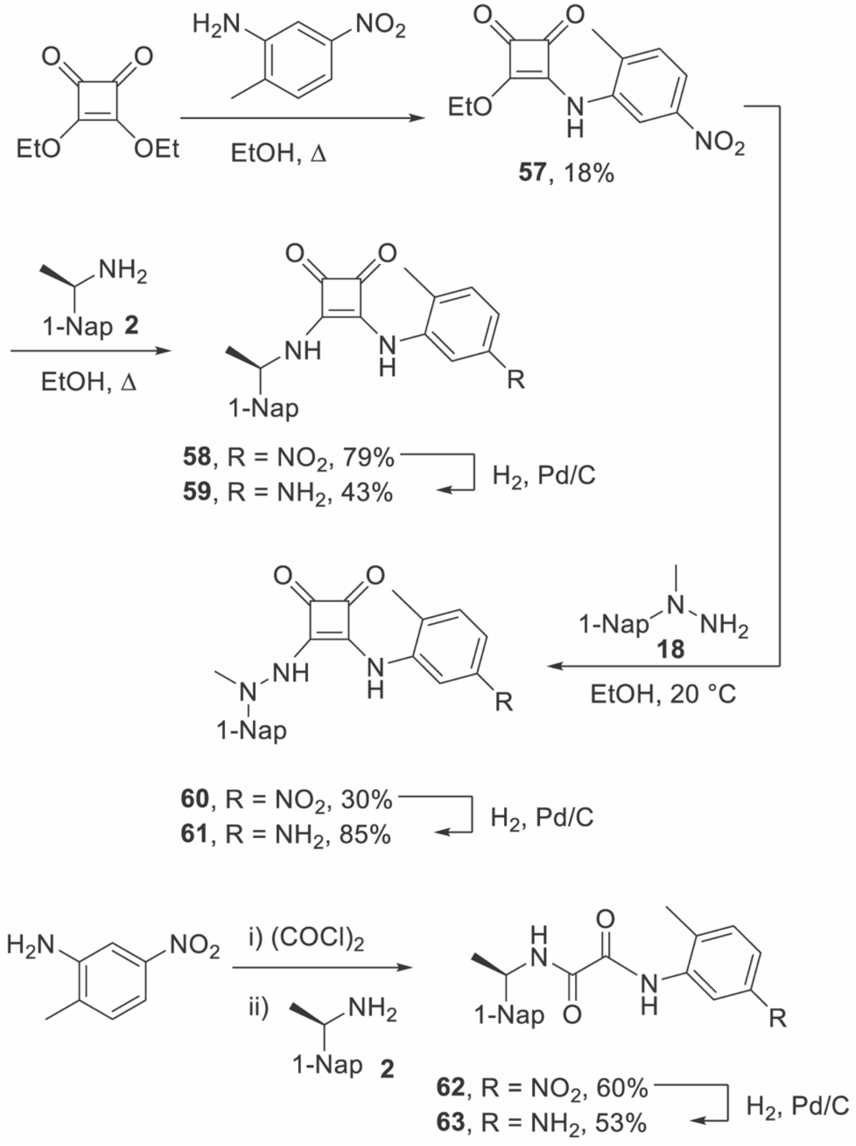
Returning to the theme of inverted amides and hybrid structures encompassing aspects of both the series I and II compounds, we first prepared **51** based on the azetidine core by standard means. In a similar vein, starting with the pentaerythritol derivative 2-phenyl-1,3-dioxane-5,5-diylbis(methylene) dimethanesulfonate **52**[57] we prepared the spirodiazetidine **55**, and coupled it to nicotinic acid affording **56** (Scheme 4.6).

Scheme 4.6. Preparation of Azetidine **51** and Spirodiazetidine **56**.



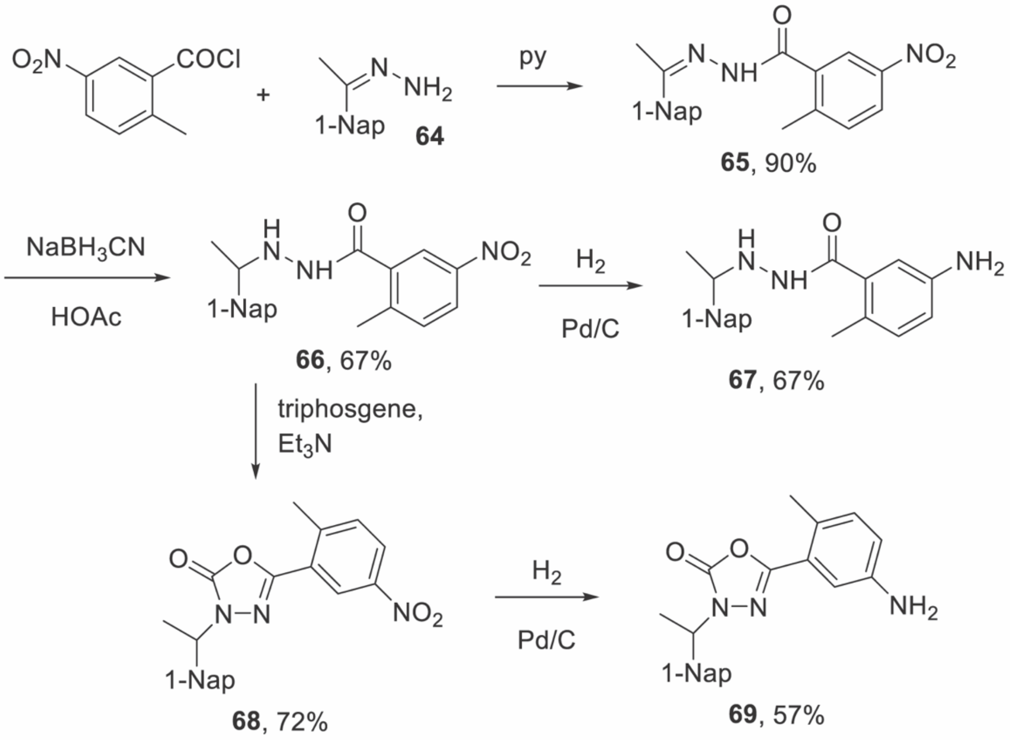
To further probe the structural requirements in the central portion of the binding pocket we turned to a brief exploration of squaramides[59, 60] and oxalamides[61]. To this end diethyl squarate was heated with 2-methyl-5-nitroaniline in ethanol with microwave irradiation to give **57** as a yellow solid, albeit only in 18% yield. Subsequent heating of **57** with 1*R*-(1-naphthyl)methylamine, again in ethanol with microwave heating afforded **58** in 79% yield, which was subjected to hydrogenolysis over palladium on charcoal to afford the desired **59** in 43% yield (Scheme 4.8). Stirring of the half-squaramide **57** with 1-(1-naphthyl)methylhydrazine **18** in ethanol at 20°C afforded the hydrazide **60** in 30%, and was followed by hydrogenolysis, which gave **61** in 85% yield (Scheme 4.7). An unsymmetrical oxalamide **62** was prepared in 60% yield by briefly stirring 2-methyl-5-nitroaniline with an excess of oxalyl chloride and potassium carbonate in dichloromethane at room temperature, followed by concentration to dryness and subsequent exposure to 1*R*-(1-naphthyl)ethylamine. Hydrogenolysis then gave the target compound **63** in 53% isolated yield (Scheme 4.7).

Scheme 4.7. Preparation of Squaramides **59** and **61**, and Oxalamide **63**.



To complete our investigation of the structural requirements of the central portion of the binding pocket, taking note of the broad utility of oxadiazoles as molecular scaffolds,[62] we prepared the oxazdiazole **69**. The synthesis began with acylation of 1-acenaphthone hydrazone **64** with 2-methyl-5-nitrobenzoyl chloride to give the acyl hydrazone **65**, which was reduced to the corresponding acyl hydrazide **66** with sodium cyanoborohydride. Hydrogenolysis of this substance over palladium on carbon gave hydrazide **67**, itself suitable for screening, whereas cyclodehydration to the oxadiazolone **68** was achieved with triphosgene, leaving only hydrogenolysis of the nitro group to complete the synthesis of **69** (Scheme 4.8).

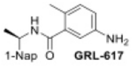
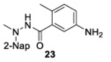
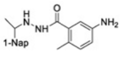
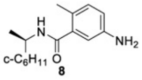
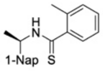
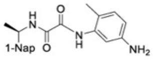
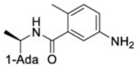
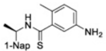
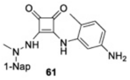
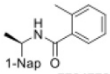
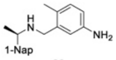
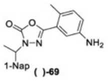
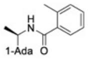
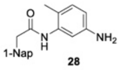
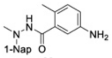
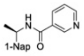
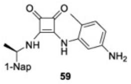
Scheme 4.8. Synthesis of Hydrazide **67** and Oxadiazolone **69**.



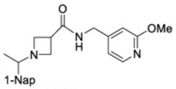
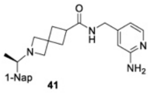
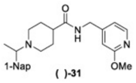
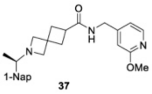
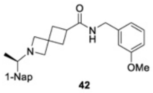
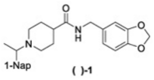
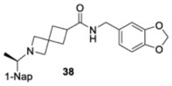
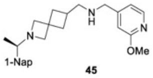
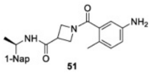
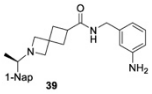
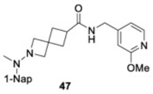
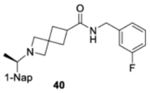
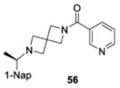
All compounds were assayed for their ability to inhibit SARS-CoV-2, SARS-CoV-1, and BtSCoV-Rf1.2004 PLpros, with an initial survey determining percentage inhibition at a fixed concentration of 50μM (Table 4.3).

Table 4.3. Naphthalene-based inhibitor efficacy against subgroup 2b PLpros. Percentage inhibition of SARS-CoV-2, SARS-CoV-1, and BtSCoV-Rf1.2004 PLpros ability to cleave peptide-AMC when incubated with series I (A) and II (B) compounds at 50µM. *Due to solubility limitations the inhibition of subgroup 2b PLpros by 50µM **8** could not be accurately assessed.

A

Compound	%age Inhibition	Compound	%age Inhibition	Compound	%age Inhibition
 GRL-617	97.2 ± 0.1	 23	22.5 ± 7.4	 ( )-67	13.8 ± 1.4
	98.3 ± 0.0		28.4 ± 3.4		NC
	97.7 ± 0.2		17.6 ± 3.7		NC
 8	*	 24	8.9 ± 2.5	 63	NC
			29.1 ± 5.3		20.0 ± 7.7
			19.3 ± 1.9		NC
 ( )-9	9.3 ± 5.4	 25	38.0 ± 3.8	 61	14.3 ± 3.8
	12.4 ± 6.2		45.6 ± 2.9		29.0 ± 1.9
	18.3 ± 2.6		18.3 ± 3.4		16.9 ± 5.1
 7724772	73.0 ± 0.8	 26	53.0 ± 2.7	 ( )-69	11.4 ± 0.8
	73.7 ± 1.0		56.6 ± 3.2		NC
	66.4 ± 1.1		66.4 ± 1.1		NC
 ( )-7	12.0 ± 1.4	 28	10.0 ± 1.1	 22	44.6 ± 2.5
	22.9 ± 3.4		29.2 ± 1.6		51.6 ± 1.3
	21.8 ± 6.6		NC		17.0 ± 5.2
 11	39.4 ± 3.0	 59	18.0 ± 3.1		
	44.4 ± 3.9		18.1 ± 5.4		
	12.9 ± 7.0		19.3 ± 1.4		

B

Compound	%age Inhibition	Compound	%age Inhibition	Compound	%age Inhibition
 ( )-34	77.6 ± 0.5	 41	63.8 ± 0.6	 ( )-31	98.4 ± 0.1
	83.2 ± 0.8		68.5 ± 2.7		94.6 ± 0.2
	76.3 ± 0.3		59.6 ± 2.8		89.1 ± 0.4
 37	82.1 ± 1.2	 42	67.6 ± 2.2	 ( )-1	97.2 ± 0.1
	92.5 ± 0.2		72.0 ± 1.2		98.0 ± 0.1
	87.2 ± 0.2		64.7 ± 0.5		97.1 ± 0.2
 38	78.3 ± 0.3	 45	21.9 ± 2.2	 51	NC
	80.0 ± 0.3		19.7 ± 3.5		19.5 ± 3.9
	69.1± 0.2		11.1 ± 1.3		17.0 ± 5.2
 39	49.3 ± 5.0	 47	9.6 ± 1.7	Structure	SARS-CoV-2 PLpro
	66.1 ± 0.8		NC		SARS-CoV-1 PLpro
	24.8 ± 2.5		8.0 ± 5.1		BtSCoV-Rf1.2004 PLpro
 40	66.0 ± 4.9	 56	NC		
	77.8 ± 0.4		NC		
	41.6 ± 4.1		5.5 ± 1.6		

Structure of BtSCoV-Rf1.2004 PLpro – 37 Complex

There is little data on how series II compounds interact with subgroup 2b PLpros, and furthermore no structure of compounds with spiroazetidine backbones. SAR data indicate that the use of either piperidine or spiroazetidine backbones will affect the potency of series II inhibitors. Perhaps more importantly, it shows that the optimal decoration of the eastern arene ring differs depending on which backbone is used. To further explore these differences, we crystallized BtSCoV-Rf1.2004 PLpro in complex with **37**. Utilizing X-ray diffraction its structure was determined to a resolution of 2.89Å in space group $P6_422$ (Table S4.1). The BtSCoV-Rf1.2004 PLpro-**GRL0617** complex structure was used as a search model for phasing the catalytic core. Upon finding a single monomer in the asymmetric unit, the Ubl domain was added. As with the **GRL0617** co-crystal structure, the Ubl domain of this structure was in a tucked conformation (Figure S4.1).

Within the active site, F_o-F_c density fitting **37** is readily observed within the P3/P4 binding pocket (Figure 4.8A). In this pocket **37** forms five hydrogen bonds with the protein. The amide scaffold forms hydrogen bonds with the main chains of L163, Y269, and Q270, while spiroazetidine nitrogen forms bonds with the side chains of D165 and Y274. Additionally, strong hydrophobic interactions are formed between the western naphthyl group and P248 and P249 as is typical of naphthalene-based PLpro inhibitors. However, when compared with a structure of SARS-CoV-1 PLpro in complex with **1** (PDB 3MJ5) we see that the naphthyl group of **37** is situated approximately 0.9Å further away from the two prolines (Figure 4.8B). The difference in naphthyl group location could be due to a shift of the BL2 loop caused by the bulky 1,3-dioxolane ring. The eastern arene groups of both **37** and **1** cause a shift in the BL2 loop that is similar to the conformation it adopts to accommodate Ubl substrates. However, this shift is more pronounced in the **1** co-crystal structure, even causing Q270 to flip away from the arene rings to avoid a steric clash.

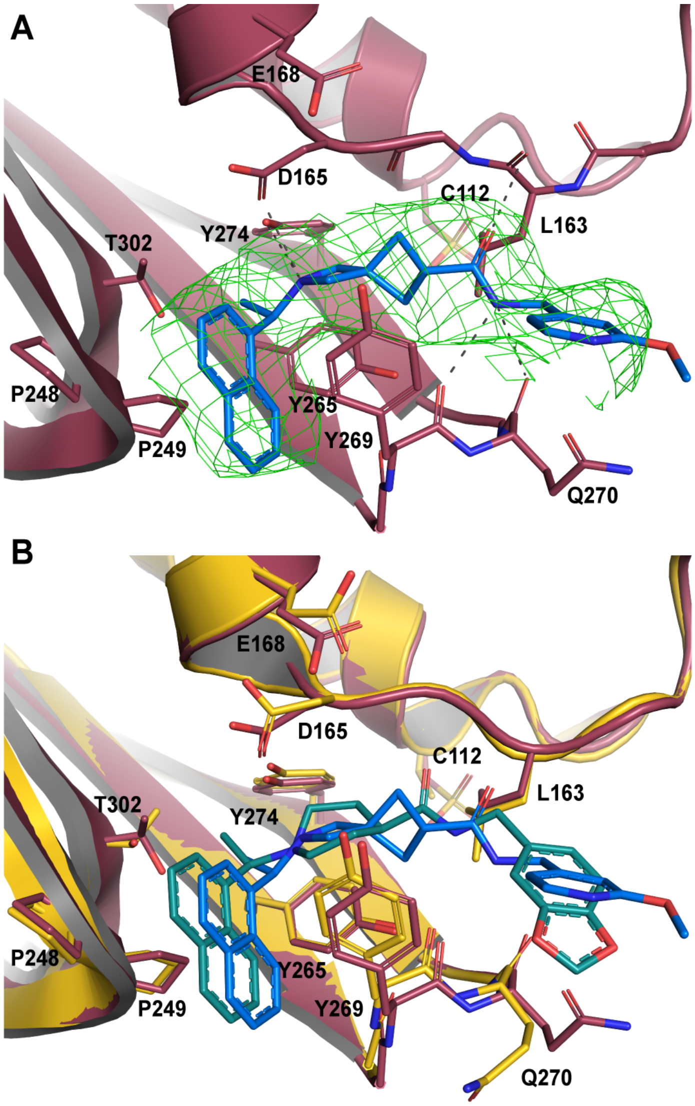


Figure 4.8. BtSCoV-Rf1 PLpro in complex with **37** A) An F_o-F_c electron density map is shown contoured at 1σ (green mesh). With **37** shown in blue and BtSCoV-Rf1 PLpro shown in raspberry. B) Overlay of **1** (teal) bound to SARS-CoV-2 PLpro (yellow).

In vitro cytotoxicity, serum binding, and stability of protease inhibitors

We determined the toxicity of these inhibitors by assessing their ability to decrease the reductive capability of numerous cells using the MTT assay (Table 4.4) and determining CC_{50} values. We compared these values to those for **GRL0617**, **1**, and **31**, all of which have been previously shown to have CC_{50}s greater than 100 μM in Vero E6 cells, with the latter two tested in HEK293 cells as well[27, 32]. To confirm this trend in immunologically relevant human cell lines all three compounds, along with **37** and **38** were assessed for CC_{50} in multiple cells, including human renal proximal tubule cells (RPTEC), Beas-2B, A549, and Sh-SH5Y cells. None of the other compounds were able to decrease MTT staining below 50% in A549 of Sh-SY5Y cells at concentrations as high as 100 μM. Focusing on RPTEC and BEAS-2b cells, **37** displayed the highest CC_{50} values in comparison to **1**, **38**, and **41**. The CC_{50} reported in these cells is folds higher than the IC_{50} reported for inhibition of PLpro degradation, as determined above. In general, amongst the newly-prepared compounds, alteration of the eastern moiety appeared to alter the CC_{50}, as compared to the west moiety or the central rings. The addition of the central ring decreased the CC_{50} as compared to **GRL0617**, as shown by comparison to other of the new series of compounds. The effect of these compounds on MTT staining was validated using cell morphology, which demonstrated morphological characteristics of cell death in combination with reduced cell number (Supplemental Information).

Table 4.4: CC_{50} values of PLpro Inhibitors in Human cell lines after 48 hr.

	RPTEC	**BEAS-2B**	**A549**	**SH-SY5Y**
GRL0617	240.7 μM	244.3 μM	427.6 μM	278 μM
37	56.07 μM	80.54 μM	>100 μM	>100 μM
38	21.84 μM	17.48 μM		
31	35.58 μM	62.23 μM		
1	41.5 μM	59.08 μM		

Discussion

Effects of Variation in Betacoronavirus PLpros on DUB activity

The PLpros of subgroup 2b viruses are highly conserved in primary, secondary, and tertiary structures, particularly at important function sites. The residues lining the active site, BL2 loop, P3 site, and P4 site are identical among all seven subgroup 2b PLpros but vary considerably in other betacoronaviruses. This results in both series I and II naphthalene-based compounds having similar inhibitory properties in all subgroup 2b PLpros but having negligible inhibition towards other betacoronaviruses[24, 63]. Furthermore, the UIM, which is known to accommodate both Ub and ISG15 binding, is fully conserved between BtSCoV-Rf1.2004 and SARS-CoV-1. SARS-CoV-2 PLpro has six differences at the UIM, T171(S), H172(Y), K196(Q), L217(F), V226(T), and Q233(K). Of the six points of difference in SARS-CoV-2, five are identical across the other subgroup 2b PLpros. The high conservation of Q233 across subgroup 2b is noteworthy because experiments have shown inserting a Q233(E) mutation in SARS-CoV-1 PLpro notably diminished DUB activity but increased deISGylase activity. Additionally, SARS-CoV-2 has a lysine residue at site 233 and has reduced DUB activity compared to SARS-CoV-1 and BtSCoV-Rf1.2004.

The lone difference within the UIM of BtSCoV-Rf1.2004 is L188, which is a valine in SARS-CoV-1 (Figure 4.9). This change may result in the BtSCoV-Rf1.2004 finger domain having stronger interactions with the hydrophobic pocket of Ub, particularly F45 when compared to SARS. Additionally, the increased proximity of L188 to I223 and V226 on the Zn finger loop seems to cause the loop to adopt a slightly more closed conformation around the Ub, increasing the proximity of the hydrophobic PLpro surface. A stronger interaction at the proximal Ub binding site would explain the lower K_m BtSCoV-Rf1.2004 PLpro has toward mono-Ub. When BtSCoV-Rf1.2004 PLpro DUB activity is examined through cleavage of various poly-Ub substrates it is apparent that its activity towards long-chain poly-Ub is more similar to SARS-CoV-1. Like other subgroup 2b viruses it follows a distinct di-distributive cleavage pattern, in that it struggles to process mono-Ub-AMC and any linkage type of di-Ub, but readily cleaves K48 linked poly-Ub chains of three or more (Figure 4.3).

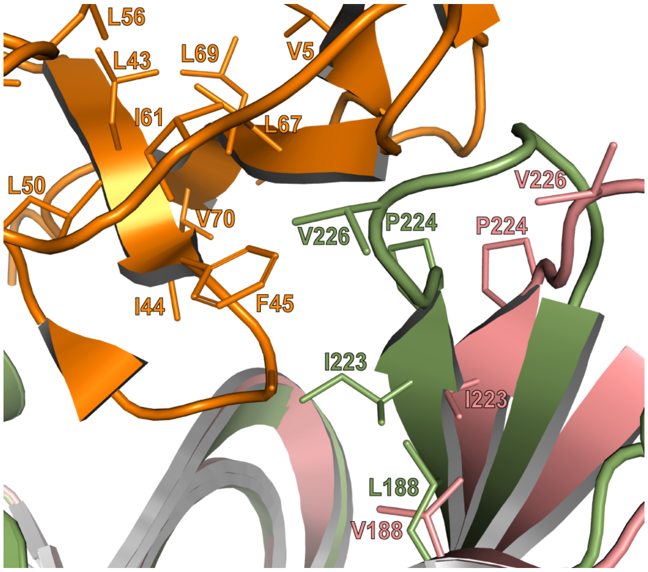


Figure 4.9. Hydrophobic interactions between Ub and PLpro finger domains. Close up view of the hydrophobic residues within the finger domain portion of the proximal Ub binding pocket of BtSCoV-Rf1.2004 PLpro (green) and SARS-CoV-1 PLpro (pink). BtSCoV-Rf1.2004 PLpro is overlaid with a structure (PDB 5E6J) of SARS-CoV-1 bound to K48 linked di-Ub (orange).

Conversely, MHV PLP and MERS-CoV PLpro share secondary structure topology with subgroup 2b PLpros but vary widely in sequence homology. These residue differences, particularly in areas such as the UIM, are responsible for differences in kinetics and substrate preferences. Both have mono-distributive DUB activity, in that they can cleave any length poly-Ub chain and do not have as strong a preference for K48 linked poly-Ub as subgroup 2b PLpros. The trend in conservation extends to the residues stabilizing the interaction between the Ubl and thumb domain as well and may provide insight into the nature of mono vs di-distributive DUB activity. These residues are highly conserved among subgroup 2b PLpros but are not found in other betacoronaviruses, which points to the possibility that the tucked conformation is a distinct feature of subgroup 2b PLpros. It is unclear if this affects PLpro activity, but the correlation between di-distributive DUB activity and the ability to adopt tucked Ubl conformations warrants further study.

A significant trait that BtSCoV-Rf1.2004 PLpro differs from SARS-CoV-1 on is a reduced K_M to mono-Ub. Like MERS-CoV and SARS-CoV-2 PLpros, BtSCoV-Rf1.2004 PLpro can be readily saturated with mono-Ub at concentrations equivalent to those found in mammalian cells, which range from 10 to 23 μM. The ability to be saturated at cellular Ub concentrations and relative inability to process mono-Ub make BtSCoV-Rf1.2004 susceptible to product inhibition similar to SARS-CoV-2 and MERS-CoV. Meanwhile, at physiological concentrations, SARS-CoV-1 PLpro would not even be at its K_M.

Trends in Subgroup 2b PLpro DeISGylating Activities.

Despite Ubls being inherently similar substrates, viral proteases, including coronavirus PLpros, display a fairly high degree of substrate specificity. Differences in poly-Ub chain linkage can affect how they fit within the UIM, and ISG15 mutations between species can affect interactions at critical specificity sites. Typically, viral proteases will be adapted to rapidly process substrates involved in pathways detrimental to their survival and may retain off-target activity towards highly similar substrates. While all betacoronaviruses prefer the ISG15 of their natural host over Ub, host species, and therefore substrate preferences vary widely between viruses. Similar to what has been seen in SARS-CoV-1 and SARS-CoV-2, the deISGylase activity of BtSCoV-Rf1.2004 PLpro exceeds its DUB activity, except in the case of long-chain K48-Ub. The consistently robust deISGylase activity displayed by subgroup 2b PLpros relative to their modest DUB activity suggests that ISG15 regulated immune pathways may be more detrimental to viral replication than pathways regulated by Ub. Despite this, the specific utility deISGylase activity serves coronaviruses is unknown. Because deISGylase activity appears critical to coronavirus replication, species-specific deISGylase activity can be a factor in determining what species a virus can productively infect. Interspecies variation in ISG15 has been shown to limit the zoonotic range of influenza B[64, 65]. The similarity in species specificity between subgroup 2b viruses indicates that these viruses probably infect many of the same host species. Humans, palm civets, pangolins, minks, and several bats have all been identified as host species for subgroup 2b viruses, and some have been shown to host multiple 2b viruses[2, 4, 5, 66]. The largely conserved species preferences of subgroup 2b viruses may also enable them to productively infect new species that already serve as reservoirs for other 2b viruses with relatively few mutations.

BtSCoV-Rf1.2004 PLpro may be able to serve as a useful tool in determining residues that affect interspecies differences between SARS-CoV-1 and SARS-CoV-2. The deISGylase activities of these PLpros differ against ISG15s from sheep, camel, northern tree shrew, and jackknife fish. The PLpro of BtSCoV-Rf1.2004 matches the activity of SARS-CoV-2 against sheep, camel, and jackknife fish, but matches that of SARS-CoV-1 against northern tree shrew. While mammalian respiratory viruses would not naturally infect fish, SARS-CoV-1, MERS-CoV, and MHV PLpros all demonstrate some off-target activity towards jackknife fish ISG15. SARS-CoV-2 was the first betacoronavirus to show no appreciable activity towards fish ISG15, but the lack of activity demonstrated by BtSCoV-Rf1.2004 PLpro, which is more closely related to SARS-CoV-1, could help discern which differences are responsible.

The ability of BtSCoV-Rf1.2004 PLpro to quickly processes horseshoe bat and vesper bat ISG15s, while having little activity towards Egyptian fruit bats is not unusual. This was observed in SARS-CoV-2 as well and is likely due to the diversity within the order *Chiroptera*. Bats make up over 20% of all mammal species and ISG15 can share as little as 60% sequence identity between bats from different families, which is no greater than the similarity expected between any two mammalian ISG15s.

Naphthalene-based inhibitors as lead compounds for coronavirus subgroup 2b therapeutic development

Initial evaluation of each novel compound at 50 μM revealed that most compounds have equivalent levels of inhibition against all three subgroup 2b PLpros regardless of inhibitor efficacy. This trend was further supported by IC_{50} testing of top compounds, which determined that IC_{50} values for each inhibitor for all three PLpros were usually within one to two-fold, and always within one order of magnitude. In addition to previously examined compounds, several of the newly designed compounds demonstrated low μM

IC_{50} values and show promise as potential lead compounds. Of the compounds tested in this study **GRL0617**, **1**, **31**, and **37** were the most potent inhibitors. In all four cases IC_{50} values are below 10 μM and CC_{50} values are more than five times the IC_{50} in RPTEC and BEAS-2B cells and greater than 100 μM in A549 and SH-SY5Y cells. Within series II the eastern arene ring was the most critical factor affecting CC_{50} compared to the western naphthyl group or backbone ring system. Alteration of the 1,3-dioxolane ring of **1** and **38** to a 2-methoxy-4-pyridine ring yielded larger therapeutic windows, but only when paired with a spiroazetidine backbone to orient the group properly within the pocket. These two modifications result in **37** having the highest CC_{50} when compared to other series II compounds tested. These compounds can be useful probes in studies of PLpro activity, and with further optimization have potential as lead candidates in the development of therapeutics for the treatment of subgroup 2b viruses.

Conclusions

The ability of all three 2b PLpros to productively cleave ISG15 from humans as well as several important livestock species further highlights the need to monitor these viruses closely and to develop effective therapeutics to mitigate the damage in the event of future spillover. Most naphthalene-based compounds inhibit all three subgroup 2b PLpros at similar levels, indicating that P3/P4 binding drugs will be useful for the treatment of future subgroup 2b coronavirus outbreaks as well as the current COVID-19 pandemic. Structural data indicate that these compounds could be furtherer optimized to bind these pockets, enhancing potency and specificity. Current naphthalene-based PLpro inhibitors are effective at treating both SARS-CoV-1 and SARS-CoV-2 *in vitro* and have been found not to have cytotoxic effects. While significant *in vivo* testing is still required before these compounds could be used therapeutically, they are promising lead compounds for treating immunocompromised patients that can't receive vaccines or 2b viruses that induce short-lived immunity.

Experimental Section

Chemicals and Reagents.

5-Amino-2-methyl-N-[(R)-1-(1-naphthyl)ethyl]benzamide (**GRL0617**) was purchased from Raystar, CN; 5-(acetylamino)-2-methyl-N-[(1R)-1-(1-napthanlenyl)ethyl]-benzamide (compound 6) was purchased from MedChem Express; Z-RLRGG-7-amino-4-methyl-courmarin (peptide-AMC) was purchased from Bachem; Ubiquitin−7-amino-4-methylcourmarin (Ub-AMC) was purchased from Boston Biochem; human ISG15−7-amino-4-methylcourmarin (ISG15-AMC) was purchased from Boston Biochem; Lys6, Lys11, Lys29, Lys33, Lys48, Lys63, and linear linked di-Ub were obtained from Boston Biochem; DL-dithiothroitol (DTT) was purchased from GoldBio; Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from GoldBio; 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was purchased from Fisher BioReagents; Imidazole was purchased from Acros Organics; tris(hydroxymethyl)aminomethane (Tris) was purchased from Fisher Scientific; Sodium chloride (NaCl) was purchased from Fisher Chemical; Bovine serum albumin (BSA) was purchased from Sigma Life Science; Dehydrated Luria-Bertani (LB) broth was purchased from Fisher Scientific; Ampicillin was purchased from GoldBio.

Construction, Expression, and Purification of Viral Deubiquitinases.

The ubiquitin-like domain (Ubl) and the catalytic core of BtSCoV-Rf1.2004 PLpro (pp1ab 1536-1850; 2-316) were cloned into pET-15b by Genscript and transformed into

T7 express E. coli. Cells were cultured in 4.5L of LB broth containing 100µg/mL ampicillin at 37°C until the OD600 reached 0.6. Once reached, the expression was induced by the addition of 0.5mM isopropyl β-D-thiogalactopyranoside (IPTG), and the culture was incubated at 18°C overnight. The culture was centrifuged at 12,000g for 10min, and the pellet was collected and stored in a −80°C freezer. The cell pellet was dissolved into lysis buffer (500mM NaCl and 50mM Tris-HCl [pH = 7.5]) and then sonicated in Fisher Scientific series 150 on ice at 50% power with 5s pulses for 6min. The lysate was centrifuged at 30,000g for 30min to remove all insoluble products. The supernatant was then filtered and placed onto Ni-nitrilotriacetic agarose resin (Qiagen). The resin was washed using five column volumes of lysis buffer containing 10mM imidazole. The protein was eluted using 5 column volumes of lysis buffer containing 250mM imidazole. Thrombin was added to the elution to remove the 6X His-tag, and the combined solution was dialyzed in size exclusion buffer (100mM NaCl, 5mM HEPES, and 2mM dithiothreitol (DTT) [pH = 7.4]) and run over a Size Exclusion Superdex 200 column (GE Healthcare, Pittsburgh PA). Purity was confirmed by gel electrophoresis.

Co-crystallization of BtSCoV-Rf1.2004 PLpro with GRL0617

Size-exclusion chromatography fractions containing BtSCoV-Rf1.2004 PLpro were pooled based on the chromatogram, confirmed by SDS–PAGE and concentrated to 10.5 mg/mL and 10.4 mg/mL for co-crystallization with **GRL0617** and **37** respectively. **GRL0617** in 100% DMSO was added to the protein sample at a 5:1 molar ratio with a final DMSO concentration of 3%. The sample was screened in 96-well hanging-drop plates against a suite of 1728 commercially available conditions (Qiagen and Hampton) using a Mosquito robot (TTP Labtech). Plate-like crystals in different conditions formed in a time ranging from one to three weeks. Hits from the screen were scaled up to hanging-drop 24-well plates containing a 500 μL reservoir of the crystallization solution and were optimized using varying salt, precipitant, pH, additive and protein concentration gradients. The final crystallization condition for the optimized BtSCoV-Rf1.2004 PLpro-**GRL0617** co-crystal was 0.2M Ammonium Acetate, 20% PEG 1000. **37** in 100% DMSO was added to the protein sample at a 5:1 molar ratio and a final DMSO concentration of 0.5%. A crystal screen was set up for the sample against a suite of 768 commercially available conditions (Qiagen) in 96-well hanging-drop plates using a Mosquito robot (TTP Labtech). The final crystallization condition for the BtSCoV-Rf1.2004 PLpro-**37** co-crystals was 0.1 M Tris-HCl pH 8.5, 0.6M sodium fluoride, resulting in bipyramidal crystals in one to three weeks.

Data Collection and Processing

All crystals were flash-cooled in liquid nitrogen. BtSCoV-Rf1.2004 PLpro-**GRL0617** complex structural data were collected under a dry N2 stream on the 17-ID (AMX) beamline at Brookhaven National Laboratory using an Eiger-9M detector. BtSCoV-Rf1.2004 PLpro-**37** complex structural data were collected on the 22-ID beamline using an Eiger-16M detector at Argonne National Laboratory under a dry N_2 stream. The data were indexed, integrated and scaled in HKL-2000[67]. Data-collection statistics are included in Table S4.1.

Structure Solution and Refinement

Phases for the BtSCoV-Rf1.2004 PLpro–**GRL0617** and **37** co-crystal structures were solved by molecular replacement in Phaser[68-70]. A homology model of BtSCoV-Rf1.2004 PLpro based on a SARS-CoV-1 PLpro–**GRL0617** co-crystal structure (PDB 3E9S) was used as a reference for the **GRL0617** complex, while the BtSCoV-Rf1.2004

PLpro–**GRL0617** structure was used as the model for the **37** complex. In both cases the finger, thumb, and palm domains of the PLpro were used as a search model, placing 2 copies in the asymmetric unit of the **GRL0617** complex and 1 in the asymmetric unit of the **37** complex. Afterwards the Ubl domains were built in manually. The phased models were modified through alternating rounds of manual building and refinement in Coot and Phenix[71]. The final models were validated in MolProbity[72], and the associated refinement statistics are included in Table S4.1.

BtSCoV-Rf1.2004 PLpro Deubiquitinase and deISGylating Assays.

All assays were run using Corning Costar half-volume 96-well plates containing AMC buffer (100mM NaCl, 50mM HEPES [pH = 7.5], 0.01mg/mL bovine serum albumin (BSA), and 5mM DTT) to a final volume of 50μL and performed in triplicate. The CLAIROstar plate reader (BMG Lab Tech, Inc.) was used to measure the fluorescence of the AMC cleavage, and the data was analyzed using MARS (BMG Lab Tech, Inc.). The AMC fluorescence was observed from the cleavage of Ub-AMC and ISG15-AMC obtained from Boston Biochem, MA. ISG15-AMC concentrations of substrate ranged from 625nM to 20μM, and Ub-AMC ranged from 235nM to 30μM. Protease concentrations used for the Ub-AMC and ISG15-AMC assays were 5 and 1nM, respectively. To calculate KM and Vmax values, the initial rates were fitted to the Michalis-Menten equation, $\upsilon = Vmax/(1 + (KM/[S]))$, using the Enzyme Kinetics (v. 1.3) module of SigmaPlot (v. 10.0, SPSS Inc.). Vmax was translated into kcat using $kcat = Vmax/[E]$.

BtSCoV-Rf1.2004 PLpro Poly-Ub Cleavage Assays.

Lys6, Lys11, Lys29, Lys33, Lys48, Lys63, and linear linked di-Ub obtained from Boston Biochem were incubated at 10 μM with 20 nM BtSCoV-Rf1.2004 PLpro.

Reactions were performed in AMC buffer at a volume of 45 µL and a temperature of 37°C. 10 µL samples were taken at the indicated time points and heat-shocked at 98°C for 5 min. Lys48 and Lys63 linked tetra-Ub obtained from Boston Biochem were incubated at 13.67 µM with 23 nM BtSCoV-Rf1.2004 PLpro. Reactions were performed in AMC buffer at a volume of 70 µL and a temperature of 37°C. 8 µL samples were taken at the indicated time points and heat-shocked at 98°C for 5 min. SDS-PAGE analysis was performed using Mini-PROTEAN TGX and Coomassie blue.

Protease Activity Assay with proISG15 Substrates.

At 37°C, 20nM BtSCoV-Rf1.2004 PLpro was run against 10 µM of each ISG15. Reaction mixtures were 100 µL in AMC buffer. 10µL samples were taken at the indicated time points, and the reaction was quenched in 2×Laemmli sample buffer followed by boiling at 98°C for 5min. SDS-PAGE analysis was performed using Mini-PROTEAN TGX Stain-Free.

BtSCoV-Rf1.2004 PLpro Inhibition IC_{50} Value Determination.

IC_{50} assays were performed using similar methods to peptide-AMC, Ub-AMC, and ISG15-AMC cleavage experiments and those described previously. BtSCoV-Rf1.2004 PLpro was run at 1 µM against 50 µM peptide-AMC in 98% AMC buffer/2% DMSO. Reactions were performed in triplicate with inhibitor concentrations ranging from 195nM to 100 µM, depending on compound tested. IC_{50} calculations were performed using Prism8 from GraphPad.

Compound Synthesis

The synthesis of series I and series II compounds is described in the supplemental information section.

Cells and Culture Conditions

Non-targeted toxicity was assessed in numerous diverse cell lines, including Renal Proximal Tubular Epithelial Cells (RPTECs, ATCC CRL-4031), BEAS-2B bronchial epithelial cells (ATCC CRL-9609), A549 adenocarcinomic alveolar basal epithelial cells (ATCC, CCL-185), and SH-SY5Y neuroblastoma cells (ATCC, CRL-2266). RPTECs were grown in DMEM/F12 (ATCC, 30-2006) supplemented with hTERT Immortalized RPTEC Growth Kit (ATCC, ACS-4007); A549 and SH-SY5Y cells were grown in DMEM (ATCC, 30-2002) supplemented with 10% Fetal Bovine Serum (VWR, 97068-085) and 1% penicillin-streptomycin solution (ATCC, 30-2300). BEAS-2B cells were grown using the BEGM BulletKit (Lonza, CC-3170) and flasks were coated with 0.03 mg/mL bovine collagen (Fisher, CB-40231) and 0.01 mg/mL human fibronectin (Fisher, CB-40008A). All cells were maintained at 37°C in a 5% CO_2 incubator.

Cytotoxicity was assessed using MTT staining and cell morphology. Cells were seeded in 48-well tissue culture plates at densities between 25 – 50,000 (A549, BEAS-2B), 100,000 (RPTEC), and 50 – 150,000 (SH-SY5Y) cells per well depending on growth rate and experimental conditions. Cells were maintained at 37°C in a 5% CO_2 incubator for a minimum of 24 hr and were at least 80% confluent prior to dosing. Inhibitors were dissolved in DMSO and diluted in different culture media to their final concentrations. Cells were treated with DMSO alone (vehicle control, no greater than 0.1% (v/v)) or 1, 5, 10, 50, 75, or 100 μM of each inhibitor; except for **GRL0617**, where cells were also treated with 175, 250, 500, 750, or 1000 μM concentrations. For serum binding assays, inhibitors (and DMSO controls) were pre-incubated in DMEM containing 10% (v/v) Fetal Bovine

Serum and 1% penicillin-streptomycin for 5 minutes prior to cell exposure. Stability was assessed by incubation of inhibitors for 24 hr at 37°C in media. Alterations in cytotoxicity as compared to un-incubated controls were indicative of serum binding or inhibitor instability. Regardless of the experiment, MTT (Sigma, M2128-1G) was added after 48 hr treatment at a final concentration of 0.1 mg/mL, and plates were incubated for 2 hr at 37°C. After media aspiration, the remaining MTT formazan crystals were dissolved in DMSO and absorbance was determined for each well at 490 nm using a BMG CLARIOstar plate reader. Experiments were performed in triplicate per passage in at least 3 distinct passages of cells. The concentration of protease inhibitor that resulted in 50% growth inhibition (CC_{50}) as compared to DMSO control was estimated from a non-linear regression curve as calculated in GraphPad Prism 7. When GraphPad could not automatically determine accurate CC_{50} values, the concentration on the linear regression curve at which MTT staining was half of control was utilized. Changes in MTT staining were compared to changes in cell morphology at either 24 or 48 hr after exposure to inhibitors using a Nikon Eclipse Ti. Figures were assembled in Photoshop.

References

1. Wang, L. F.; Shi, Z.; Zhang, S.; Field, H.; Daszak, P.; Eaton, B. T., Review of bats and SARS. *Emerg Infect Dis* **2006,** *12* (12), 1834-40.

2. Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J. H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H.; Zhang, J.; McEachern, J.; Field, H.; Daszak, P.; Eaton, B. T.; Zhang, S.; Wang, L. F., Bats are natural reservoirs of SARS-like coronaviruses. *Science* **2005,** *310* (5748), 676-9.

3. Ji, W.; Wang, W.; Zhao, X.; Zai, J.; Li, X., Cross-species transmission of the newly identified coronavirus 2019-nCoV. *J Med Virol* **2020,** *92* (4), 433-440.

4. Menachery, V. D.; Yount, B. L., Jr.; Debbink, K.; Agnihothram, S.; Gralinski, L. E.; Plante, J. A.; Graham, R. L.; Scobey, T.; Ge, X. Y.; Donaldson, E. F.; Randell, S. H.; Lanzavecchia, A.; Marasco, W. A.; Shi, Z. L.; Baric, R. S., A SARS-like cluster

of circulating bat coronaviruses shows potential for human emergence. *Nat Med* **2015,** *21* (12), 1508-13.

5. Lam, T. T.; Jia, N.; Zhang, Y. W.; Shum, M. H.; Jiang, J. F.; Zhu, H. C.; Tong, Y. G.; Shi, Y. X.; Ni, X. B.; Liao, Y. S.; Li, W. J.; Jiang, B. G.; Wei, W.; Yuan, T. T.; Zheng, K.; Cui, X. M.; Li, J.; Pei, G. Q.; Qiang, X.; Cheung, W. Y.; Li, L. F.; Sun, F. F.; Qin, S.; Huang, J. C.; Leung, G. M.; Holmes, E. C.; Hu, Y. L.; Guan, Y.; Cao, W. C., Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **2020,** *583* (7815), 282-285.

6. Dudley, M. Z.; Halsey, N. A.; Omer, S. B.; Orenstein, W. A.; O'Leary, S. T.; Limaye, R. J.; Salmon, D. A., The state of vaccine safety science: systematic reviews of the evidence. *Lancet Infect Dis* **2020,** *20* (5), e80-e89.

7. Chiu, M.; Bao, C.; Sadarangani, M., Dilemmas With Rotavirus Vaccine: The Neonate and Immunocompromised. *Pediatr Infect Dis J* **2019,** *38* (6S Suppl 1), S43-S46.

8. Dilucca, M.; Forcelloni, S.; Georgakilas, A. G.; Giansanti, A.; Pavlopoulou, A., Codon Usage and Phenotypic Divergences of SARS-CoV-2 Genes. *Viruses* **2020,** *12* (5).

9. Thiel, V.; Ivanov, K. A.; Putics, A.; Hertzig, T.; Schelle, B.; Bayer, S.; Weissbrich, B.; Snijder, E. J.; Rabenau, H.; Doerr, H. W.; Gorbalenya, A. E.; Ziebuhr, J., Mechanisms and enzymes involved in SARS coronavirus genome expression. *J Gen Virol* **2003,** *84* (Pt 9), 2305-2315.

10. Daczkowski, C. M.; Dzimianski, J. V.; Clasman, J. R.; Goodwin, O.; Mesecar, A. D.; Pegan, S. D., Structural Insights into the Interaction of Coronavirus Papain-Like Proteases and Interferon-Stimulated Gene Product 15 from Different Species. *J Mol Biol* **2017,** *429* (11), 1661-1683.

11. Barretto, N.; Jukneliene, D.; Ratia, K.; Chen, Z.; Mesecar, A. D.; Baker, S. C., The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. *J Virol* **2005,** *79* (24), 15189-98.

12. Mielech, A. M.; Deng, X.; Chen, Y.; Kindler, E.; Wheeler, D. L.; Mesecar, A. D.; Thiel, V.; Perlman, S.; Baker, S. C., Murine coronavirus ubiquitin-like domain is important for papain-like protease stability and viral pathogenesis. *J Virol* **2015,** *89* (9), 4907-17.

13. Ratia, K.; Saikatendu, K. S.; Santarsiero, B. D.; Barretto, N.; Baker, S. C.; Stevens, R. C.; Mesecar, A. D., Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc Natl Acad Sci U S A* **2006,** *103* (15), 5717-22.

14. Zhao, C.; Denison, C.; Huibregtse, J. M.; Gygi, S.; Krug, R. M., Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc Natl Acad Sci U S A* **2005,** *102* (29), 10200-5.

15. Davis, M. E.; Gack, M. U., Ubiquitination in the antiviral immune response. *Virology* **2015,** *479-480*, 52-65.

16. Sadler, A. J.; Williams, B. R., Interferon-inducible antiviral effectors. *Nat Rev Immunol* **2008,** *8* (7), 559-68.

17. Frieman, M.; Ratia, K.; Johnston, R. E.; Mesecar, A. D.; Baric, R. S., Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling. *J Virol* **2009,** *83* (13), 6689-705.

18. Devaraj, S. G.; Wang, N.; Chen, Z.; Chen, Z.; Tseng, M.; Barretto, N.; Lin, R.; Peters, C. J.; Tseng, C. T.; Baker, S. C.; Li, K., Regulation of IRF-3-dependent innate immunity by the papain-like protease domain of the severe acute respiratory syndrome coronavirus. *J Biol Chem* **2007,** *282* (44), 32208-21.

19. Scholte, F. E. M.; Zivcec, M.; Dzimianski, J. V.; Deaton, M. K.; Spengler, J. R.; Welch, S. R.; Nichol, S. T.; Pegan, S. D.; Spiropoulou, C. F.; Bergeron, E., Crimean-Congo Hemorrhagic Fever Virus Suppresses Innate Immune Responses via a Ubiquitin and ISG15 Specific Protease. *Cell Rep* **2017,** *20* (10), 2396-2407.

20. Zhao, C.; Sridharan, H.; Chen, R.; Baker, D. P.; Wang, S.; Krug, R. M., Influenza B virus non-structural protein 1 counteracts ISG15 antiviral activity by sequestering ISGylated viral proteins. *Nat Commun* **2016,** *7*, 12754.

21. Dzimianski, J. V.; Scholte, F. E. M.; Bergeron, E.; Pegan, S. D., ISG15: It's Complicated. *J Mol Biol* **2019,** *431* (21), 4203-4216.

22. Freitas, B. T.; Scholte, F. E. M.; Bergeron, E.; Pegan, S. D., How ISG15 combats viral infection. *Virus Res* **2020,** *286*.

23. Shen, Z.; Ratia, K.; Cooper, L.; Kong, D.; Lee, H.; Kwon, Y.; Li, Y.; Alqarni, S.; Huang, F.; Dubrovskyi, O.; Rong, L.; Thatcher, G. R.; Xiong, R., Potent, Novel SARS-CoV-2 PLpro Inhibitors Block Viral Replication in Monkey and Human Cell Cultures. *bioRxiv* **2021**.

24. Shin, D.; Mukherjee, R.; Grewe, D.; Bojkova, D.; Baek, K.; Bhattacharya, A.; Schulz, L.; Widera, M.; Mehdipour, A. R.; Tascher, G.; Geurink, P. P.; Wilhelm, A.; van der Heden van Noort, G. J.; Ovaa, H.; Muller, S.; Knobeloch, K. P.; Rajalingam, K.; Schulman, B. A.; Cinatl, J.; Hummer, G.; Ciesek, S.; Dikic, I., Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **2020,** *587* (7835), 657-662.

25. Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M.; Olsen, S. K., Activity profiling and structures of inhibitor-bound SARS-CoV-2-PLpro protease provides a framework for anti-COVID-19 drug design. *bioRxiv* **2020**.

26. Chen, X.; Chou, C. Y.; Chang, G. G., Thiopurine analogue inhibitors of severe acute respiratory syndrome-coronavirus papain-like protease, a deubiquitinating and deISGylating enzyme. *Antivir Chem Chemother* **2009,** *19* (4), 151-6.

27. Baez-Santos, Y. M.; Barraza, S. J.; Wilson, M. W.; Agius, M. P.; Mielech, A. M.; Davis, N. M.; Baker, S. C.; Larsen, S. D.; Mesecar, A. D., X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *J Med Chem* **2014,** *57* (6), 2393-412.

28. Baez-Santos, Y. M.; St John, S. E.; Mesecar, A. D., The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral Res* **2015,** *115*, 21-38.

29. Freitas, B. T.; Durie, I. A.; Murray, J.; Longo, J. E.; Miller, H. C.; Crich, D.; Hogan, R. J.; Tripp, R. A.; Pegan, S. D., Characterization and Noncovalent Inhibition of the Deubiquitinase and deISGylase Activity of SARS-CoV-2 Papain-Like Protease. *ACS Infect Dis* **2020,** *6* (8), 2099-2109.

30. Klemm, T.; Ebert, G.; Calleja, D. J.; Allison, C. C.; Richardson, L. W.; Bernardini, J. P.; Lu, B. G.; Kuchel, N. W.; Grohmann, C.; Shibata, Y.; Gan, Z. Y.; Cooney, J. P.; Doerflinger, M.; Au, A. E.; Blackmore, T. R.; van der Heden van Noort, G. J.; Geurink, P. P.; Ovaa, H.; Newman, J.; Riboldi-Tunnicliffe, A.; Czabotar, P. E.; Mitchell, J. P.; Feltham, R.; Lechtenberg, B. C.; Lowes, K. N.; Dewson, G.; Pellegrini, M.; Lessene, G.; Komander, D., Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *EMBO J* **2020,** *39* (18), e106275.

31. Pitsillou, E.; Liang, J.; Ververis, K.; Lim, K. W.; Hung, A.; Karagiannis, T. C., Identification of Small Molecule Inhibitors of the Deubiquitinating Activity of the SARS-CoV-2 Papain-Like Protease: in silico Molecular Docking Studies and in vitro Enzymatic Activity Assay. *Front Chem* **2020,** *8*, 623971.

32. Ratia, K.; Pegan, S.; Takayama, J.; Sleeman, K.; Coughlin, M.; Baliji, S.; Chaudhuri, R.; Fu, W.; Prabhakar, B. S.; Johnson, M. E.; Baker, S. C.; Ghosh, A. K.; Mesecar, A. D., A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proc Natl Acad Sci U S A* **2008,** *105* (42), 16119-24.

33. Sander, A. L.; Yadouleton, A.; Moreira-Soto, A.; Tchibozo, C.; Hounkanrin, G.; Badou, Y.; Fischer, C.; Krause, N.; Akogbeto, P.; E, F. d. O. F.; Dossou, A.; Brunink, S.; Drosten, C.; Aissi, M. A. J.; Harouna Djingarey, M.; Hounkpatin, B.; Nagel, M.; Drexler, J. F., An Observational Laboratory-Based Assessment of SARS-CoV-2 Molecular Diagnostics in Benin, Western Africa. *mSphere* **2021,** *6* (1).

34. Baez-Santos, Y. M.; Mielech, A. M.; Deng, X.; Baker, S.; Mesecar, A. D., Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus. *J Virol* **2014,** *88* (21), 12511-27.

35. Chen, Y.; Savinov, S. N.; Mielech, A. M.; Cao, T.; Baker, S. C.; Mesecar, A. D., X-ray Structural and Functional Studies of the Three Tandemly Linked Domains of Non-structural Protein 3 (nsp3) from Murine Hepatitis Virus Reveal Conserved Functions. *J Biol Chem* **2015,** *290* (42), 25293-306.

36. Bekes, M.; Rut, W.; Kasperkiewicz, P.; Mulder, M. P.; Ovaa, H.; Drag, M.; Lima, C. D.; Huang, T. T., SARS hCoV papain-like protease is a unique Lys48 linkage-specific di-distributive deubiquitinating enzyme. *Biochem J* **2015,** *468* (2), 215-26.

37. Langley, C.; Goodwin, O.; Dzimianski, J. V.; Daczkowski, C. M.; Pegan, S. D., Structure of interferon-stimulated gene product 15 (ISG15) from the bat species Myotis davidii and the impact of interdomain ISG15 interactions on viral protein engagement. *Acta Crystallogr D Struct Biol* **2019,** *75* (Pt 1), 21-31.

38. Dzimianski, J. V.; Scholte, F. E. M.; Williams, I. L.; Langley, C.; Freitas, B. T.; Spengler, J. R.; Bergeron, E.; Pegan, S. D., Determining the molecular drivers of species-specific interferon-stimulated gene product 15 interactions with nairovirus ovarian tumor domain proteases. *PLoS One* **2019,** *14* (12), e0226415.

39. Deaton, M. K.; Dzimianski, J. V.; Daczkowski, C. M.; Whitney, G. K.; Mank, N. J.; Parham, M. M.; Bergeron, E.; Pegan, S. D., Biochemical and Structural Insights into the Preference of Nairoviral DeISGylases for Interferon-Stimulated Gene Product 15 Originating from Certain Species. *J Virol* **2016,** *90* (18), 8314-27.

40. Bekes, M.; van der Heden van Noort, G. J.; Ekkebus, R.; Ovaa, H.; Huang, T. T.; Lima, C. D., Recognition of Lys48-Linked Di-ubiquitin and Deubiquitinating Activities of the SARS Coronavirus Papain-like Protease. *Mol Cell* **2016,** *62* (4), 572-85.

41. Daczkowski, C. M.; Goodwin, O. Y.; Dzimianski, J. V.; Farhat, J. J.; Pegan, S. D., Structurally Guided Removal of DeISGylase Biochemical Activity from Papain-Like Protease Originating from Middle East Respiratory Syndrome Coronavirus. *J Virol* **2017,** *91* (23).

42. Rogers-Evans, M.; Knust, H.; Plancher, J. M.; Carreira, E. M.; Wuitschik, G.; Burkhard, J.; Li, D. B.; Guerot, C., Adventures in drug-like chemistry space: from oxetanes to spiroazetidines and beyond! *Chimia (Aarau)* **2014,** *68* (7-8), 492-9.

43. Ghosh, A. K.; Takayama, J.; Rao, K. V.; Ratia, K.; Chaudhuri, R.; Mulhearn, D. C.; Lee, H.; Nichols, D. B.; Baliji, S.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D., Severe Acute Respiratory Syndrome-Coronavirus Papain-Like Protease Inhibitors: Design, Synthesis, Protein-Ligand X-ray Structure and Biologial Evaluation *J. Med. Chem.* **2010,** *53*, 4968-4979.

44. Báez-Santos, y. M.; St. John, S. E.; Mesecar, A. D., The SARS-coronavirus Papain-Like Protease: Structure, function and Inhibition by Designed Antiviral Compounds. *Antiviral Res.* **2015,** *115*, 21-38.

45. Freitas, B. T.; Durie, I. A.; Crabtree, J.; Longo, J. E.; Miller, H. C.; Crich, D.; Hogan, R. J.; Tripp, R. A.; Pegan, S. D., Characterization and Noncovalent Inhibition of the Deubiquitinase and DeISGylase Activity of SARS-COV-2 Papain-Like Protease. *ACS Infect. Dis.* **2020,** *6*, 2099-2109.

46. Dhanju, S.; Upadhyaya, K.; Rice, C. A.; Pegan, S. D.; Media, J.; Valeriote, F. A.; Crich, D., Synthesis, Cytotoxicity and Genotoxicity of 10-Aza-9-oxakalkitoxin, An N,N,O-Trisubstituted Hydroxylamine Analog, or Hydroxalog, of a Marine Natural Product. *J. Am. Chem. Soc.* **2020,** *142*, 9147-9151.

47. Hill, J.; Hettikankanamalage, A. A.; Crich, D., Diversity-Oriented Synthesis of N,N,O-Trisubstituted Hydroxylamines from Alcohols and Amines by N-O Bond Formation. *J. Am. Chem. Soc.* **2020,** *142*, 14820-14825.

48. Portoghese, P. S.; Sultana, M.; Takemori, A. E., Design of peptidomimetic delta-opioid receptor antagonists using the message address concept. *J. Med. Chem.* **1990,** *33*, 1714-1720.

49. B., R.; Reddy S., R.; Reddy A., P.; Bandicchor, R., Amide Activation by TMSCl: Reduction of Amides to Amines by $LiAlH_4$ Under Mild Conditions. **2013,** *54*, 4908-4913.

50. Li, G.; Huan, Y.; Yuan, B.; Wang, J.; Jiang, Q.; Lin, Z.; Shen, Z.; Huang, H., Discovery of novel xanthine compounds targeting DPP-IV and GPR119 as anti-diabetic agents. *Eur. J. Med. Chem.* **2016,** *124*, 103-116.

51. Griffiths, D.; Johnstone, C. Preparation of Ureido-Carboxamido Thiophenes as Inhibitors of IKK2 Kinase. WO2003010163, 2003.

52. Palmer, A. M.; Chiesa, V.; Schmid, A.; Munch, G.; Grobbel, B.; Zimmermann, P. J.; Brehm, C.; Buhr, W.; Simon, W.-A.; Kromer, W.; Postius, S.; Volz, J.; Hess, D., Tetrahydrochromenoimidazoles as Potassium-Competitive Acid Blockers (P-CABs): Structure-Activity Relationship of Their Antisecretory Properties and Their Affinity toward the hERG Channel. *J. Med. Chem.* **2010,** *53*, 3645-3674.

53. Radchenko, D. S.; Grygorenko, O. O.; Komarov, I. V., Synthesis of 2-azaspiro[3.3]heptane-derived amino acids: ornitine and GABA analogues. *Amino Acids* **2010,** *39*, 515-521.

54. Lafrance, D.; Bowles, P.; Leeman, K.; Rafka, R., Mild Decarboxylative Activation of Malonic Acid Derivatives by 1,1'-Carbonyldiimidazole. *Org. Lett.* **2011,** *13*, 2322-2325.

55. Rogers-Evans, M.; Knust, H.; Plancher, J.-M.; Carreira, E. M.; Wuitschik, G.; Burkhard, J.; Li, D. B.; Guérot, C., Adventures in Drug-like Chemistry Space: From Oxetanes to Spiroazetidines and Beyond! *Chimia* **2014,** *68*, 492-499.

56. Burkhard, J. A.; Wagner, B.; Fischer, H.; Schuler, F.; Müller, K.; Carreira, E. M., Synthesis of Azaspirocycles and their Evaluation in Drug Discovery *Angew. Chem. Int. Ed.* **2010,** *49*, 3524-3527.

57. El Ashry, E. S. K., Y.; Hamid, H. A.; El-Zemity, S. R.; Boghdady, S., Synthesis of functionalised derivatives of pentaerythritol. *J. Chem. Res., Miniprint* **2003**, 111-128.

58. Marion, F.; Calvet, S.; Marié, J.-C.; Courillon, C.; Malacria, M., Silylated Vinyloxiranes – Recent Advances and Synthetic Applications. *Eur. J. Org. Chem.* **2006**, 453-462.

59. Marchetti, L. A.; Kumawat, L. K.; Mao, N.; Stephens, J. C.; Elmes, R. B. P., The Versatility of Squaramides: From Supramolecular Chemistry to Chemical Biology. *Chem* **2019,** *5*, 1398-1485.

60. Wurm, F. R.; Klok, H.-A., Be squared: expanding the horizon of squaric acid-mediated conjugations. *Chem. Soc. Rev.* **2013,** *42*, 8220-8236.

61. Motati, D. R.; Uredi, D.; Watkins, e. B., The Discovery and Development of Oxalamide and Pyrrole Small Molecule Inhibitors of gp120 and HIV Entry - A Review *Curr. Topics Med. Chem,* **2019,** *19*, 1650-1675.

62. Bostöm, J.; Hogner, A.; Llinàs, A.; Wellner, E.; Plowright, A. T., Oxadiazoles in Medicinal Chemistry. *J. Med. Chem.* **2012,** *55*, 1817-1830.

63. Kilianski, A.; Baker, S. C., Cell-based antiviral screening against coronaviruses: developing virus-specific and broad-spectrum inhibitors. *Antiviral Res* **2014,** *101*, 105-12.

64. Sridharan, H.; Zhao, C.; Krug, R. M., Species specificity of the NS1 protein of influenza B virus: NS1 binds only human and non-human primate ubiquitin-like ISG15 proteins. *J Biol Chem* **2010,** *285* (11), 7852-6.

65. Versteeg, G. A.; Hale, B. G.; van Boheemen, S.; Wolff, T.; Lenschow, D. J.; Garcia-Sastre, A., Species-specific antagonism of host ISGylation by the influenza B virus NS1 protein. *J Virol* **2010,** *84* (10), 5423-30.

66. Oreshkova, N.; Molenaar, R. J.; Vreman, S.; Harders, F.; Oude Munnink, B. B.; Hakze-van der Honing, R. W.; Gerhards, N.; Tolsma, P.; Bouwstra, R.; Sikkema, R. S.; Tacken, M. G.; de Rooij, M. M.; Weesendorp, E.; Engelsma, M. Y.; Bruschke, C. J.; Smit, L. A.; Koopmans, M.; van der Poel, W. H.; Stegeman, A., SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Euro Surveill* **2020,** *25* (23).

67. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **1997,** *276*, 307-26.

68. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 235-42.

69. 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* (Pt 4), 658-674.

70. 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 D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

71. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.

72. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 1), 12-21.

CHAPTER 5

AVIAN OLIGOADENYLATE SYNTHETASE-LIKE PROTEIN CONJUGATION AND IMPLICATIONS TO AVIAN IMMUNE REGULATION[4]

---

[4] Freitas, B.T., J.D. Shepard, S. Rodriguez, B.M. O'Boyle, K. Baker, J.E. Longo, H.C. Miller, E. Bergeron, and S.D. Pegan. To be submitted to *Frontiers in Immunology*

Abstract

Post-translational modification of host and viral proteins by ubiquitin and ubiquitin-like proteins plays a key role in a host's ability to mount an effective immune response. Avian species lack a ubiquitin-like protein, interferon stimulated gene product 15 (ISG15), found in mammals and other non-avian reptiles. ISG15 serves as a messenger molecule and can be conjugated to both host and viral proteins leading these proteins to be stabilized, degraded, or sequestered. Structurally, ISG15 is comprised of a tandem ubiquitin-like domain. Tandem ubiquitin-like domains are also found in 2'-5' oligoadenylate synthetase-like (OASL) near its C-terminus and bind OASL to retinoic acid inducible gene-I increasing its sensitivity to viral nucleic acids, resulting in enhanced production of type 1 interferons and robust immune response. Unlike human and other mammalian OASL homologs, avian OASLs terminate their tandem ubiquitin-like domains with the same LRLRGG motif found in ubiquitin and ISG15, a motif required for their conjugation to proteins. We show that the chicken OASL Ubl domain structurally shares characteristics with ISG15s and that these characteristics are likely conserved among avian OASLs. Here, the ISG15-like features of avian OASLs are explored, as well as how they impact interactions with viral deubiquitinases and deISGylases.

Introduction

Oligoadenylate synthetase (OAS) proteins are a group of enzymes that act as pattern recognition receptors (PRR)s and detect pathogen-associated molecular patterns (PAMPs)(1). Upon recognition of PAMPs, PRRs initiate signaling cascades that induce host defense mechanisms(2). One of the key PAMPs that PRRs detect is double-stranded RNA. The OAS family of proteins, a type of PRR, is made up of one or more OAS domains, at least one of which is catalytically active and contains an RNA binding site(3, 4). Following RNA binding, the OAS enzyme is activated and catalyzes the synthesis of 2'-5'-oligoadenylates (2-5A)s, which activate latent RNAse L to begin degradation of cytoplasmic RNAs(5, 6). Another member of the OAS family, OAS-like (OASL), is also shown to play a role in anti-viral mechanisms in response to double-stranded RNA, but with a different mechanism of action than its OAS counterpart(3, 4, 7).

OASLs are unique members of the OAS family in that they contain a single OAS domain that may be catalytically active or inactive across different species. The functional diversity of OASL is exemplified by the enzymatically active mouse OASL2 and its paralog, human OASL (hOASL), which shows no 2-5A synthetase activity(3, 5, 8). Mouse OASL1, however, is enzymatically inactive, while duck, goose, ostrich, and chicken OASLs maintain their 2-5A synthetase activity(5, 9). In addition to an OAS domain, OASLs also contain a tandem ubiquitin (Ub)-like protein (Ubl) domain at their C-terminus(3).

Ub is a post-translational modifier that is conjugated to target proteins at its C-terminal LRLRGG motif by E3 ligase proteins. In the presence of double-stranded RNA in humans, retinoic acid inducible gene-I (RIG-I) undergoes a conformational change that exposes its N-terminal caspase activation and recruitment domains (CARDS) which are then conjugated by K63 poly-Ub by the ligase, tripartite motif containing protein 25 (TRIM25), initiating a signaling cascade that upregulates type I interferons (IFN) and IFN stimulated genes (ISG)s such as hOASL(4). Following the upregulation of hOASL, its Ubl domain is thought to mimic poly-Ub and conjugates to RIG-I without TRIM25, enhancing the sensitivity of RIG-I(4, 10). Structurally, the Ubl domains of mammalian OASLs resemble that of native Ub but lack the LRLRGG motif that allows Ub to be conjugated to its target host and viral proteins. It is therefore unknown how hOASL mediates this interaction between the Ubl domain and RIG-I. This motif is found however on the Ubl domain of avian OASLs.

Unlike most mammalian OASLs, avian OASLs are catalytically active and act on the OAS/RNase L pathway, as well as induce RIG-I in a Ubl-dependent manner. The removal of the Ubl domains from avian OASLs destroys their ability to activate either pathway, whereas mammalian OASLs activate RNase L in a Ubl independent manner(11). The dual functionality of the avian OASLs may be due to the reduced number of OAS and OASL proteins avian species express compared to mammals. Most birds express only OASL with some *Ratitae* species, such as ostriches, expressing OASL and OAS1(12). Meanwhile, mammals can have up to three OAS proteins, up to two OASLs, and the NTase cyclic GMP-AMP synthetase (cGAS). Additionally, birds lack another important immunoregulatory Ubl, ISG15, that is involved in many aspects of the mammalian antiviral innate immune response(13). This reduced diversity of OAS proteins and increased reliance on Ubl domains potentially renders this a target of viral proteases for immune system avoidance in mammals.

Both Ub and ISG15 are important mediators of the mammalian antiviral innate immune response. They are involved in the regulation of many antiviral pathways including RIG-I activation, NF-κB inflammation, cytokine and chemokine production, and immune cell activation(14-17). Additionally, ubiquitination and ISGylation of viral proteins can cause them to be degraded or inactivated(18-20). Because of this, Ub and ISG15 have become targets of proteases such as the ovarian tumor domain proteases (OTU)s of Nairoviruses and papain-like proteases (PLpro) of coronaviruses as a means of suppressing their host's immune response(14, 21-23). These deubiquitinases (DUB)s and deISGylases reverse post-translational modifications by Ub and ISG15 by cleaving the conjugation created at their C-terminal LRLRGG motifs. These viral proteases usually preferentially cleave immunologically relevant poly-Ub chains and ISG15s from their virus' host species. In both cases, the proteases process the viral polypeptide into its functional units, however, by adapting to also reverse post-translational modifications by Ubls, they can delay initiation of the host's innate immune response long enough to replicate(24, 25).

Here we examine the structure of the domestic chicken OASL tandem Ubl domain and show that it contains features resembling those found in mammalian ISG15s. Analysis of OASL sequences from six diverse species indicates that these features are likely conserved among avian OASLs just as they are among ISG15s but are not shared by mammalian OASLs. Additionally, we examined the ability of OTUs and PLpros from a diverse group of viruses to cleave the Ubl domain of chicken OASL. Several OTUs that lack DUB and deISGylase activity were found to have moderate deOASLylase activity, with deOASLylase activity in both OTUs and PLpros mirroring viral host preferences.

Examination of the sites of interaction between chicken OASL and viral proteases revealed primary and tertiary structural factors responsible for the substrate preferences of Nairovirus OTUs and coronavirus PLpros

Materials and Methods

Chemicals and Reagents

Poly-ethylene glycol (PEG) 3350 was purchased from Sigma Life Sciences, tri-ammonium citrate was purchased from Sigma Life Sciences, Ampicillin was purchased from GoldBio, dehydrated Luria-Bertani Broth (LB) was purchased from Fisher Scientific, DL-dithiothroitol (DTT) was purchased from GoldBio, and isopropyl-β-d-thiogalactopyranoside (IPTG) was purchased from GoldBio. 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was purchased from Fisher BioReagents. Imidazole was purchased from Acros Organics; tris(hydroxymethyl)aminomethane (Tris) was purchased from Fisher Scientific. Sodium chloride (NaCl) was purchased from Fisher Chemical, and bovine serum albumin (BSA) was purchased from Sigma Life Science.

Construction, Expression, and Purification of Proteases and Ubls

The tandem Ubl domain of chicken OASL (342-476, 3-169) were cloned into pET-15b by Genscript and transformed into T7 express *E. coli*. Cells were cultured in 9 L of LB broth containing 100 mg/mL ampicillin at 37°C until the OD_{600} reached 0.6. Once reached, the expression was induced by the addition of 1 mM isopropyl β-d-thiogalactopyranoside (IPTG), and the culture was incubated at 18°C overnight. The culture was centrifuged at 12,000 g for 10 min, and then the pellet was collected and stored in a -80°C freezer. The cell pellet was dissolved into lysis buffer (500 mM NaCl and 50 mM Tris-HCl [pH = 7.0]) with lysozyme and then sonicated in Fisher Scientific series 150 on ice at 50% power with

5 s pulses for 6 min. The lysate was centrifuged at 64,000 g for 30 min to remove all insoluble products. The supernatant was then filtered and placed onto Ni-nitrilotriacetic agarose resin (Qiagen). The resin was washed using five column volumes of lysis buffer containing 10 mM imidazole. The protein was eluted using 5 column volumes of lysis buffer containing 300 mM imidazole. Thrombin was added to the elution to remove the 6X His-tag, and the combined solution was dialyzed in size exclusion buffer (200 mM NaCl, 50 mM Tris-HCl [pH = 7.0]) and run over a Size Exclusion Superdex 75 column (GE Healthcare, Pittsburgh PA). Purity was confirmed by gel electrophoresis. Nairovirus OTUs were expressed and purified as previously described(26).

Protease Activity Assay with proOASL

Activity assays of OTUs originating from Crimean-Congo hemorrhagic fever virus (CCHFV), Dugbe virus (DUGV), Erve virus (ERVEV), Nairobi sheep disease virus (NSDV), Ganjam virus (GANV), Taggart virus (TAGV), Qalyub virus (QYBV), Farallon virus (FARV), Huángpí tick virus 1 (HpTV-1), Issyk-kul virus (ISKV), Leopards Hill virus (LPHV), Dera Ghazi Khan virus (DGKV), Hazara virus (HAZV), and Kupe virus (KUPEV) with purified chicken OASL Ubl were adapted from previously reported methods(27). For 24 h, 10 mM OASL was incubated at 37°C with 20 nM of each Nairovirus OTU. At indicated timepoints 10 µL samples were taken from the reaction tubes and quenches in 2x Laemmli buffer and boiled at 98°C for 5 min. Samples were run on BioRad Mini-PROTEAN® TGX Stain-Free™ pre-cast gels. Visualization of timepoints relied on Stain-Free technology that enhances the fluorescence of endogenous tryptophan. The gels were UV-activated for five minutes and subsequently imaged in a BioRad ChemiDoc™ Imaging system according to the manufacturer's recommendations.

Crystallization of Chicken OASL Ubl Domains

The Ubl domain of chicken OASL was screened against a series of Qiagen NeXtal suites by hanging drop using a TTP Labtech Mosquito (TTP Labtech, Herfordshire, United Kingdom). The initial screens produced spindly, starburst shaped crystals from a condition containing 0.18 M tri-ammonium citrate and 20% w/v PEG 3350. This condition was then optimized by varying concentrations of PEG 3350 and tri-ammonium citrate as well as with additive screens. The final optimized crystal that the Ubl structure was collected from was cubic in shape and was generated through hanging drop with a final mother liquor of 0.18 M tri-ammonium citrate and 24% w/v PEG 3350, 16.3 mg/mL protein, and a 30% w/v galactose additive from Hampton research. The drop was 4 μL and contained a 4:1:5 ratio of mother liquor to additive to protein. The crystals were flash cooled in a cryoprotective solution containing 0.18 M tri-ammonium citrate and 28% w/v PEG 3350. The data set for chicken OASL Ubl domain was collected at the National Synchrotron Light Source II (Brookhaven National Laboratory, Upton, NY) on Life Science Biomedical Technology Research AMX beamline 17-ID-1 using a Eiger9M detector. Data were collected using wavelength 1 Å.

Data Processing and Structure Solution

All X-ray images were indexed, strategized, integrated, and scaled using HKL2000(28). To create a cross-validation set from a random 5% of the reflections to be used throughout refinement, the CCP4 software suite was employed(29). The initial phase solution for the structure of chicken OASL Ubl domain was obtained by molecular replacement via Phaser(30). A homology model of the Ubl domain based on mouse ISG15 (5CHW) was generated using MODELLER(31) for use as a search model. The structures were refined initially using Autobuild(32), then alternating rounds of manual editing in Coot(33), and automated refinement with Phenix(34). Molprobity was used to examine the final model of each structure to confirm the quality of the structures. The data collection and refinement statistics for each structure along are listed in Table S5.1. The structure of chicken OASL Ubl has been deposited in the protein data bank.

Results

Structural Analysis of chicken OASL tandem Ubl domains

Like Ub, OASL Ubls possess the ability to interact with the CARDs of RIG-I to initiate an antiviral signaling cascade in the presence of viral RNA(7, 10, 11, 35). To determine the degree to which OASL Ubl domains resemble poly-Ub or ISG15 an X-ray crystal structure was obtained of chicken OASL from residue 350 to 505. The structure was determined to a resolution of 2.23 Å in the space group $P2_1$ (Table S5.1). Molecular replacement phasing was performed using homology models of both β-grasp domains based on Ub. Each β-grasp was phased independently with linker regions built in later. Two copies of the OASL di-Ubl domain were identified in the asymmetric unit. Electron density was found for all but the final two glycine residues which were disordered and a small density gap within a flexible region of the N-terminal β-grasp between β-sheets three and four.

The OASL di-Ubl domain consists of two β-grasp folds (Figure 5.1). The N-terminal β-grasp contains four β-sheets that wrap around a three-turn α-helix, located between β-sheets two and three. Both β-grasps also contain two 3_{10}-helixes bracketing the third β-sheet, which is consistent with the secondary structure topology of Ub and ISG15. The C-terminal β-grasp consists of the same secondary structure features in the same relative positioning to one another, with the addition of a short fifth β-sheet. Resembling ISG15 more so than linear di-Ub, the two β-grasp folds are connected by a four-residue hinge region but closely associated with each other. In chicken OASL, the hinge consists of a T-E-P-Q motif that forms one internal, main chain hydrogen bond as well as a hydrogen bond to R477. The two β-grasp folds are oriented slightly closer together in OASL than in ISG15 and far more than linear di-Ub.

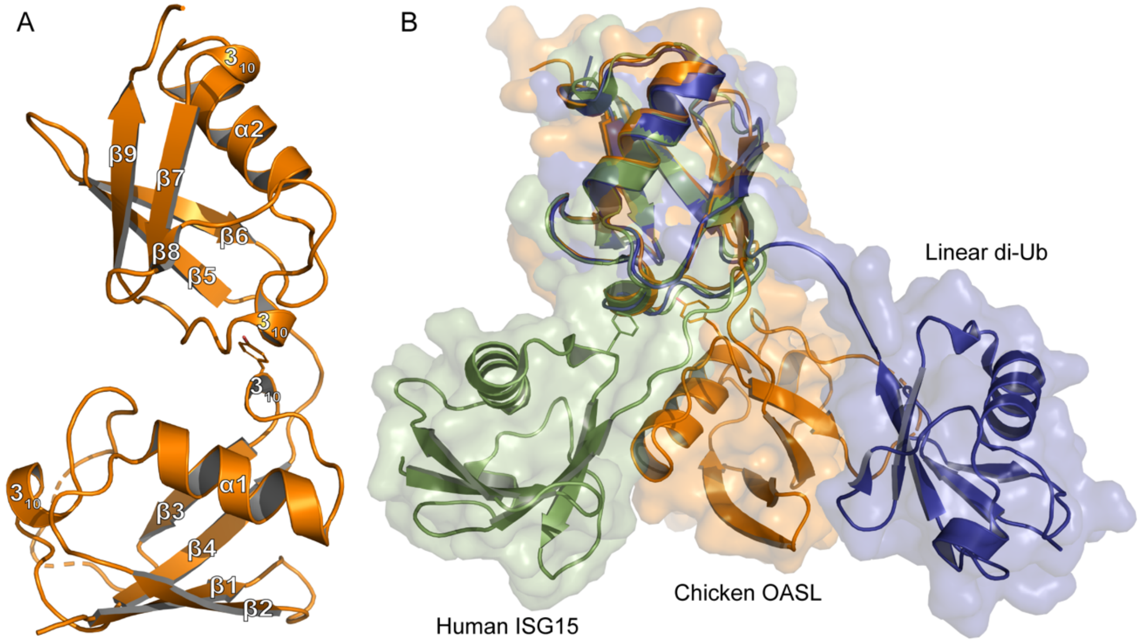


Figure 5.1. Tertiary structure comparison of immunoregulatory Ubls (A) Cartoon representation of the Ubl domain of chicken OASL (Orange) with secondary structure labels corresponding to those in Figure 5.4. (B) Cartoon and surface representations of chicken OASL Ubl domain, human ISG15 (Green)(PDB 1Z2M), and linear di-Ub (Blue)(PDB 2W9N) overlaid at their C-terminal β-grasps.

The close orientation of the folds is due largely to Y388 of the N-terminal β-grasp (Figure 5.2A). The aromatic ring of Y388 interacts with an extremely hydrophobic pocket within the C-terminal β-grasp of the Ubl domain, while the hydroxyl group forms hydrogen bonds with the main chain of V490. PISA analysis reveals this interface to be 281.1 $Å^2$ spanned by four hydrogen bonds(36). In addition to the hydrogen bonds between Y388 and V490, H489 forms bonds with R387 and W373. Because of these interactions, the two Ubls adopt a compact, rigid tandem Ubl conformation rather than two distinct Ubl domains. ISG15s have similar hydrophobic interactions between F41 and a hydrophobic pocket within their C-terminal β-grasp (Figure 5.2B). However, ISG15s do not appear to have conserved hydrogen bonds across the interface, and as a result, the two β-grasps are more flexible in their orientation to each other. In human ISG15 this interface is 182.2 $Å^2$ and is spanned by two hydrogen bonds from E139 to the main chain of F41. While E139 is highly conserved among mammalian ISG15s, it is not always oriented in a manner that would allow for interactions across the interface. When the C-terminal β-grasps are overlaid, the N-terminal β-grasp of chicken OASL is rotated approximately 65 degrees relative to human ISG15. The rotational difference between human ISG15 and chicken OASL resembles the difference between human ISG15 and bat ISG15, which is reported to be 76 degrees(37). However, despite the similarity in the rotation of the N-terminal domains of bat ISG15 and chicken OASL, the two do not overlap when the C-terminal domains are overlaid due to the more compact tertiary structure of chicken OASL.

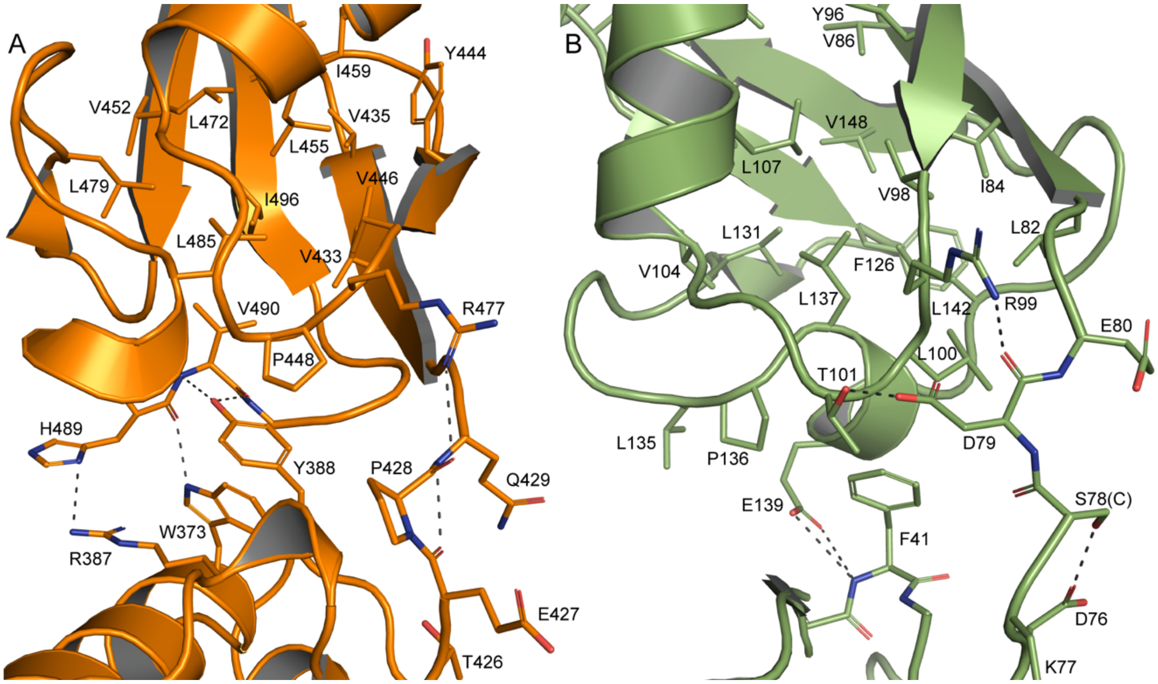


Figure 5.2. Interdomain interactions dictate Ubl β-grasp orientation (A) Interactions between the two β-grasps of chicken OASL (Orange) along with the connecting hinge region highlighting hydrogen bonds at the interface and hydrophobic interactions between Y388 and the C-terminal hydrophobic pocket. (B) Similar interactions between the two β-grasps of human ISG15 (Green).

Examination of the electrostatic surfaces of chicken OASL tandem Ubl domain confirms the presence of a negatively charged pocket between the β-grasps of OASL (Figure 5.3A). The two β-grasps are tightly packed around this charged pocket, making it difficult to distinguish between the two β-grasps in surface renderings. This differs from Ub and ISG15, which do not have highly charged inter-domain interfaces and have visibly distinct Ubl domains (Figure 5.3B). Most of the surface of the C-terminal β-grasp of chicken OASL is positively charged. Evaluation of the level of conservation of surface residues reveals that several of the most charged regions of OASL, including a positively charged pocket of the C-terminus, seem to be only moderately conserved.

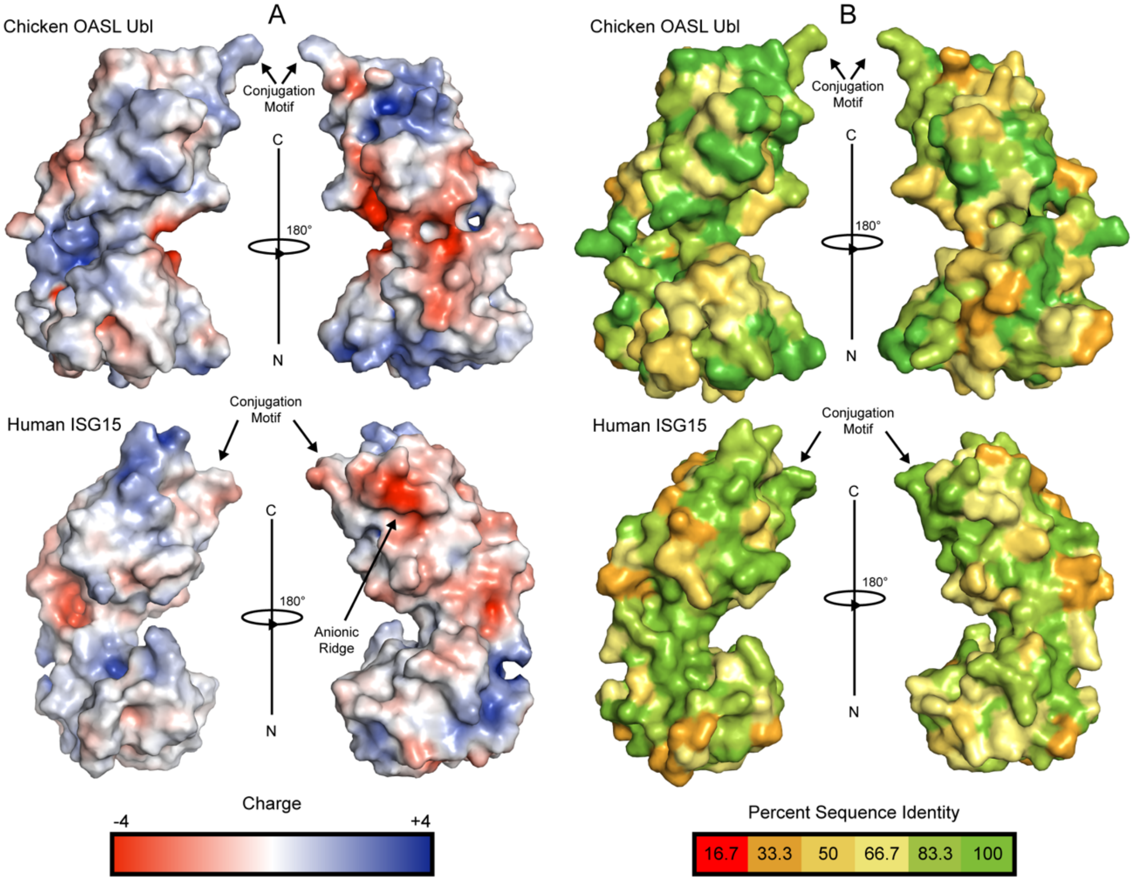
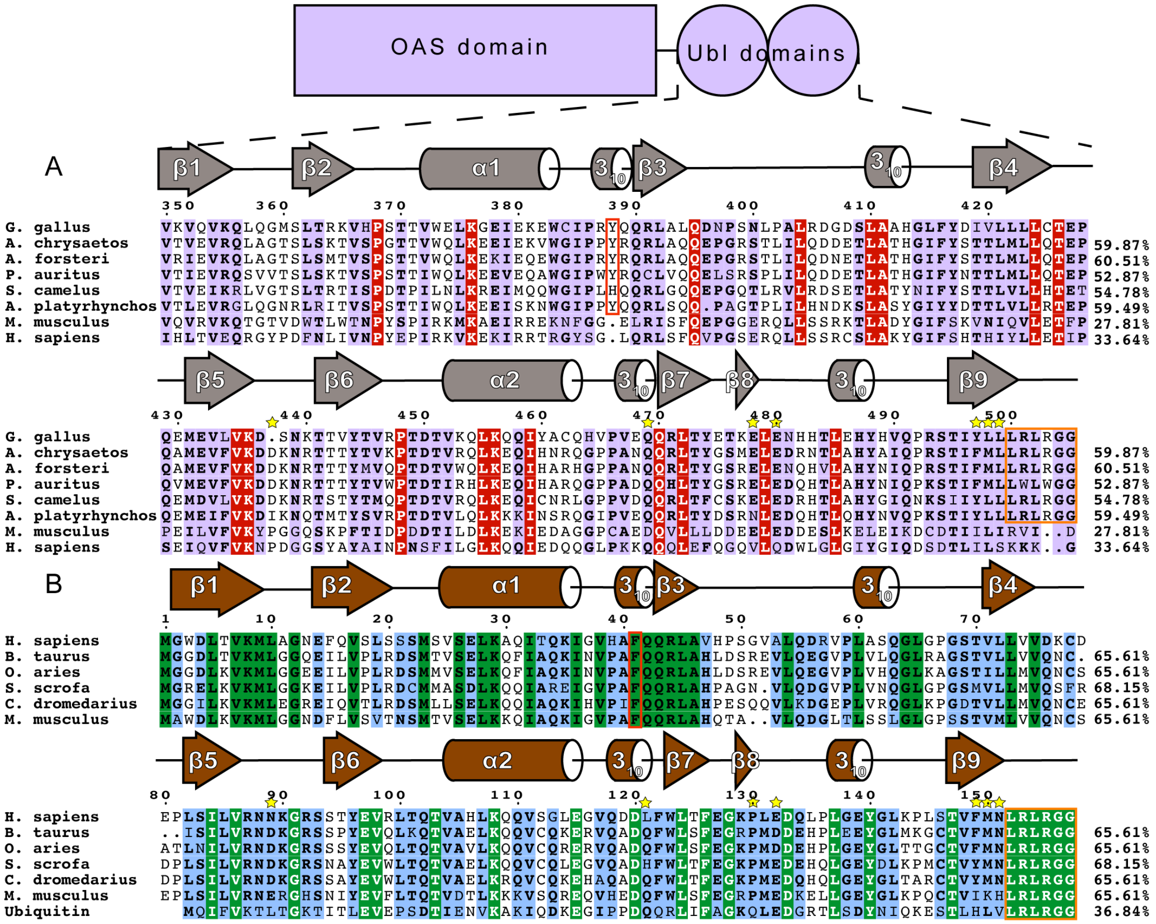


Figure 5.3. Electrostatic surface potential and residue conservation of chicken OASL Ubls and human ISG15. (A) Surface rendering of each Ubl with surface potential ranging from +4 to -4. Negatively charged regions are shown in red and positively charged regions are shown in blue. Potentials were generated using the PDB2PQR server and the surface was rendered using the adaptive Poisson-Boltzmann solver (APBS). (B) Surface rendering of each Ubl with surface residues color coded to represent degree of conservation across species based on the alignments in Figure 5.4. Color scale ranges from bright green for perfectly conserved residues to orange for residues that are one third conserved.

Interestingly, despite avian OASL Ubl domains being approximately 55 percent conserved as a whole, their hinge regions are extremely conserved at 95.8 percent. The same region of ISG15 is only about 54 percent conserved, despite ISG15s being approximately 60 percent conserved as a whole (Figure 5.4). Closer examination reveals that not only is this region highly conserved in OASLs, but it is likely far more rigid as well. In human ISG15 none of the residues are particularly rigid, and none of the three hydrogen bonds being formed within this region are likely to be conserved. Meanwhile, in avian OASLs both T426 and P428 are highly conserved and rigid residues. Additionally, at least one of the hydrogen bonds within this region is likely to be conserved.

Figure 5.4. Sequence and secondary structure comparison of OASL and ISG15 (A) Avian OASL Ubl domains aligned using ClustalW CLC Sequence Viewer. Percentages show the sequence identity relative to chicken OASL Ubl. Sequences displayed are from the following species: (*G. gallus*; NP_990372.2), Golden Eagle (*A. chrysaetos*; XP_029899442.1), Emperor Penguin (*A. forsteri*; XP_019326421.1), Double Crested Cormorant (*P. auratus*; SAMN06226698), Southern Ostrich (*S. camelus australis*; XP_009671383.1), Mallard (*A. platyrhynchos*; ARS01326.1), House Mouse (*M. musculus*; NP_035984.2), Human (*H. sapiens*; NP_003724.1). Ubl domain secondary structure based on Define Secondary Structure of Proteins (DSSP) algorithm calculations for chicken OASL is shown in gray. The aromatic residue found at the interface between the domains is boxed in red and the C-terminal LRLRGG conjugation motif is boxed in purple. Yellow stars indicate residues that form interactions at known OTU selectivity determination sites. (B) Mammalian ISG15s and Ubiquitin aligned using ClustalW CLC Sequence Viewer. Percentages show the sequence identity relative to human ISG15. Sequences displayed are from the following species: Human (*H. sapiens*; AAH09507.1), Cattle (*B. taurus*, NP_776791.1), Sheep (*O. aries,* AF152103.1), Boar (*S. scrofa*, ACB87600.1), Dromedary Camel (*C. dromedarius*, XP_010997700.1), House Mouse (*M. musculus*; AAB02697.1), and Ubiquitin (XP_022245348.1). General ISG15 secondary structure based on Define Secondary Structure of Proteins (DSSP) algorithm calculations for mouse ISG15 is shown in gray.



Sequence Analysis of Immunoregulatory Ubls

To identify which features of chicken OASL Ubl may be conserved across avian OASL, the sequences of six diverse bird species were examined. Human and mouse OASLs were also used for comparison. Upon examination of the OASL Ubl domains of these species it is apparent that some of the most highly conserved motifs of mammalian ISG15 and Ub are present in avian OASLs as well, but not in mammalian OASLs (Figure 5.4). Most notably, the LRLRGG conjugation motif is highly conserved in birds but completely absent in mammals. The presence and degree of conservation of this motif in avian OASLs, suggest a strong possibility that they are conjugating to target proteins similar to other Ubls. Conversely, the absence of this site in mammals makes it unlikely that they could conjugate to a target, indicating that if mammalian OASL Ubls are functional it would not be through conjugation in a Ub or ISG15-like manner.

Structural analysis of mammalian ISG15s reveals a critical phenylalanine residue at the interface of the two Ubl domains(37). In ISG15 F41 causes the two domains to more closely associate and has a profound impact on ISG15 tertiary structure(37). While F41 is not present in any OASLs, avian OASLs have a residue with an aromatic side chain that appears to fulfill the same purpose. All but one of the species examined have a tyrosine residue at the equivalent site. The lone outlier, ostrich OASL, has a histidine at that location. Human and mouse OASL do not contain obvious analogs. Both have aromatic side chains two positions upstream of their ISG15 counterparts, but the difference in location might place these residues outside of the domain interface. In addition to F41, ISG15s have a QQRLA motif at this site that makes up a 3_{10}-helix followed by a short β-sheet. This motif is fully conserved in Ub and the N-terminal domain of ISG15. In avian OASLs, it is moderately conserved on the N-terminal domain and highly conserved on the C-terminal domain. It is not well conserved on either domain of mammalian OASL.

Other similarities can be found in the degree and regions of homology between avian OASLs and ISG15s. In general, ISG15s share approximately 60% sequence identity even among distantly related mammals(26). Likewise, the base level of sequence identity between avian OASL Ubls appears to be 50-55%. The regions of conservation appear to be similar as well. In addition to those already mentioned, there are several sections of ISG15 that have highly conserved sequences across species, and while they are not all similar to the sequences found in avian OASLs, the same regions of OASL share high degrees of conservation internally.

Cleavage of chicken OASL tandem Ubl domains and Nairovirus OTUs

Given that avian OASLs possess high similarity to the C-terminal Ubl tail known to form isopeptide linkages with target proteins, OASLs may be conjugating to targets similar to Ub and ISG15. The sequence and structural similarities between avian OASL Ubls and ISG15s allow for the possibility that OASL could perform similar functions in the avian immune system to those left in the absence of ISG15. If conjugation of OASL Ubls to a target has antiviral effects, there may be an evolutionary pressure for some viruses to adapt the ability to counter this mechanism similar to viruses reversing ubiquitination and ISGylation. Similar to ISG15, avian OASLs are translated in an immature form with several amino acids downstream of their LRLRGG conjugation motif that would have to be cleaved off before conjugation. To explore if viral proteases may be able to process OASL, the pro-form of chicken OASL tandem Ubl domain was expressed with a C-terminal 6X His-Tag and was incubated with OTUs from 14 Nairovirus species for 24 hours with timepoint samples taken at specified intervals to determine approximate turnover (Figure 5.5). Of the 14 OTUs tested, 7 were capable of cleaving chicken OASL to some degree, and 4 totally cleaved the OASL within 24 hours. None fully processed the substrate in less than 2 hours. The rate of cleavage demonstrated by these OTUs is similar to what is seen when known deISGylases are incubated with ISG15s from non-host species(26, 38).

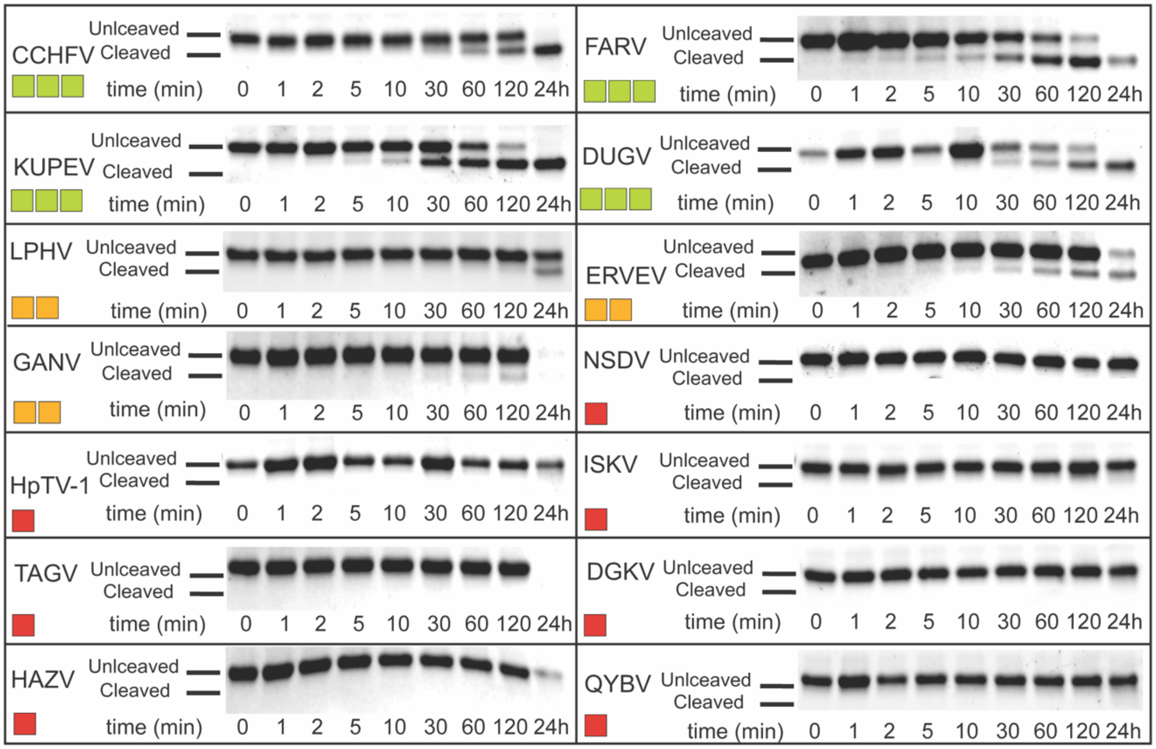


Figure 5.5. DeOASLylase activity of Nairovirus OTUs in chicken. OTUs from CCHFV, FARV, KUPEV, DUGV, LPHV, ERVEV, GANV, NSDV, HpTV-1, ISKV, TAGV, DGKV, HAZV, and QYBV were evaluated for their cleavage activity towards proOASL Ubl from chicken at 37°C, 10 µM of OASL Ubl was incubated with 20 nM of each OTU for at least 24 hr with samples taken at the time points indicated. The summary of chicken OASL cleavage by the different Nairovirus OTUs is presented as a heat map. Colors range from dark red (no cleavage) to light green (moderate cleavage).

Additionally, we tested the deOASLylase activity of six PLpros representing all four classes of coronaviruses for the ability to cleave chicken OASL (Figure 5.6). In general, coronavirus PLpros were less active against chicken OASL than Nairovirus OTUs. In both cases, half of the enzymes tested demonstrated no activity, however the OTUs that did demonstrate deOASLylase activity cleaved more than their PLpros counterparts. The three PLpros that demonstrated some degree of deOASLylase activity were found in each category except for alphacoronavirus. Meanwhile, all of the non-cleaving PLpros were found in either the alpha or betacoronavirus subgroup. One of the active chicken deOASLylating PLpros was from avian infectious bronchitis virus (IBV), which causes severe respiratory distress in chickens(39).

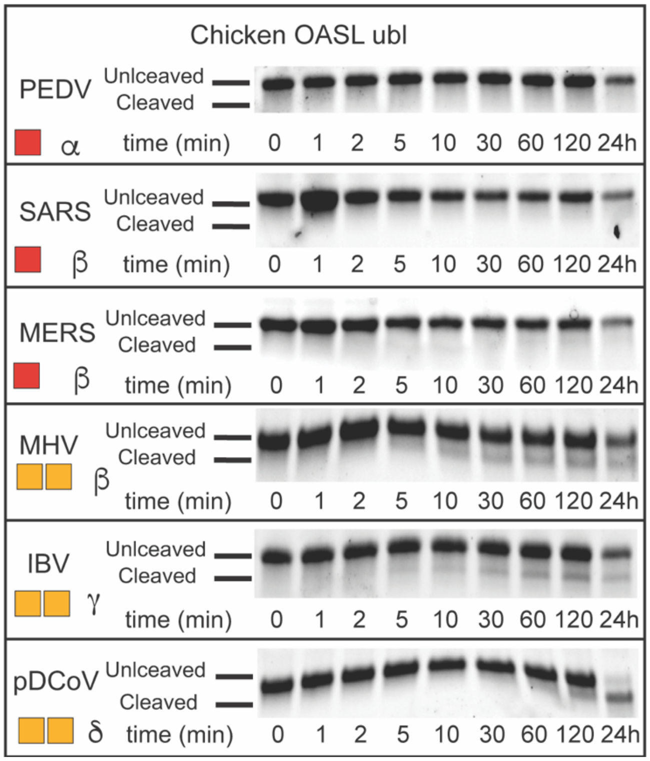


Figure 5.6. DeOASLylase activity of coronavirus PLpros in chicken. PLpros from PEDV, SARS, MERS, MHV, IBV, and pDCoV were evaluated for their cleavage activity towards proOASL Ubl from chicken at 37°C, 10 µM of OASL Ubl was incubated with 20 nM of each OTU for at least 24 hr with samples taken at the time points indicated. The summary of chicken OASL cleavage by the different coronavirus PLpros is presented as a heat map. Colors range from dark red (no cleavage) to orange (weak cleavage). The subgroup each coronavirus belongs to is denoted next to the respective heat map.

Discussion

Significance of Ubl conformation and Y388 in OASL activity

Compared to linear di-Ub, ISG15 has a relatively compact conformation but retains rotational flexibility around the interface between its domains. The β-grasps of chicken OASL's tandem Ubl domain are even more compact than those of most ISG15s and the presence of four hydrogen bonds at the interface likely reduces rotational flexibility. Sequence data suggests that three of the four hydrogen bonds are highly conserved among tandem Ubl motifs found in bird OASLs. As the hinge region of chicken OASL forms a hydrogen bond with a residue on the C-terminal domain as well as one within itself, it would likely have limited flexibility. The high conservation observed between avian OASL hinge regions suggests that the rigidity of the region would be conserved as well. Conversely, the hinge of ISG15 is flexible and highly variable(37). Interestingly, both chicken OASL and human ISG15 have an arginine at the same position on their C-terminal domains that forms a hydrogen bond with the hinge main chain, however, this arginine is not conserved among either OASLs or ISG15s. The arginine is found in only two of the six OASLs examined and is not seen on any of the other ISG15s examined here. However, other side chains with primary amines can be found at this site with glutamines being found at this site in two of the OASLs, and two of the ISG15s, as well as a lysine being present on one of the OASLs.

The increased number of interactions dictating OASL tertiary structure along with the higher degree of conservation seen in these residues suggests that this conformation of the Ubl domain may contribute to the function of avian OASL. The removal of the tandem Ubl domain from duck and ostrich OASL has been shown to hinder their ability to activate not only the RIG-I pathway but also the RNase L pathway by reducing their capacity to bind viral RNA(11). Mutation of F41 to a lysine in human ISG15 altered the conformation to the point where SARS-CoV-1 no longer bound the Ubl(37). If this tandem Ubl conformation is critical to OASL's involvement in either pathway it is likely that a mutation to Y388 could produce a similar loss of activity.

Potential implications of Nairovirus OTU deOASLylase activity

The ability of some Nairoviruses to productively process chicken OASL may be the result of off-target activity toward a similar substrate, or the result of an evolutionary pressure on Nairoviruses to counter OASL dependent immune responses. Both may be true depending on which virus and OTU are being examined. The OTU of the Ganjam virus (GANV) has robust DUB and deISGylase activity along with weak deOASLylase activity and is likely cleaving chicken OASL only due to its similarity to Ub and ISG15. On the other hand, the OTUs of FARV, KUPEV, and Dugbe virus (DUGV) all demonstrate moderate activity towards chicken OASL and have been shown not to have much activity toward Ub or human ISG15. It seems possible that for some of these viruses OASL may be a natural substrate along with DUB and deISGylase activity. Interestingly the OTU of CCHFV, which showed moderate chicken deOASLylase activity, has significant but relatively modest activity towards both Ub and human ISG15 when compared to OTUs that specialize in those substrates. While no OTU tested here demonstrated robust deOASLylase activity in chickens, the cleavage appears to match that seen when known deISGylases are incubated with ISG15 from species distantly related to their natural hosts.

Host species preferences may be responsible for reduced viral deOASLylase activity

Viral deISGylases have species preferences, usually aligning with the species they productively infect. For most OTUs and PLpros this results in the enzyme rapidly cleaving ISG15 from its host species along with highly conserved ISG15s, usually from closely related species. However, many retain the ability to process ISG15s from distantly related species at a reduced rate, similar to the rate chicken OASL was cleaved by several OTUs here. KUPEV and DUGV belong to the *Nairobi sheep disease* serogroup of OTUs and were isolated from ticks found on cattle. They are thought to infect bovine and ovine species, and unsurprisingly cleave domestic sheep and cattle ISG15 most rapidly, however, the rate that they cleave chicken OASL is similar to the rate at which they process camel ISG15. Similar trends are seen in the OTUs of CCHFV and ERVEV. FARV, the most active chicken deOASLylase, belongs to the *Hughes* serogroup, which is thought to primarily infect seabirds(40). If the natural host of FARV were to be identified, it is possible that its OASL would be a preferred substrate of FARV OTU.

Structural factors affecting deOASLylase activity of OTUs and PLpros

A functionally important similarity between chicken OASL and other immunoregulatory Ubls is the location of the C-terminal LRLRGG motif. As with Ub and ISG15, this motif of chicken OASL is solvent-exposed and would be readily accessible to viral proteases (Figure 5.7A). When chicken OASL is overlaid with a structure of sheep ISG15 in complex with the KUPEV OTU as well as FARV OTU, it appears that the OASL would be capable of forming some of the same interactions that allow OTUs to cleave ISG15s and Ub. Most crucially, the LRLRGG OTU cleavage site, while not properly oriented in this structure due to a lack of stabilizing interactions, would fit into the active site adjacent binding site (Figure 5.7C). Aside from the LRLRGG binding pocket, three key sites have been determined to be important for OTU selectivity(26). OASL forms similar interaction at two of the three and neither site is identical (Figure 5.7B). The differences at these recognition sites result in KUPEV OTU having significantly higher activity toward sheep ISG15 than chicken OASL. While the residues at these selectivity sites are not shared between ISG15s and avian OASLs they are highly conserved within their respective groups (Figure 5.1).

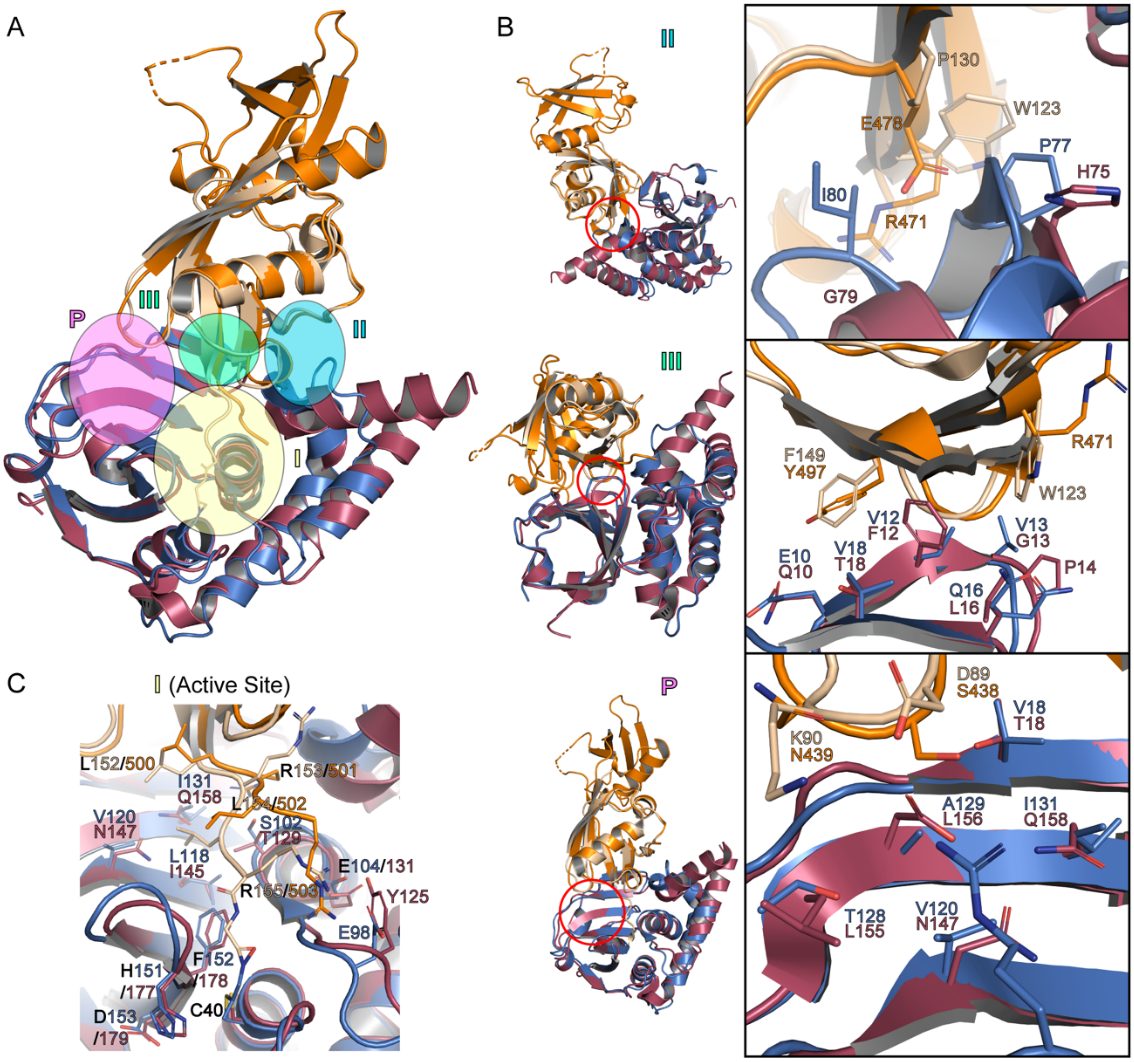


Figure 5.7. Interactions between Ubls and OTUs that determine substrate specificity (A) Cartoon representation of chicken OASL Ubl (Orange), FARV OTU (Purple) (PDB 6DX5), sheep ISG15 C-terminal domain (Beige) and KEPEV OTU (Blue) (PDB 6OAR). Ubls are overlaid at their C-terminal β-grasps and OTUs are overlaid with the Ubl LRLRGG motifs withing their active sites. (B) Close up views of OTU selectivity determination sites, highlighting interactions that determine substrate preferences. (C) Close of view of FARV and KUPEV OTU active sites with the C-terminus of Ubl substrates bound.

When the chicken OASL Ubl domain is similarly overlaid with a structure K48 linked di-Ub bound to the PLpro from SARS-CoV-1 we see that the orientation of the OASL β-grasps would prevent it from fitting into the active site of the PLpro (Figure 5.8A).

The orientation of the two OASL β-grasps to each other prevents both from fitting into their respective binding pockets while also fitting the C-terminal LRLRGG motif into the active site (Figure 5.8C). Furthermore, overlaying the C-terminal β-grasps of the two Ubls so both fit into the active site causes the N-terminal β-grasp of chicken OASL to overlap with the zinc finger of the PLpro (Figure 5.8B) in a way that would not be possible. In this case, it seems that the lack of flexibility between the two β-grasps protects the Ubl. Because poly-Ub, and to a lesser extent ISG15, have weaker interactions between their domains they are free to adopt conformations that are more favorable to cleavage by PLpros if those conformations are not already their natural state. Some of the increased deOASLylase activity OTUs have relative to PLpros may be attributable to them only interacting with Ubl substrates at the C-terminal β-grasp, so they cannot be blocked by the N-terminal domain.

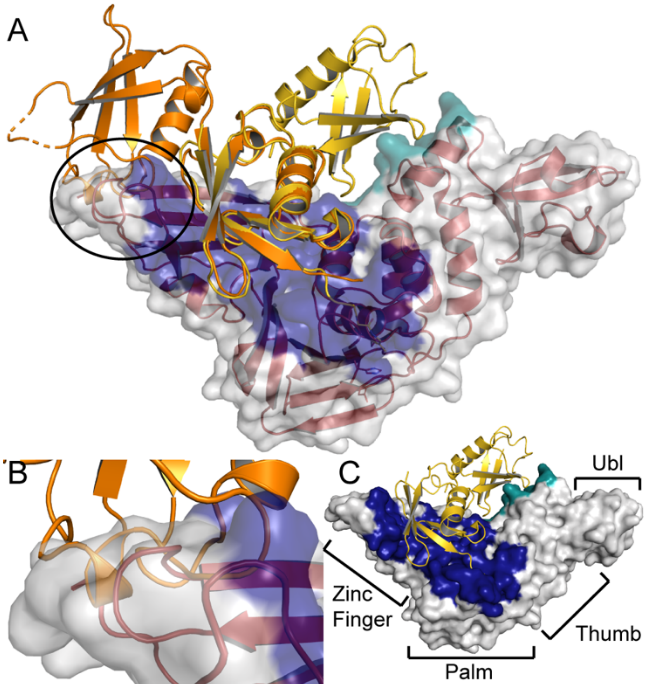


Figure 5.8. Steric hindrance at the zinc finger prevents SARS-CoV PLpro from cleaving chicken OASL (A) Chicken OASL Ubl domain (Orange) overlaid at the C-terminal β-grasp with K48 linked di-Ub (Yellow) bound to SARS-CoV PLpro (Red cartoon and white surface) (PDB 5E6J). (B) Close up view of the overlap between the PLpro zinc finger domain and N-terminal β-grasp of chicken OASL that causes a steric hindrance, preventing cleavage. (C) Surface view of SARS-CoV PLpro with the four domains labeled and the proximal Ub-binding pocket (Blue) and the distal Ub-binding pocket (Teal) highlighted.

References

1. Vajjhala PR, Ve T, Bentham A, Stacey KJ, Kobe B. The molecular mechanisms of signaling by cooperative assembly formation in innate immunity pathways. Mol Immunol. 2017;86:23-37.

2. Pandey S, Kawai T, Akira S. Microbial sensing by Toll-like receptors and intracellular nucleic acid sensors. Cold Spring Harb Perspect Biol. 2014;7(1):a016246.

3. Choi UY, Kang JS, Hwang YS, Kim YJ. Oligoadenylate synthase-like (OASL) proteins: dual functions and associations with diseases. Exp Mol Med. 2015;47:e144.

4. Zhu J, Ghosh A, Sarkar SN. OASL-a new player in controlling antiviral innate immunity. Curr Opin Virol. 2015;12:15-9.

5. Kristiansen H, Gad HH, Eskildsen-Larsen S, Despres P, Hartmann R. The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. J Interferon Cytokine Res. 2011;31(1):41-7.

6. Li XL, Ezelle HJ, Hsi TY, Hassel BA. A central role for RNA in the induction and biological activities of type 1 interferons. Wiley Interdiscip Rev RNA. 2011;2(1):58-78.

7. Ibsen MS, Gad HH, Andersen LL, Hornung V, Julkunen I, Sarkar SN, et al. Structural and functional analysis reveals that human OASL binds dsRNA to enhance RIG-I signaling. Nucleic Acids Res. 2015;43(10):5236-48.

8. Eskildsen S, Justesen J, Schierup MH, Hartmann R. Characterization of the 2'-5'-oligoadenylate synthetase ubiquitin-like family. Nucleic Acids Res. 2003;31(12):3166-73.

9. Tag-El-Din-Hassan HT, Morimatsu M, Agui T. Functional analysis of duck, goose, and ostrich 2'-5'-oligoadenylate synthetase. Infect Genet Evol. 2018;62:220-32.

10. Zhu J, Zhang Y, Ghosh A, Cuevas RA, Forero A, Dhar J, et al. Antiviral activity of human OASL protein is mediated by enhancing signaling of the RIG-I RNA sensor. Immunity. 2014;40(6):936-48.

11. Rong E, Wang X, Chen H, Yang C, Hu J, Liu W, et al. Molecular Mechanisms for the Adaptive Switching Between the OAS/RNase L and OASL/RIG-I Pathways in Birds and Mammals. Front Immunol. 2018;9:1398.

12. Tatsumi R, Sekiya S, Nakanishi R, Mizutani M, Kojima S, Sokawa Y. Function of ubiquitin-like domain of chicken 2'-5'-oligoadenylate synthetase in conformational stability. J Interferon Cytokine Res. 2003;23(11):667-76.

13. Magor KE, Miranzo Navarro D, Barber MR, Petkau K, Fleming-Canepa X, Blyth GA, et al. Defense genes missing from the flight division. Dev Comp Immunol. 2013;41(3):377-88.

14. Frieman M, Ratia K, Johnston RE, Mesecar AD, Baric RS. Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling. J Virol. 2009;83(13):6689-705.

15. Napolitano A, van der Veen AG, Bunyan M, Borg A, Frith D, Howell S, et al. Cysteine-Reactive Free ISG15 Generates IL-1beta-Producing CD8alpha(+) Dendritic Cells at the Site of Infection. J Immunol. 2018;201(2):604-14.

16. Recht M, Borden EC, Knight E, Jr. A human 15-kDa IFN-induced protein induces the secretion of IFN-gamma. J Immunol. 1991;147(8):2617-23.

17. Swaim CD, Scott AF, Canadeo LA, Huibregtse JM. Extracellular ISG15 Signals Cytokine Secretion through the LFA-1 Integrin Receptor. Mol Cell. 2017;68(3):581-90 e5.

18. Kim YJ, Kim ET, Kim YE, Lee MK, Kwon KM, Kim KI, et al. Consecutive Inhibition of ISG15 Expression and ISGylation by Cytomegalovirus Regulators. PLoS Pathog. 2016;12(8):e1005850.

19. Tang Y, Zhong G, Zhu L, Liu X, Shan Y, Feng H, et al. Herc5 attenuates influenza A virus by catalyzing ISGylation of viral NS1 protein. J Immunol. 2010;184(10):5777-90.

20. Kwon YT, Ciechanover A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. Trends Biochem Sci. 2017;42(11):873-86.

21. Devaraj SG, Wang N, Chen Z, Chen Z, Tseng M, Barretto N, et al. Regulation of IRF-3-dependent innate immunity by the papain-like protease domain of the severe acute respiratory syndrome coronavirus. J Biol Chem. 2007;282(44):32208-21.

22. Frias-Staheli N, Giannakopoulos NV, Kikkert M, Taylor SL, Bridgen A, Paragas J, et al. Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. Cell Host Microbe. 2007;2(6):404-16.

23. Holzer B, Bakshi S, Bridgen A, Baron MD. Inhibition of interferon induction and action by the nairovirus Nairobi sheep disease virus/Ganjam virus. PLoS One. 2011;6(12):e28594.

24. Denison MR, Perlman S. Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. J Virol. 1986;60(1):12-8.

25. Bergeron E, Albarino CG, Khristova ML, Nichol ST. Crimean-Congo hemorrhagic fever virus-encoded ovarian tumor protease activity is dispensable for virus RNA polymerase function. J Virol. 2010;84(1):216-26.

26. Dzimianski JV, Scholte FEM, Williams IL, Langley C, Freitas BT, Spengler JR, et al. Determining the molecular drivers of species-specific interferon-stimulated gene product 15 interactions with nairovirus ovarian tumor domain proteases. PLoS One. 2019;14(12):e0226415.

27. Deaton MK, Dzimianski JV, Daczkowski CM, Whitney GK, Mank NJ, Parham MM, et al. Biochemical and Structural Insights into the Preference of Nairoviral DeISGylases for Interferon-Stimulated Gene Product 15 Originating from Certain Species. J Virol. 2016;90(18):8314-27.

28. Otwinowski Z, Minor W. [20] Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997;276:307-26.

29. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):235-42.

30. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40(Pt 4):658-74.

31. Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics. 2016;54:5 6 1-5 6 37.

32. Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr D Biol Crystallogr. 2008;64(Pt 1):61-9.

33. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60(Pt 12 Pt 1):2126-32.

34. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):213-21.

35. Dhar J, Cuevas RA, Goswami R, Zhu J, Sarkar SN, Barik S. 2'-5'-Oligoadenylate Synthetase-Like Protein Inhibits Respiratory Syncytial Virus Replication and Is Targeted by the Viral Nonstructural Protein 1. J Virol. 2015;89(19):10115-9.

36. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372(3):774-97.

37. Langley C, Goodwin O, Dzimianski JV, Daczkowski CM, Pegan SD. Structure of interferon-stimulated gene product 15 (ISG15) from the bat species Myotis davidii and the impact of interdomain ISG15 interactions on viral protein engagement. Acta Crystallogr D Struct Biol. 2019;75(Pt 1):21-31.

38. Daczkowski CM, Dzimianski JV, Clasman JR, Goodwin O, Mesecar AD, Pegan SD. Structural Insights into the Interaction of Coronavirus Papain-Like Proteases and Interferon-Stimulated Gene Product 15 from Different Species. J Mol Biol. 2017;429(11):1661-83.

39. Cavanagh D. Coronavirus avian infectious bronchitis virus. Vet Res. 2007;38(2):281-97.

40. Converse JD, Hoogstraal H, Moussa MI, Feare CJ, Kaiser MN. Soldado virus (Hughes group) from Ornithodoros (Alectorobius) capensis (Ixodoidea: Argasidae) infesting Sooty Tern colonies in the Seychelles, Indian Ocean. Am J Trop Med Hyg. 1975;24(6 Pt 1):1010-8.

CHAPTER 6

DISCUSSION

PLpro DUB activity provides insight into immune evasion strategies

A considerable number of pathways, including those involved with the innate immune system, are regulated through post-translational modification by Ub and Ubls. Therefore, it can be difficult to discern which pathways individual DUBs and deISGylases are targeting. However, most Ub and Ubl regulated pathways are mediated by specific Ubls or poly-Ub linkage types, therefore the specificity of an enzyme can provide clues to the pathway or pathways it has adapted to counteract. For instance, it is common for viral DUBs to have preferences for K48 and K63 linked poly-Ub because of their association with proteasomal degradation and innate immune system signal transduction[103]. This can provide insight into which immune pathways are most detrimental to the virus's success.

While all three subgroup 2b PLpros examined here show a preference for the ISG15 of their natural host over poly-Ub they retain moderate to robust DUB activity towards long-chain K48 poly-Ub, with some activity toward long-chain K63 poly-Ub as well. K48 poly-Ub is associated with the proteasomal degradation pathway and is likely being targeted to keep viral proteins intact. When comparing long-chain K48 DUB activity, BtSCoV-Rf1.2004 cleaves the fastest of the three and SARS-CoV-2 cleaves the slowest. Interestingly, while BtSCoV-Rf1.2004 cleaves long-chain poly-Ub faster than the other two, SARS-CoV-1 cleaves mono and di-Ub the fastest of the three. SARS-CoV-2 processes all forms of Ub, including mono-Ub, more slowly than other subgroup 2b

PLpros. This appears to be due to PLpro residue 233 being a lysine in SARS-CoV-2 and a glutamine in all other subgroup 2b viruses. Interestingly, there are only two PLpro mutations among the emerging SARS-CoV-2 strains, A146D and K233Q. A146D likely has no functional impact on PLpro activity due to its location. However, K233Q is likely to bring SARS-CoV-2 DUB activity more in line with that of SARS-CoV-1 and BtSCoV-Rf1.2004 as residue 233 has previously been identified as a determinant of PLpro specificity. Mutation of SARS-CoV-1 Q233 to a glutamate resulted in DUB activity being halved and human deISGylase activity more than doubling[104]. As Ub is a highly conserved substrate across all possible coronavirus hosts and ISG15 can be highly variable, increasing human deISGylase activity at the expense of DUB activity may be detrimental unless a virus is highly specialized to infect humans. Therefore, retention of K48 DUB activity may indicate a selective pressure among subgroup 2b viruses towards being versatile in the hosts they infect and that proteasomal degradation of viral proteins is one of the major impediments to coronavirus replication. The presence of K233 in SARS-CoV-2 may have been an adaption that improved deISGylase specialization within its natural reservoir, and the K233Q mutation would likely increase DUB activity, helping it survive in its novel human hosts.

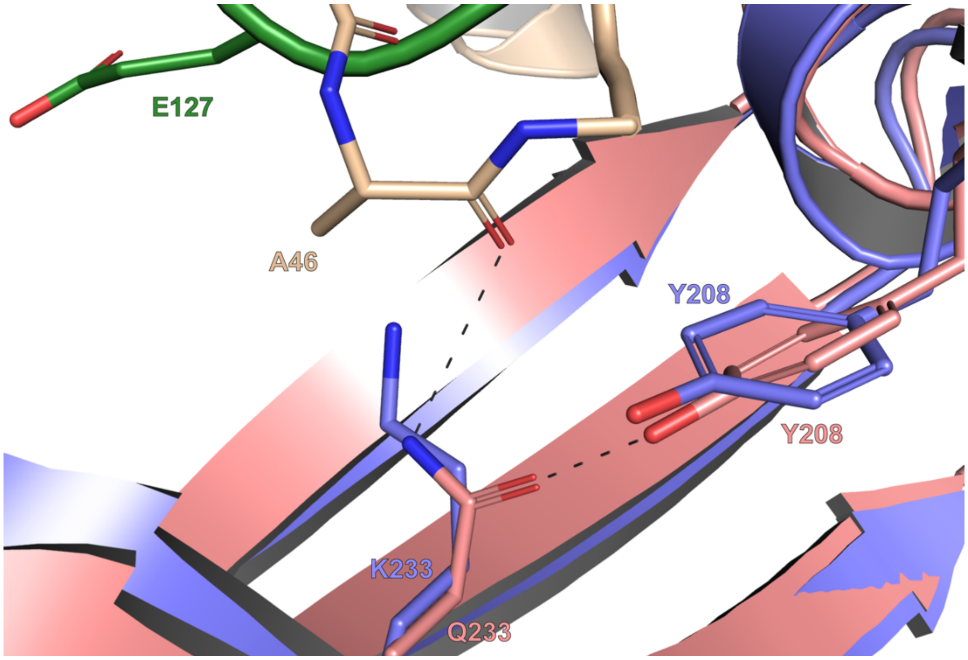


Figure 6.1. PLpro residue 233 Ubl interactions. Overlaid structure of SARS-CoV-1 PLpro (pink) bound to K48 di-Ub (tan), with SARS-CoV-2 PLpro (blue), and C-terminal human ISG15 (green).

In addition to suppressing the proteasomal degradation pathway through the reversal of K48 linked poly-ubiquitination, subgroup 2b PLpros retain mild to moderate K63 DUB activity. The level of K63 DUB activity varies between subgroup 2b viruses, usually proportional to K48 poly-Ub cleavage rates, but is generally less than what is seen in other betacoronaviruses such as MHV and MERS-CoV[36, 105]. K63 poly-Ub is an important mediator of signal transduction in many antiviral signaling cascades and reversal of K63 poly-ubiquitination would result in a weakened type-I IFN response[71, 89]. Based on a recent study, the type I IFN pathway is a target of the MHV PLP DUB activity as MHV PLP mutants with ablated DUB activity but functional deISGylase activity activated the IFN response earlier, leading to reduced pathogenesis[45]. This suggests that subgroup 2b viruses might also retain some level of K63 DUB activity to aid in circumventing the type-

I IFN response. However, as subgroup 2b viruses have lower K63 DUB activity than other betacoronaviruses, it is possible that the antiviral signaling cascades mediated by K63 poly-Ub are not as detrimental to subgroup 2b viruses as other coronaviruses. Testing this through the use of subgroup 2b viruses with PLpro mutations to ablate either deISGylase or DUB activity, similar to the experiment on DUB activity in MHV, would yield valuable data on both the role of deISGylase activity and K63 DUB activity in the subgroup 2b immunosuppression strategy.

Distinct characteristics of subgroup 2b PLpros

Despite retaining K48 and K63 poly-Ub cleavage activity, the PLpros of all three subgroup 2b viruses were examined (SARS-CoV-1, SARS-CoV-2, and BtSCoV-Rf1.2004) demonstrated a markedly reduced capacity to process mono-Ub and di-Ub of any linkage type. This cleavage pattern was discovered in SARS-CoV-1 in 2015, known as a di-distributive cleavage pattern, in which proteases recognize two Ub domains as a single unit for cleavage rather than mono-Ub [39]. On the other hand, PLpros with mono-distributive cleavage patterns, such as the betacoronaviruses MHV and MERS-CoV, recognize individual Ub domains as distinct units and can therefore cleave Ub regardless of chain length. As subgroup 2b PLpros have a di-distributive cleavage pattern, they recognize di-Ub within chains of three or more and rapidly process them into mono and di-Ub, but struggle to further process di-Ub into mono-Ub. It has been proposed that this recognition pattern is linked to the distal Ub binding pocket of subgroup 2b PLpros having stronger interactions with Ubls than that of PLpros that show a mono-distributive cleavage pattern[36, 39, 106]. This is supported by the observed interactions between the distal Ub binding pocket of all three subgroup 2b PLpros and Ub including stronger hydrophobic interactions as well as more hydrogen bonds than with MERS-CoV PLpro and MHV PLP2. As Ubls preferentially binding to the distal site would require chain lengths or three or more to reach the active site, enhanced interactions at the distal site could explain the preference of subgroup 2b PLpros towards longer chain Ubls.

Additionally, the di-distributive cleavage pattern may be an adaption to facilitate deISGylase specialization while retaining K48 DUB activity. Preferentially binding Ub at the distal Ub binding pocket over the proximal Ub binding pocket would make removal of di-Ubl modifications more efficient than removal of mono-Ubl modifications. A PLpro's capacity to remove mono-Ub would be reduced, however, it would not affect the removal of poly-Ub chains and could potentially aid in deISGylase activity.

Another feature that is unique to subgroup 2b PLpros, and highly conserved among them, is the ability to adopt a tucked Ubl domain conformation. Both SARS-CoV-1 and BtSCoV-Rf1.2004 PLpros have been shown to adopt this conformation, and it does not appear to be induced by crystal contacts in any structure in which it has appeared[104]. The conformation is stabilized by a seam of electrostatic interactions between the Ubl and thumb domains. Sequence analysis shows that the residues involved in these interactions are all highly conserved among subgroup 2b PLpros, but are not found outside of subgroup 2b. The degree of conservation of this feature may indicate a selective pressure to retain it, but it serves no obvious function and currently does not appear to interact with any substrate. Specifically, a published structure of SARS-CoV-1 PLpro bound to the C-terminal domain of mouse ISG15 demonstrated that the tucked Ubl domain does not form contacts with the Ubl substrate[104]. Despite this, the Ubl domain may play a role in PLpro stability as the removal of the Ubl domain of MHV PLP2 results in decreased thermal stability, ultimately leading to reduced activity and attenuated pathogenicity[35]. Therefore, adopting a tucked Ubl conformation may protect the domain and be beneficial for the stability of subgroup 2b PLpros.

Implications of coronavirus PLpro deISGylase species specificity

The relative rate at which PLpros cleave ISG15 from various species can provide information as to what hosts the virus may be infecting or has infected previously. For example, MHV PLP2 is highly specific for mouse ISG15, and to date, MHV has only been shown to productively infect mice[107]. Meanwhile, some coronaviruses with more versatile PLpros appear to infect a much wider range of hosts. For instance, MERS-CoV encodes a PLpro with moderate to robust activity against ISG15s from humans, mice, camels, bats, and fish[104]. While some of these are likely to be off-target effects, MERS-CoV has been found in humans, camels, bats, and alpacas in the wild and is capable of infecting marmots as well as several species of new and old-world monkeys in lab settings[108]. Similarly, all three subgroup 2b viruses tested so far have been shown to productively process ISG15s from a diverse range of species[104, 109] and to have an equally diverse group of natural reservoirs including humans, palm civets, pangolins, camels, marmosets, mice, cats, rhesus macaques, and several species of bats[2, 3, 108, 110].

It is unlikely that deISGylases capable of processing ISG15 from a wide range of species would be found in viruses with diverse hosts by chance. Viral immunosuppressive mechanisms are usually adapted to maximize efficiency within their preferred host species. For a coronavirus to encode a PLpro with moderate deISGylase activity toward a species it is likely that at one point or another it infected the species or a species with a similar ISG15. Subgroup 2b viruses are thought to primarily infect various bat species, and occasionally spill over into species that inhabit the same region. All three subgroup 2b PLpros examined here displayed robust activity toward vesper bat ISG15, and when BtSCoV-Rf1.2004 was incubated with two horseshoe bat ISG15s it fully processed both even more quickly than vesper bat. The species preferences of the three subgroup 2b PLpros are consistent, with all three cleaving most ISG15s at roughly the same rate. The similarity in deISGylase activity between PLpros from various subgroup 2b viruses suggests that if they are not already infecting many of the same species, they would likely be well suited to do so.

PLpro and OTU processing of unexplored Ubl post-translational modifiers

In addition to ISG15 and various forms of poly-Ub, some viral proteases may be targeting other immunologically relevant Ubls. Based on the cleavage demonstrated by various PLpros and OTUs, it seems possible that some of these proteases are not in fact DUBs or deISGylases, but are adapted to process another Ubl substrate and retain some off-target activity towards Ub and ISG15. This can be difficult to assess due to how specialized many deISGylases are to their preferred species, so it is possible that for many we simply have not identified their preferred host species. This is particularly likely for Nairoviruses, which are often isolated from ticks and not from host animals, so their natural hosts are often unknown. Along with, ISG15 there are several known Ubl post-translational modifiers that have been implicated in immunoregulation including oligoadenylate synthetase-like (OASL); small Ub-related modifier (SUMO); neural precursor cell-expressed, developmentally down-regulated 8 (Nedd8); and HLA-F adjacent transcript 10 (FAT10). OASL is thought to activate the RLR signaling pathway in response to dsRNA. SUMOylation of the IKK complex induces NF-κB signaling in response to DNA damage and blocking neddylation in macrophages has been shown to downregulate proinflammatory cytokine production by decreasing NF-κB signaling[111, 112]. Less is known about the function of FAT10, but it is induced in response to TNF-α and IFN-γ and is involved in modulating various aspects of the IFN response[113]. Like ISG15, FAT10 consists of two Ubl domains, whereas SUMO and Nedd8 resemble mono-Ub[114, 115]. All three are conjugated to target proteins at a C-terminal Gly-Gly motif, however, only Nedd8 contains upstream residues similar to Ub and ISG15 that are recognized by PLpros and OTUs proteases. Each is regulated by a system of ligases and proteases similar to Ub and ISG15, however, their role in immunity might make them targets for viral proteases as well.

As we have seen in cleavage experiments utilizing chicken OASL, several OTUs processed the substrate at a similar rate to that seen when known deISGylases process ISG15 from a species other than its preferred host. It's possible that some of these OTUs have adapted to process OASL, but are not chicken viruses and would more rapidly process the OASL of their natural host. Studies examining Ubl cleavage preferences by viral proteases found that CCHFV OTU does not appreciably process Nedd8, but that SARS-CoV-1 PLpro can process it at a similar rate to mono-Ub[116, 117]. Because Ubl turnover by viral proteases is highly specific to each virus, this turnover may not be reflective of their respective protease type. In the case of PLpros in particular, SARS-CoV-1 and other subgroup 2b PLpros follow a di-distributive cleavage pattern and would not be the ideal model for testing PLpro activity against a substrate resembling mono-Ub. Further testing of Nedd8 cleavage by more diverse OTUs and PLpros with mono-distributive cleavage patterns such as MERS-CoV may illuminate a novel mechanism of immunosuppression by these viruses.

Designing small molecule PLpro inhibitors for the treatment of novel coronaviruses

Due to the importance of PLpros to the success of coronaviruses, the development of non-covalent small molecule PLpro inhibitors is a promising strategy for treating SARS-CoV-2, as well as subgroup 2b viruses that emerge in the future. Naphthalene-based inhibitors have proven to be equally effective against all subgroup 2b viruses that they have been tested against. RNA viruses are particularly susceptible to mutation, yet sequence analysis shows that the residues forming interactions with these inhibitors are fully conserved within subgroup 2b. This suggests that these residues are critical to the function of subgroup 2b PLpros and that it is unlikely a novel subgroup 2b virus will emerge with mutations in the P3/P4 pocket that make it resistant to naphthalene-based inhibitors. While further optimization may be needed before use as therapeutics, these compounds have displayed minimal cytotoxicity in several human cell lines and have been shown to be effective at reducing replication of both SARS-CoV-1 and SARS-CoV-2 in Vero E6 cells.

Structure-activity relationship (SAR) analysis of series I compounds reveals key functional groups that are essential to inhibitor efficacy, as well as regions that allow for variation and optimization. The most critical interaction to maintain is the hydrophobic interaction of the western naphthyl group with the P4 pocket and Y269. Any mutations to these residues or substitutions of the western group result in a near-total loss of inhibitor efficacy[56]. Another major determinant of series I inhibitor efficacy is the backbone as compounds utilizing a simple amide backbone have proven most effective. This is due in part to their ability to form a hydrogen bond between their amide backbone and the main chain of the BL2 loop. When this hydrogen bond is lost through changes to the backbone, efficacy is diminished by up to 65 percent. These two groups appear to be indispensable for maintaining strong interactions between the series I compounds and subgroup 2b PLpros as changes to these moieties usually result in a loss of efficacy, therefore changing them is not likely to be beneficial in future optimization efforts.

Conversely, the eastern arene group appears to be the only series I functional group that can be modified without significant loss of activity. This group interacts at the P3 site and has, with several different ring decorations, demonstrated low µM IC_{50}s against subgroup 2b PLpro activity. So far, the 1-methyl-4-amino decoration of this ring has been the most effective, however further optimization of this site may improve binding specificity. It may be possible to improve series I inhibitors by modifying the 4-amino group to more closely mimic the arginine that naturally binds to the P3 site. Due to these factors, the eastern arene group would be an excellent site of focus in future attempts to optimize series I naphthalene-based inhibitors.

The optimal decoration of the eastern arene group is highly dependent on the backbone linker being used to attach it to the western naphthyl group. Unlike series I compounds, the backbones of series II compounds can be modified so long as the eastern arene group is modified accordingly. The most effective series II compounds can be categorized into those that utilize piperidine backbones and those that utilize spiroazetidine backbones. Within series II the difference in backbone length alters the location of the eastern arene group, resulting in a 2-methoxy-4-pyridyl group being optimal for both cytotoxicity and inhibition when paired with a spiroazetidine linkage, but equal to a 1,3-dioxolane group when paired with a piperidine linkage. The combination of a spiroazetidine backbone and a 2-methoxy-4-pyridyl arene group as in **37** results in an inhibitor with similar efficacy to top series II piperidine compounds, but with decreased cytotoxicity. By focusing on modifying backbone linkages in consort with arene rings we may be able to further optimize series I and series II therapeutics for the treatment of subgroup 2b viruses, which would safeguard against the possibility of a mutation emerging that renders one series ineffective.

Naphthalene-based compounds, such as **GRL0617**, **1**, **31**, and **37**, can be utilized as both chemical probes and lead compounds in future studies. These compounds have the potential to be useful chemical probes in the study of coronavirus biology through their ability to inhibit PLpro activity. This can be useful in examining the functions of novel or previously unexplored subgroup 2b PLpros and the NSPs they cleave. Furthermore, they are promising lead compounds for treating subgroup 2b coronavirus infections. With further optimization, they could become effective therapeutics for SARS-CoV-2 as well as future subgroup 2b coronaviruses.

REFERENCES

1. Ji, W.; Wang, W.; Zhao, X.; Zai, J.; Li, X., Cross-species transmission of the newly identified coronavirus 2019-nCoV. *J Med Virol* **2020,** *92* (4), 433-440.

2. Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J. H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H.; Zhang, J.; McEachern, J.; Field, H.; Daszak, P.; Eaton, B. T.; Zhang, S.; Wang, L. F., Bats are natural reservoirs of SARS-like coronaviruses. *Science* **2005,** *310* (5748), 676-9.

3. Wang, L. F.; Shi, Z.; Zhang, S.; Field, H.; Daszak, P.; Eaton, B. T., Review of bats and SARS. *Emerg Infect Dis* **2006,** *12* (12), 1834-40.

4. Hung, L. S., The SARS epidemic in Hong Kong: what lessons have we learned? *J R Soc Med* **2003,** *96* (8), 374-8.

5. LeDuc, J.; Barry, M. A., SARS, the First Pandemic of the 21st Century. *Emerging Infectious Disease journal* **2004,** *10* (11).

6. Chew, S. K., SARS: how a global epidemic was stopped. *Bull World Health Organ* **2007,** *85* (4), 324-324.

7. Zaki, A. M.; van Boheemen, S.; Bestebroer, T. M.; Osterhaus, A. D.; Fouchier, R. A., Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **2012,** *367* (19), 1814-20.

8. Nishiura, H.; Endo, A.; Saitoh, M.; Kinoshita, R.; Ueno, R.; Nakaoka, S.; Miyamatsu, Y.; Dong, Y.; Chowell, G.; Mizumoto, K., Identifying determinants of heterogeneous transmission dynamics of the Middle East respiratory syndrome (MERS) outbreak in the Republic of Korea, 2015: a retrospective epidemiological analysis. *BMJ Open* **2016,** *6* (2), e009936.

9. Nishiura, H.; Miyamatsu, Y.; Mizumoto, K., Objective Determination of End of MERS Outbreak, South Korea, 2015. *Emerg Infect Dis* **2016,** *22* (1), 146-8.

10. Park, H. Y.; Lee, E. J.; Ryu, Y. W.; Kim, Y.; Kim, H.; Lee, H.; Yi, S. J., Epidemiological investigation of MERS-CoV spread in a single hospital in South Korea, May to June 2015. *Euro Surveill* **2015,** *20* (25), 1-6.

11. Al-Omari, A.; Rabaan, A. A.; Salih, S.; Al-Tawfiq, J. A.; Memish, Z. A., MERS coronavirus outbreak: Implications for emerging viral infections. *Diagn Microbiol Infect Dis* **2019,** *93* (3), 265-285.

12. Wu, D.; Wu, T.; Liu, Q.; Yang, Z., The SARS-CoV-2 outbreak: What we know. *Int J Infect Dis* **2020,** *94*, 44-48.

13. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W.; China Novel Coronavirus, I.; Research, T., A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **2020,** *382* (8), 727-733.

14. Hozhabri, H.; Piceci Sparascio, F.; Sohrabi, H.; Mousavifar, L.; Roy, R.; Scribano, D.; De Luca, A.; Ambrosi, C.; Sarshar, M., The Global Emergency of Novel Coronavirus (SARS-CoV-2): An Update of the Current Status and Forecasting. *Int J Environ Res Public Health* **2020,** *17* (16).

15. Cui, J.; Li, F.; Shi, Z. L., Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **2019,** *17* (3), 181-192.

16. V'Kovski, P.; Kratzel, A.; Steiner, S.; Stalder, H.; Thiel, V., Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol* **2021,** *19* (3), 155-170.

17. Drosten, C.; Gunther, S.; Preiser, W.; van der Werf, S.; Brodt, H. R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A.; Berger, A.; Burguiere, A. M.; Cinatl, J.; Eickmann, M.; Escriou, N.; Grywna, K.; Kramme, S.; Manuguerra, J. C.; Muller, S.; Rickerts, V.; Sturmer, M.; Vieth, S.; Klenk, H. D.; Osterhaus, A. D.; Schmitz, H.; Doerr, H. W., Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **2003,** *348* (20), 1967-76.

18. Woo, P. C.; Lau, S. K.; Lam, C. S.; Lau, C. C.; Tsang, A. K.; Lau, J. H.; Bai, R.; Teng, J. L.; Tsang, C. C.; Wang, M.; Zheng, B. J.; Chan, K. H.; Yuen, K. Y., Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. *J Virol* **2012,** *86* (7), 3995-4008.

19. Lam, T. T. Y.; Jia, N.; Zhang, Y. W.; Shum, M. H. H.; Jiang, J. F.; Zhu, H. C.; Tong, Y. G.; Shi, Y. X.; Ni, X. B.; Liao, Y. S.; Li, W. J.; Jiang, B. G.; Wei, W.; Yuan, T. T.; Zheng, K.; Cui, X. M.; Li, J.; Pei, G. Q.; Qiang, X.; Cheung, W. Y. M.; Li, L. F.; Sun, F. F.; Qin, S.; Huang, J. C.; Leung, G. M.; Holmes, E. C.; Hu, Y. L.; Guan, Y.; Cao, W. C., Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **2020,** *583* (7815), 282-+.

20. Menachery, V. D.; Yount, B. L.; Debbink, K.; Agnihothram, S.; Gralinski, L. E.; Plante, J. A.; Graham, R. L.; Scobey, T.; Ge, X. Y.; Donaldson, E. F.; Randell, S. H.; Lanzavecchia, A.; Marasco, W. A.; Shi, Z. L. L.; Baric, R. S., A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med* **2015,** *21* (12), 1508-+.

21. Wong, N. A.; Saier, M. H., Jr., The SARS-Coronavirus Infection Cycle: A Survey of Viral Membrane Proteins, Their Functional Interactions and Pathogenesis. *Int J Mol Sci* **2021,** *22* (3).

22. Tahir Ul Qamar, M.; Alqahtani, S. M.; Alamri, M. A.; Chen, L. L., Structural basis of SARS-CoV-2 3CL(pro) and anti-COVID-19 drug discovery from medicinal plants. *J Pharm Anal* **2020,** *10* (4), 313-319.

23. EA, J. A.; Jones, I. M., Membrane binding proteins of coronaviruses. *Future Virol* **2019,** *14* (4), 275-286.

24. Malik, Y. A., Properties of Coronavirus and SARS-CoV-2. *Malays J Pathol* **2020,** *42* (1), 3-11.

25. Neuman, B. W.; Kiss, G.; Kunding, A. H.; Bhella, D.; Baksh, M. F.; Connelly, S.; Droese, B.; Klaus, J. P.; Makino, S.; Sawicki, S. G.; Siddell, S. G.; Stamou, D. G.; Wilson, I. A.; Kuhn, P.; Buchmeier, M. J., A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol* **2011,** *174* (1), 11-22.

26. Masters, P. S., The molecular biology of coronaviruses. *Adv Virus Res* **2006,** *66*, 193-292.

27. Li, W.; Moore, M. J.; Vasilieva, N.; Sui, J.; Wong, S. K.; Berne, M. A.; Somasundaran, M.; Sullivan, J. L.; Luzuriaga, K.; Greenough, T. C.; Choe, H.; Farzan, M., Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **2003,** *426* (6965), 450-4.

28. Wang, N.; Shi, X.; Jiang, L.; Zhang, S.; Wang, D.; Tong, P.; Guo, D.; Fu, L.; Cui, Y.; Liu, X.; Arledge, K. C.; Chen, Y. H.; Zhang, L.; Wang, X., Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res* **2013,** *23* (8), 986-93.

29. Tresnan, D. B.; Levis, R.; Holmes, K. V., Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J Virol* **1996,** *70* (12), 8669-74.

30. Fehr, A. R.; Perlman, S., Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* **2015,** *1282*, 1-23.

31. Mody, V.; Ho, J.; Wills, S.; Mawri, A.; Lawson, L.; Ebert, M.; Fortin, G. M.; Rayalam, S.; Taval, S., Identification of 3-chymotrypsin like protease (3CLPro) inhibitors as potential anti-SARS-CoV-2 agents. *Commun Biol* **2021,** *4* (1), 93.

32. Harcourt, B. H.; Jukneliene, D.; Kanjanahaluethai, A.; Bechill, J.; Severson, K. M.; Smith, C. M.; Rota, P. A.; Baker, S. C., Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J Virol* **2004,** *78* (24), 13600-12.

33. Baez-Santos, Y. M.; St John, S. E.; Mesecar, A. D., The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral Res* **2015,** *115*, 21-38.

34. Ratia, K.; Saikatendu, K. S.; Santarsiero, B. D.; Barretto, N.; Baker, S. C.; Stevens, R. C.; Mesecar, A. D., Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc Natl Acad Sci U S A* **2006,** *103* (15), 5717-22.

35. Mielech, A. M.; Deng, X.; Chen, Y.; Kindler, E.; Wheeler, D. L.; Mesecar, A. D.; Thiel, V.; Perlman, S.; Baker, S. C., Murine coronavirus ubiquitin-like domain is important for papain-like protease stability and viral pathogenesis. *J Virol* **2015,** *89* (9), 4907-17.

36. Baez-Santos, Y. M.; Mielech, A. M.; Deng, X.; Baker, S.; Mesecar, A. D., Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus. *J Virol* **2014,** *88* (21), 12511-27.

37. Bekes, M.; van der Heden van Noort, G. J.; Ekkebus, R.; Ovaa, H.; Huang, T. T.; Lima, C. D., Recognition of Lys48-Linked Di-ubiquitin and Deubiquitinating Activities of the SARS Coronavirus Papain-like Protease. *Mol Cell* **2016,** *62* (4), 572-85.

38. Kaiser, S. E.; Riley, B. E.; Shaler, T. A.; Trevino, R. S.; Becker, C. H.; Schulman, H.; Kopito, R. R., Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nat Methods* **2011,** *8* (8), 691-6.

39. Bekes, M.; Rut, W.; Kasperkiewicz, P.; Mulder, M. P.; Ovaa, H.; Drag, M.; Lima, C. D.; Huang, T. T., SARS hCoV papain-like protease is a unique Lys48 linkage-specific di-distributive deubiquitinating enzyme. *Biochem J* **2015,** *468* (2), 215-26.

40. Bailey-Elkin, B. A.; Knaap, R. C.; Johnson, G. G.; Dalebout, T. J.; Ninaber, D. K.; van Kasteren, P. B.; Bredenbeek, P. J.; Snijder, E. J.; Kikkert, M.; Mark, B. L., Crystal structure of the Middle East respiratory syndrome coronavirus (MERS-CoV) papain-like protease bound to ubiquitin facilitates targeted disruption of deubiquitinating activity to demonstrate its role in innate immune suppression. *J Biol Chem* **2014,** *289* (50), 34667-82.

41. Devaraj, S. G.; Wang, N.; Chen, Z.; Chen, Z.; Tseng, M.; Barretto, N.; Lin, R.; Peters, C. J.; Tseng, C. T.; Baker, S. C.; Li, K., Regulation of IRF-3-dependent innate immunity by the papain-like protease domain of the severe acute respiratory syndrome coronavirus. *J Biol Chem* **2007,** *282* (44), 32208-21.

42. Frieman, M.; Ratia, K.; Johnston, R. E.; Mesecar, A. D.; Baric, R. S., Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling. *J Virol* **2009,** *83* (13), 6689-705.

43. Clasman, J. R.; Everett, R. K.; Srinivasan, K.; Mesecar, A. D., Decoupling deISGylating and deubiquitinating activities of the MERS virus papain-like protease. *Antiviral Res* **2020,** *174*, 104661.

44. Daczkowski, C. M.; Goodwin, O. Y.; Dzimianski, J. V.; Farhat, J. J.; Pegan, S. D., Structurally Guided Removal of DeISGylase Biochemical Activity from Papain-Like Protease Originating from Middle East Respiratory Syndrome Coronavirus. *J Virol* **2017,** *91* (23).

45. Deng, X.; Chen, Y.; Mielech, A. M.; Hackbart, M.; Kesely, K. R.; Mettelman, R. C.; O'Brien, A.; Chapman, M. E.; Mesecar, A. D.; Baker, S. C., Structure-Guided Mutagenesis Alters Deubiquitinating Activity and Attenuates Pathogenesis of a Murine Coronavirus. *J Virol* **2020,** *94* (11).

46. Baez-Santos, Y. M.; Barraza, S. J.; Wilson, M. W.; Agius, M. P.; Mielech, A. M.; Davis, N. M.; Baker, S. C.; Larsen, S. D.; Mesecar, A. D., X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. *J Med Chem* **2014,** *57* (6), 2393-412.

47. Lee, H.; Lei, H.; Santarsiero, B. D.; Gatuz, J. L.; Cao, S.; Rice, A. J.; Patel, K.; Szypulinski, M. Z.; Ojeda, I.; Ghosh, A. K.; Johnson, M. E., Inhibitor recognition specificity of MERS-CoV papain-like protease may differ from that of SARS-CoV. *ACS Chem Biol* **2015,** *10* (6), 1456-65.

48. Chiu, M.; Bao, C.; Sadarangani, M., Dilemmas With Rotavirus Vaccine: The Neonate and Immunocompromised. *Pediatr Infect Dis J* **2019,** *38* (6S Suppl 1), S43-S46.

49. Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R., Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003,** *300* (5626), 1763-7.

50. Needle, D.; Lountos, G. T.; Waugh, D. S., Structures of the Middle East respiratory syndrome coronavirus 3C-like protease reveal insights into substrate specificity. *Acta Crystallogr D Biol Crystallogr* **2015,** *71* (Pt 5), 1102-11.

51. Abian, O.; Ortega-Alarcon, D.; Jimenez-Alesanco, A.; Ceballos-Laita, L.; Vega, S.; Reyburn, H. T.; Rizzuti, B.; Velazquez-Campoy, A., Structural stability of SARS-CoV-2 3CLpro and identification of quercetin as an inhibitor by experimental screening. *Int J Biol Macromol* **2020,** *164*, 1693-1703.

52. De Clercq, E., Potential antivirals and antiviral strategies against SARS coronavirus infections. *Expert Rev Anti Infect Ther* **2006,** *4* (2), 291-302.

53. Chen, X.; Chou, C. Y.; Chang, G. G., Thiopurine analogue inhibitors of severe acute respiratory syndrome-coronavirus papain-like protease, a deubiquitinating and deISGylating enzyme. *Antivir Chem Chemother* **2009,** *19* (4), 151-6.

54. Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M.; Olsen, S. K., Activity profiling and structures of inhibitor-bound SARS-CoV-2-PLpro protease provides a framework for anti-COVID-19 drug design. *bioRxiv* **2020**.

55. Shen, Z.; Ratia, K.; Cooper, L.; Kong, D.; Lee, H.; Kwon, Y.; Li, Y.; Alqarni, S.; Huang, F.; Dubrovskyi, O.; Rong, L.; Thatcher, G. R.; Xiong, R., Potent, Novel SARS-CoV-2 PLpro Inhibitors Block Viral Replication in Monkey and Human Cell Cultures. *bioRxiv* **2021**.

56. Shin, D.; Mukherjee, R.; Grewe, D.; Bojkova, D.; Baek, K.; Bhattacharya, A.; Schulz, L.; Widera, M.; Mehdipour, A. R.; Tascher, G.; Geurink, P. P.; Wilhelm, A.; van der Heden van Noort, G. J.; Ovaa, H.; Muller, S.; Knobeloch, K. P.; Rajalingam, K.; Schulman, B. A.; Cinatl, J.; Hummer, G.; Ciesek, S.; Dikic, I., Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **2020,** *587* (7835), 657-662.

57. Ghosh, A. K.; Samanta, I.; Mondal, A.; Liu, W. R., Covalent Inhibition in Drug Discovery. *ChemMedChem* **2019,** *14* (9), 889-906.

58. Ratia, K.; Pegan, S.; Takayama, J.; Sleeman, K.; Coughlin, M.; Baliji, S.; Chaudhuri, R.; Fu, W.; Prabhakar, B. S.; Johnson, M. E.; Baker, S. C.; Ghosh, A. K.; Mesecar, A. D., A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proc Natl Acad Sci U S A* **2008,** *105* (42), 16119-24.

59. Davis, M. E.; Gack, M. U., Ubiquitination in the antiviral immune response. *Virology* **2015,** *479-480*, 52-65.

60. Ivashkiv, L. B.; Donlin, L. T., Regulation of type I interferon responses. *Nat Rev Immunol* **2014,** *14* (1), 36-49.

61. Sadler, A. J.; Williams, B. R., Interferon-inducible antiviral effectors. *Nat Rev Immunol* **2008,** *8* (7), 559-68.

62. Vajjhala, P. R.; Ve, T.; Bentham, A.; Stacey, K. J.; Kobe, B., The molecular mechanisms of signaling by cooperative assembly formation in innate immunity pathways. *Mol Immunol* **2017,** *86*, 23-37.

63. Celhar, T.; Magalhaes, R.; Fairhurst, A. M., TLR7 and TLR9 in SLE: when sensing self goes wrong. *Immunol Res* **2012,** *53* (1-3), 58-77.

64. Kawai, T.; Akira, S., The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol* **2010,** *11* (5), 373-84.

65. Kawasaki, T.; Kawai, T., Toll-like receptor signaling pathways. *Front Immunol* **2014,** *5*, 461.

66. Blasius, A. L.; Beutler, B., Intracellular toll-like receptors. *Immunity* **2010,** *32* (3), 305-15.

67. Lei, J.; Hilgenfeld, R., RNA-virus proteases counteracting host innate immunity. *FEBS Lett* **2017,** *591* (20), 3190-3210.

68. Lin, S. C.; Lo, Y. C.; Wu, H., Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature* **2010,** *465* (7300), 885-90.

69. Jiang, Z.; Ninomiya-Tsuji, J.; Qian, Y.; Matsumoto, K.; Li, X., Interleukin-1 (IL-1) receptor-associated kinase-dependent IL-1-induced signaling complexes phosphorylate TAK1 and TAB2 at the plasma membrane and activate TAK1 in the cytosol. *Mol Cell Biol* **2002,** *22* (20), 7158-67.

70. Kollewe, C.; Mackensen, A. C.; Neumann, D.; Knop, J.; Cao, P.; Li, S.; Wesche, H.; Martin, M. U., Sequential autophosphorylation steps in the interleukin-1 receptor-associated kinase-1 regulate its availability as an adapter in interleukin-1 signaling. *J Biol Chem* **2004,** *279* (7), 5227-36.

71. Chen, Z. J., Ubiquitination in signaling to and activation of IKK. *Immunol Rev* **2012,** *246* (1), 95-106.

72. Ajibade, A. A.; Wang, H. Y.; Wang, R. F., Cell type-specific function of TAK1 in innate immune signaling. *Trends Immunol* **2013,** *34* (7), 307-16.

73. Wang, C.; Deng, L.; Hong, M.; Akkaraju, G. R.; Inoue, J.; Chen, Z. J., TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* **2001,** *412* (6844), 346-51.

74. Besse, A.; Lamothe, B.; Campos, A. D.; Webster, W. K.; Maddineni, U.; Lin, S. C.; Wu, H.; Darnay, B. G., TAK1-dependent signaling requires functional interaction with TAB2/TAB3. *J Biol Chem* **2007,** *282* (6), 3918-28.

75. Cheung, P. C.; Nebreda, A. R.; Cohen, P., TAB3, a new binding partner of the protein kinase TAK1. *Biochem J* **2004,** *378* (Pt 1), 27-34.

76. Ono, K.; Ohtomo, T.; Sato, S.; Sugamata, Y.; Suzuki, M.; Hisamoto, N.; Ninomiya-Tsuji, J.; Tsuchiya, M.; Matsumoto, K., An evolutionarily conserved motif in

the TAB1 C-terminal region is necessary for interaction with and activation of TAK1 MAPKKK. *J Biol Chem* **2001,** *276* (26), 24396-400.

77. Beinke, S.; Ley, S. C., Functions of NF-kappaB1 and NF-kappaB2 in immune cell biology. *Biochem J* **2004,** *382* (Pt 2), 393-409.

78. Hornung, V.; Hartmann, R.; Ablasser, A.; Hopfner, K. P., OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids. *Nat Rev Immunol* **2014,** *14* (8), 521-8.

79. Wu, X.; Wu, F. H.; Wang, X.; Wang, L.; Siedow, J. N.; Zhang, W.; Pei, Z. M., Molecular evolutionary and structural analysis of the cytosolic DNA sensor cGAS and STING. *Nucleic Acids Res* **2014,** *42* (13), 8243-57.

80. Ibsen, M. S.; Gad, H. H.; Andersen, L. L.; Hornung, V.; Julkunen, I.; Sarkar, S. N.; Hartmann, R., Structural and functional analysis reveals that human OASL binds dsRNA to enhance RIG-I signaling. *Nucleic Acids Res* **2015,** *43* (10), 5236-48.

81. Chebath, J.; Benech, P.; Revel, M.; Vigneron, M., Constitutive expression of (2'-5') oligo A synthetase confers resistance to picornavirus infection. *Nature* **1987,** *330* (6148), 587-8.

82. Hovanessian, A. G.; Wood, J. N., Anticellular and antiviral effects of pppA(2'p5'A)n. *Virology* **1980,** *101* (1), 81-90.

83. Schroder, H. C.; Suhadolnik, R. J.; Pfleiderer, W.; Charubala, R.; Muller, W. E., (2'-5')Oligoadenylate and intracellular immunity against retrovirus infection. *Int J Biochem* **1992,** *24* (1), 55-63.

84. Choi, U. Y.; Kang, J. S.; Hwang, Y. S.; Kim, Y. J., Oligoadenylate synthase-like (OASL) proteins: dual functions and associations with diseases. *Exp Mol Med* **2015,** *47*, e144.

85. Zhang, X.; Shi, H.; Wu, J.; Zhang, X.; Sun, L.; Chen, C.; Chen, Z. J., Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell* **2013,** *51* (2), 226-35.

86. Wu, J.; Sun, L.; Chen, X.; Du, F.; Shi, H.; Chen, C.; Chen, Z. J., Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* **2013,** *339* (6121), 826-30.

87. Kristiansen, H.; Gad, H. H.; Eskildsen-Larsen, S.; Despres, P.; Hartmann, R., The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. *J Interferon Cytokine Res* **2011,** *31* (1), 41-7.

88. Peisley, A.; Wu, B.; Xu, H.; Chen, Z. J.; Hur, S., Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I. *Nature* **2014,** *509* (7498), 110-4.

89. Okamoto, M.; Kouwaki, T.; Fukushima, Y.; Oshiumi, H., Regulation of RIG-I Activation by K63-Linked Polyubiquitination. *Front Immunol* **2017,** *8*, 1942.

90. Saha, S. K.; Pietras, E. M.; He, J. Q.; Kang, J. R.; Liu, S. Y.; Oganesyan, G.; Shahangian, A.; Zarnegar, B.; Shiba, T. L.; Wang, Y.; Cheng, G., Regulation of antiviral responses by a direct and specific interaction between TRAF3 and Cardif. *EMBO J* **2006,** *25* (14), 3257-63.

91. Zuin, A.; Isasa, M.; Crosas, B., Ubiquitin signaling: extreme conservation as a source of diversity. *Cells* **2014,** *3* (3), 690-701.

92. Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J., Structure of ubiquitin refined at 1.8 A resolution. *J Mol Biol* **1987,** *194* (3), 531-44.

93. Kwon, Y. T.; Ciechanover, A., The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem Sci* **2017,** *42* (11), 873-886.

94. Grice, G. L.; Nathan, J. A., The recognition of ubiquitinated proteins by the proteasome. *Cell Mol Life Sci* **2016,** *73* (18), 3497-506.

95. Hu, H.; Sun, S. C., Ubiquitin signaling in immune responses. *Cell Res* **2016,** *26* (4), 457-83.

96. Fu, B.; Wang, L.; Ding, H.; Schwamborn, J. C.; Li, S.; Dorf, M. E., TRIM32 Senses and Restricts Influenza A Virus by Ubiquitination of PB1 Polymerase. *PLoS Pathog* **2015,** *11* (6), e1004960.

97. Seissler, T.; Marquet, R.; Paillart, J. C., Hijacking of the Ubiquitin/Proteasome Pathway by the HIV Auxiliary Proteins. *Viruses* **2017,** *9* (11).

98. Zheng, N.; Shabek, N., Ubiquitin Ligases: Structure, Function, and Regulation. *Annu Rev Biochem* **2017,** *86*, 129-157.

99. Basters, A.; Knobeloch, K. P.; Fritz, G., USP18 - a multifunctional component in the interferon response. *Biosci Rep* **2018,** *38* (6).

100. Ye, Y.; Rape, M., Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **2009,** *10* (11), 755-64.

101. Zou, W.; Zhang, D. E., The interferon-inducible ubiquitin-protein isopeptide ligase (E3) EFP also functions as an ISG15 E3 ligase. *J Biol Chem* **2006,** *281* (7), 3989-94.

102. Zou, W.; Wang, J.; Zhang, D. E., Negative regulation of ISG15 E3 ligase EFP through its autoISGylation. *Biochem Biophys Res Commun* **2007,** *354* (1), 321-7.

103. Dzimianski, J. V.; Beldon, B. S.; Daczkowski, C. M.; Goodwin, O. Y.; Scholte, F. E. M.; Bergeron, E.; Pegan, S. D., Probing the impact of nairovirus genomic diversity on viral ovarian tumor domain protease (vOTU) structure and deubiquitinase activity. *PLoS Pathog* **2019,** *15* (1), e1007515.

104. Daczkowski, C. M.; Dzimianski, J. V.; Clasman, J. R.; Goodwin, O.; Mesecar, A. D.; Pegan, S. D., Structural Insights into the Interaction of Coronavirus Papain-Like Proteases and Interferon-Stimulated Gene Product 15 from Different Species. *J Mol Biol* **2017,** *429* (11), 1661-1683.

105. Chen, Y.; Savinov, S. N.; Mielech, A. M.; Cao, T.; Baker, S. C.; Mesecar, A. D., X-ray Structural and Functional Studies of the Three Tandemly Linked Domains of Non-structural Protein 3 (nsp3) from Murine Hepatitis Virus Reveal Conserved Functions. *J Biol Chem* **2015,** *290* (42), 25293-306.

106. Ratia, K.; Kilianski, A.; Baez-Santos, Y. M.; Baker, S. C.; Mesecar, A., Structural Basis for the Ubiquitin-Linkage Specificity and deISGylating activity of SARS-CoV papain-like protease. *PLoS Pathog* **2014,** *10* (5), e1004113.

107. Korner, R. W.; Majjouti, M.; Alcazar, M. A. A.; Mahabir, E., Of Mice and Men: The Coronavirus MHV and Mouse Models as a Translational Approach to Understand SARS-CoV-2. *Viruses* **2020,** *12* (8).

108. Gong, S. R.; Bao, L. L., The battle against SARS and MERS coronaviruses: Reservoirs and Animal Models. *Animal Model Exp Med* **2018,** *1* (2), 125-133.

109. Freitas, B. T.; Durie, I. A.; Murray, J.; Longo, J. E.; Miller, H. C.; Crich, D.; Hogan, R. J.; Tripp, R. A.; Pegan, S. D., Characterization and Noncovalent Inhibition of the Deubiquitinase and deISGylase Activity of SARS-CoV-2 Papain-Like Protease. *ACS Infect Dis* **2020,** *6* (8), 2099-2109.

110. Centers for Disease, C.; Prevention, Prevalence of IgG antibody to SARS-associated coronavirus in animal traders--Guangdong Province, China, 2003. *MMWR Morb Mortal Wkly Rep* **2003,** *52* (41), 986-7.

111. Stilmann, M.; Hinz, M.; Arslan, S. C.; Zimmer, A.; Schreiber, V.; Scheidereit, C., A nuclear poly(ADP-ribose)-dependent signalosome confers DNA damage-induced IkappaB kinase activation. *Mol Cell* **2009,** *36* (3), 365-78.

112. Yao, J.; Liang, X.; Liu, Y.; Zheng, M., Neddylation: A Versatile Pathway Takes on Chronic Liver Diseases. *Front Med (Lausanne)* **2020,** *7*, 586881.

113. Mah, M. M.; Roverato, N.; Groettrup, M., Regulation of Interferon Induction by the Ubiquitin-Like Modifier FAT10. *Biomolecules* **2020,** *10* (6).

114. Aichem, A.; Anders, S.; Catone, N.; Rossler, P.; Stotz, S.; Berg, A.; Schwab, R.; Scheuermann, S.; Bialas, J.; Schutz-Stoffregen, M. C.; Schmidtke, G.; Peter, C.; Groettrup, M.; Wiesner, S., The structure of the ubiquitin-like modifier FAT10 reveals an alternative targeting mechanism for proteasomal degradation. *Nat Commun* **2018,** *9* (1), 3321.

115. Theng, S. S.; Wang, W.; Mah, W. C.; Chan, C.; Zhuo, J.; Gao, Y.; Qin, H.; Lim, L.; Chong, S. S.; Song, J.; Lee, C. G., Disruption of FAT10-MAD2 binding inhibits tumor progression. *Proc Natl Acad Sci U S A* **2014,** *111* (49), E5282-91.

116. Capodagli, G. C.; McKercher, M. A.; Baker, E. A.; Masters, E. M.; Brunzelle, J. S.; Pegan, S. D., Structural analysis of a viral ovarian tumor domain protease from the Crimean-Congo hemorrhagic fever virus in complex with covalently bonded ubiquitin. *J Virol* **2011,** *85* (7), 3621-30.

117. Lindner, H. A.; Lytvyn, V.; Qi, H.; Lachance, P.; Ziomek, E.; Menard, R., Selectivity in ISG15 and ubiquitin recognition by the SARS coronavirus papain-like protease. *Arch Biochem Biophys* **2007,** *466* (1), 8-14.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

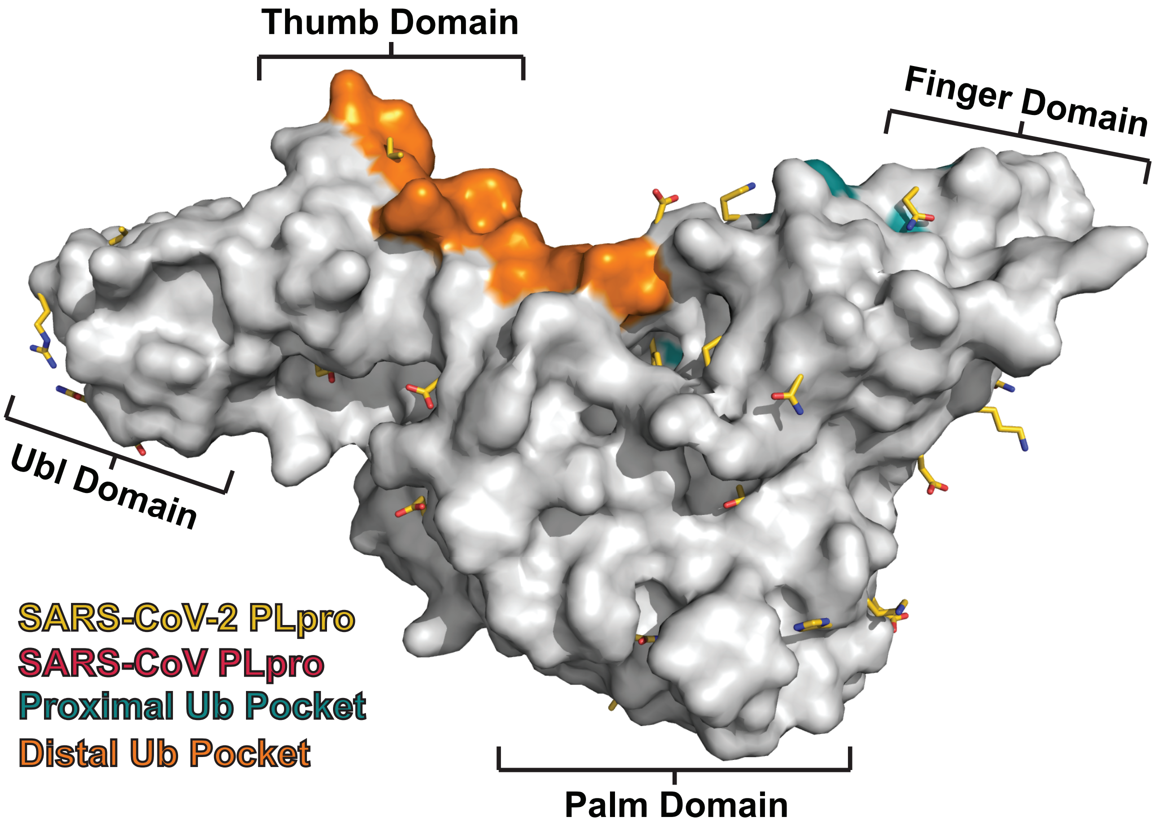


Figure S3.1. Model of the PLpro of SARS-CoV-2. A surface rendering of a homology model for SARS-CoV-2 PLpro (white), highlighting the proximal Ub binding pocket (teal) and the distal Ub binding pocket (orange) built from SARS-CoV PLpro (PDB 5E6J). Residues differing between SARS-CoV and SARS-CoV-2 are shown in yellow.

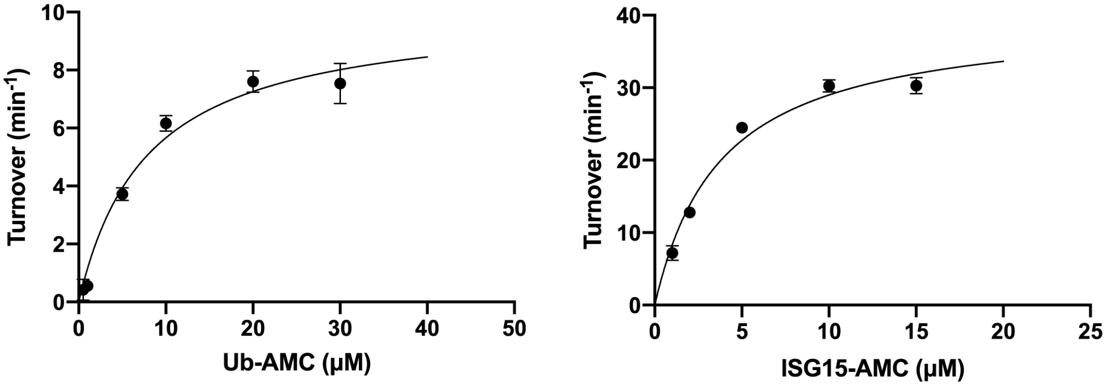


Figure S3.2. SARS-CoV-2 PLpro activities with substrates. (left panel) Activity of SARS-CoV-2 PLpro at varying concentrations of Ub-AMC. (right panel) Activity of SARS-CoV-2 PLpro at varying concentrations of ISG15-AMC.

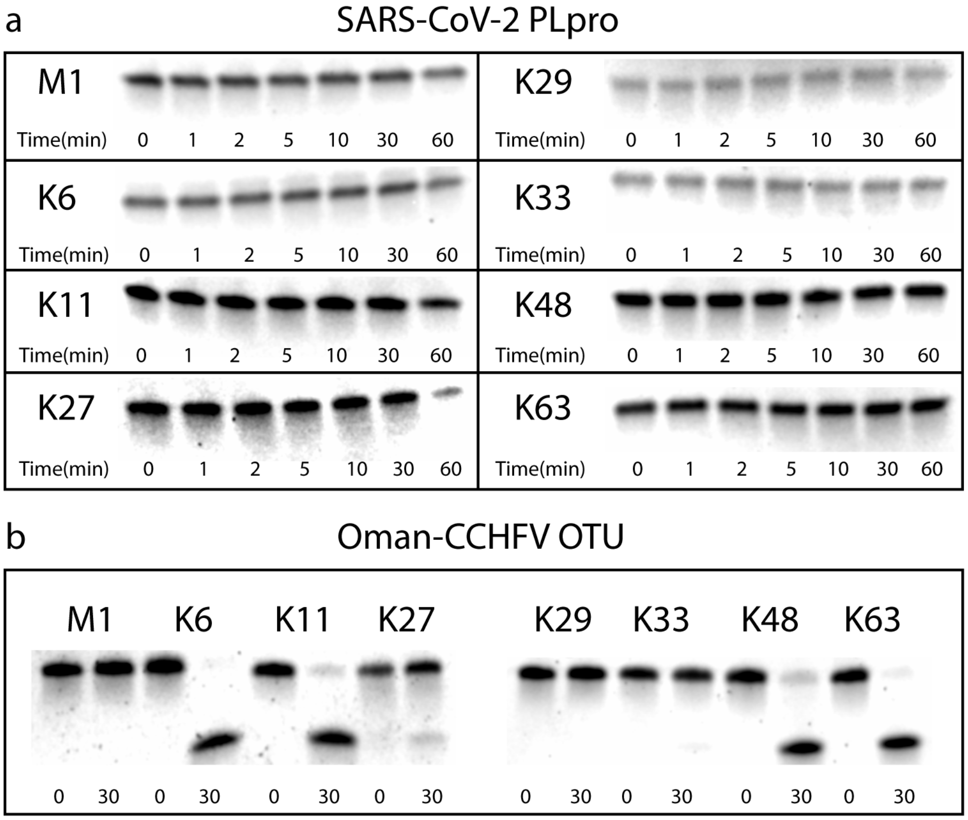


Figure S3.3. Gel cleavage assay of unlabeled di-Ub, visualized by Coomassie Blue staining. (a) Assays were performed at 37℃ with 10μM di-Ub and 20 nM SARS-CoV-2 PLpro. (b) Oman CCHFV OTU control performed at 37℃ with 13.7 μM di-Ub and 23 nM Oman CCHFV OTU.

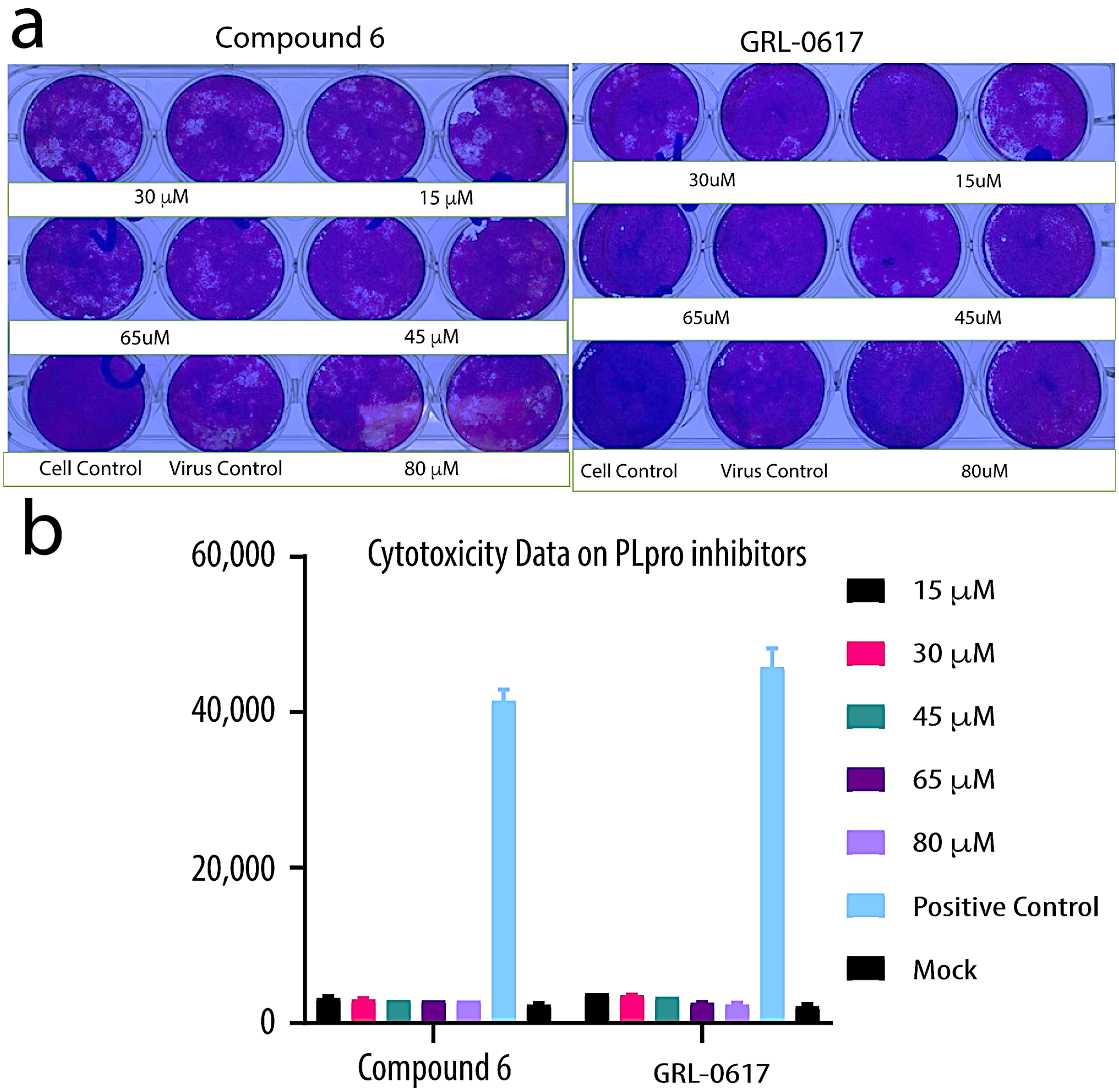


Figure S3.4. Plaque assay plates and cytotoxicity data for Compound 6 and GRL-0617. (a) Corresponding plaque assay plates for antiviral assessment of Compound 6 and GRL-0617. (b) Vero E6 cytotoxicity data for Compound 6 and GRL-0617 performed using Lonza Toxilight bioassay.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Table S4.1 Data collection and refinement statistics

	BtSCoV-Rf1.2004 PLpro/**GRL0617**	BtSCoV-Rf1.2004 PLpro/**37**
Data collection		
Space group	P 2_1 2_1 2	P 6_4 2 2
Cell dimensions		
a, *b*, *c* (Å)	67.4, 67.2, 165.3	176.3, 176.3, 79.8
α, β, γ (°)	90, 90, 90	90, 90, 120
Resolution (Å)	50.0 – 3.15 (3.20 – 3.15)	50.0 – 2.90 (2.95 – 2.90)
$R_{pim}(\%)$	9.1 (52.8)	4.9 (27.3)
R_{merge}(%)	25.0 (131.7)	17.9 (97.9)
$\mathrm{cc}_{1/2}$	0.756 (0.602)	0.982 (0.762)
$I/\sigma I$	11.6 (1.2)	15.5 (1.5)
Completeness (%)	100.0 (100.0)	99.5 (99.3)
Redundancy	4.4 (3.9)	7.4 (6.4)
Refinement		
Resolution (Å)	45.72 – 3.16 (3.28 – 3.16)	38.62 – 2.89 (3.00 – 2.89)
No. reflections	13414	16525
$R_{\mathrm{work}}(\%)/R_{\mathrm{free}}(\%)$	23.32/25.98	20.09/23.97
No. atoms		
Protein	4883	2507
Ligand/ion	50	37
Water	51	25
B-factors		
Protein	72.3	66.8
Ligand/ion	53.2	89.6
Water	51.8	60.8
R.m.s deviations		
Bond lengths (Å)	0.005	0.002
Bond angles (º)	0.86	0.48

*Highest resolution shell is shown in parenthesis.

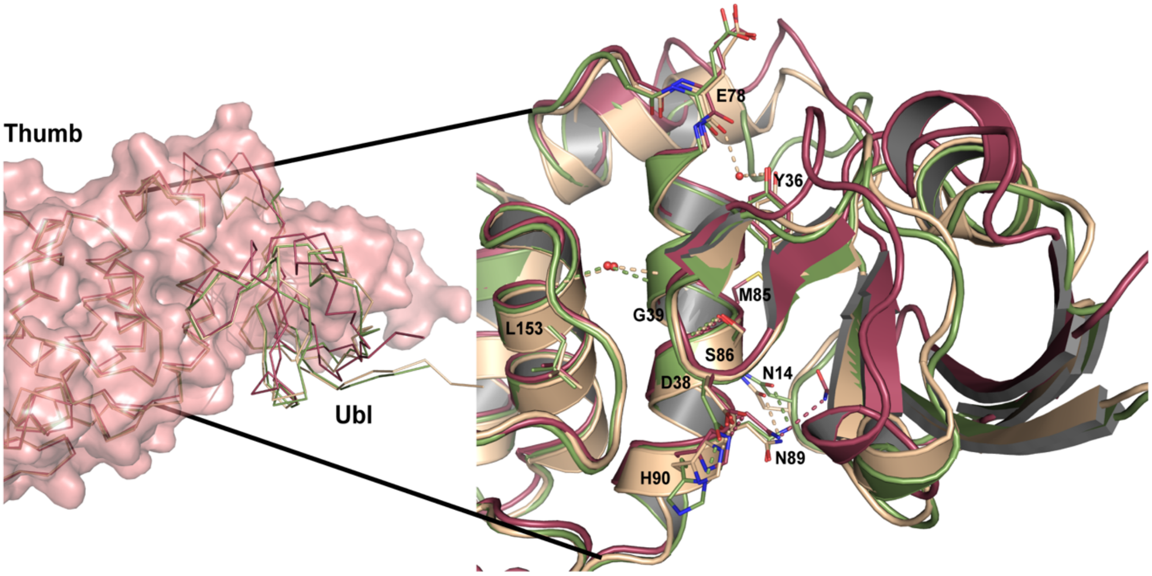


Figure S4.1. Tucked conformation of PLpro Ubl domains. Variation in position of Ubl domains from SARS-CoV-1 PLpro bound to **GRL0617** (pink surface: PDB 3E9S) compared with SARS-CoV-1 PLpro in complex with mouse ISG15 (tan ribbon/cartoon: PDB 5TL7), BtSCoV-Rf1.2004 PLpro bound to **GRL0617** (green ribbon/cartoon), and BtSCoV-Rf1.2004 PLpro bound to **37** (raspberry ribbon/cartoon).

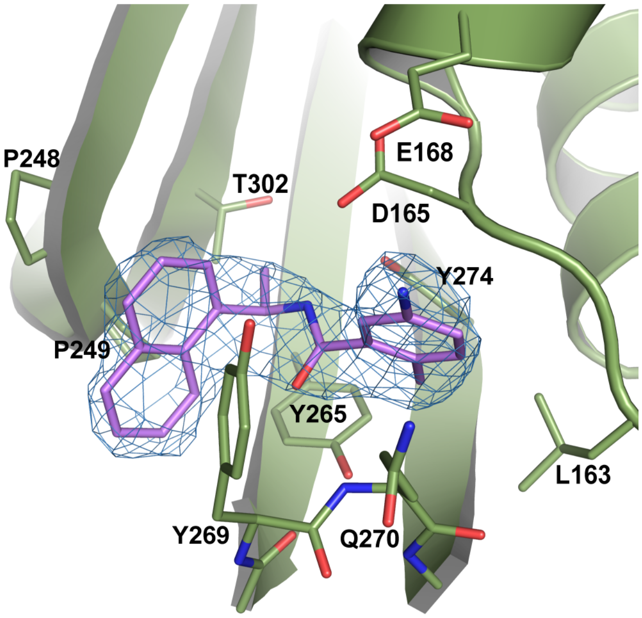


Figure S4.2: **GRL0617** density in the P3/P4 pocket. Inhibitor binding pocket of BtSCoV-Rf1.2004 PLpro (green) in complex with **GRL0617** (purple). A $2F_o$-F_c electron density map is shown contoured at 1σ (blue mesh).

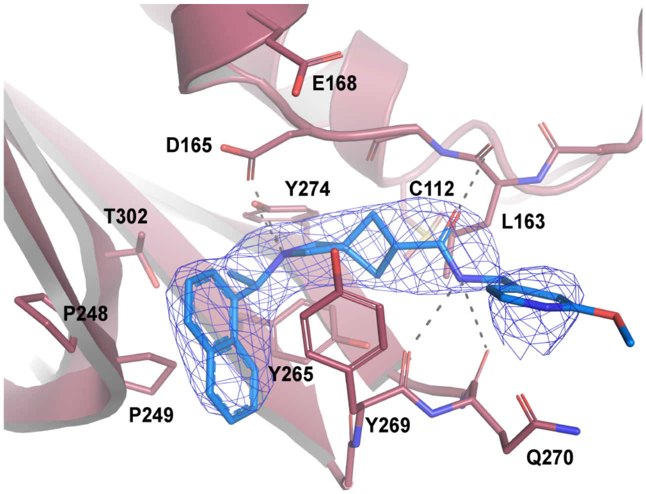


Figure S4.3: **37** density in the P3/P4 pocket. Inhibitor binding pocket of BtSCoV-Rf1.2004 PLpro (raspberry) in complex with **37** (blue). A $2F_o$-F_c electron density map is shown contoured at 1σ (blue mesh).

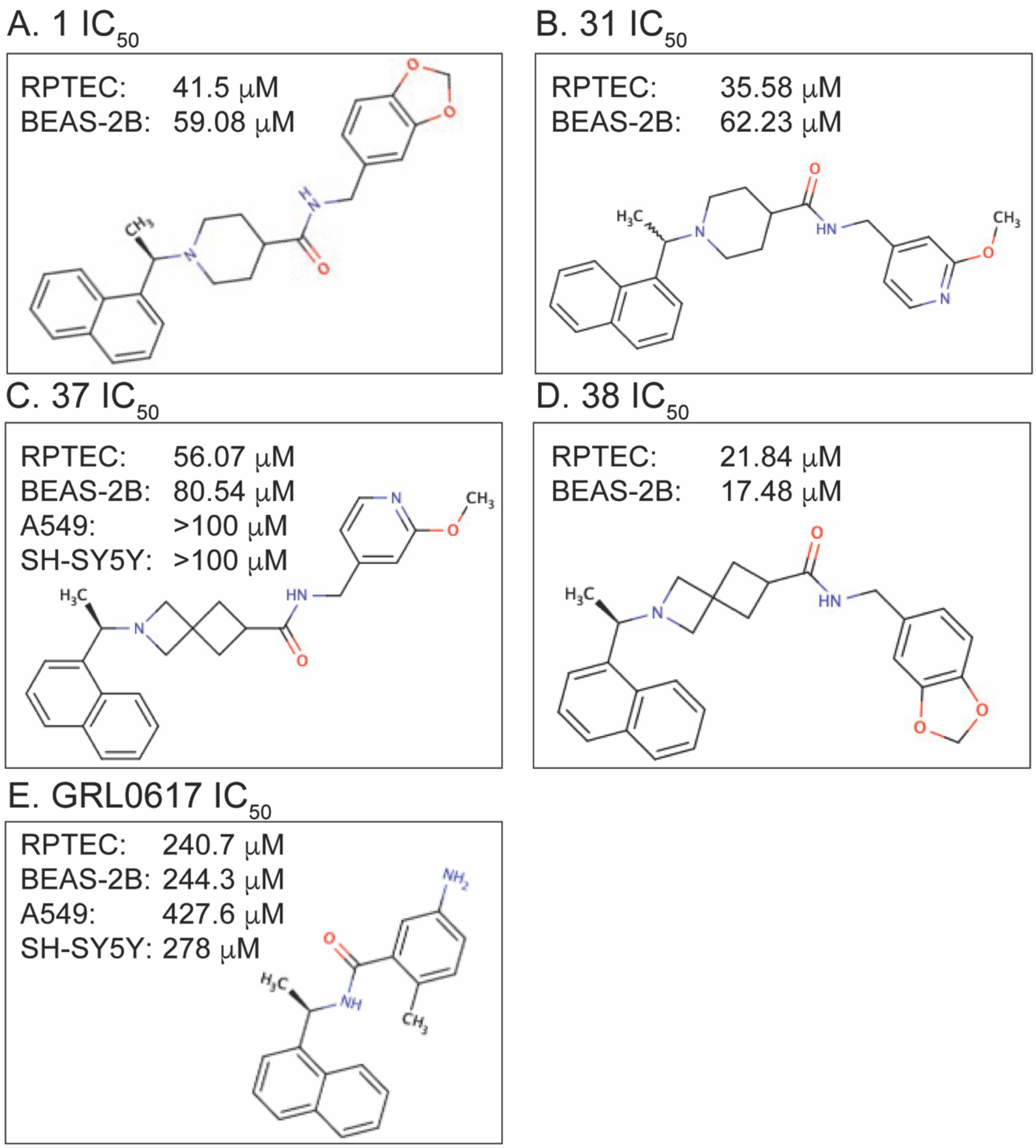


Figure S4.4: Structure-activity relationship between PLpro inhibitors and their off-target toxicity. **A**) **37**, **B**) **38**, **C**) **31**, D) **1**, and **E**) **GRL0617** structures were matched with the respective IC_{50}s in multiple human cell lines.

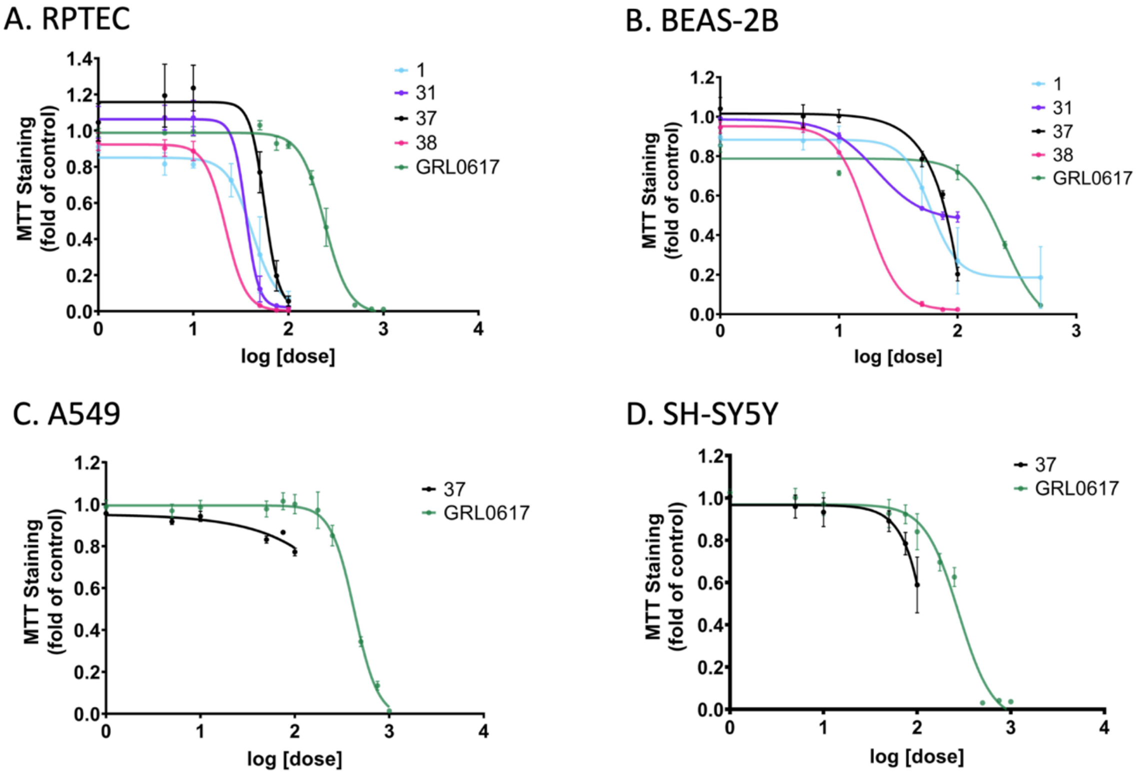


Figure S4.5: Cell viability of multiple human cell lines upon 48 hr exposure to PLpro inhibitors. MTT linear regression curves of **A**) RPTEC, **B**) BEAS-2B, **C**) A549, and **D**) SH-SY5Y cells upon 48 hr exposure to **37**, **38**, **31**, **1**, or **GRL0617**.

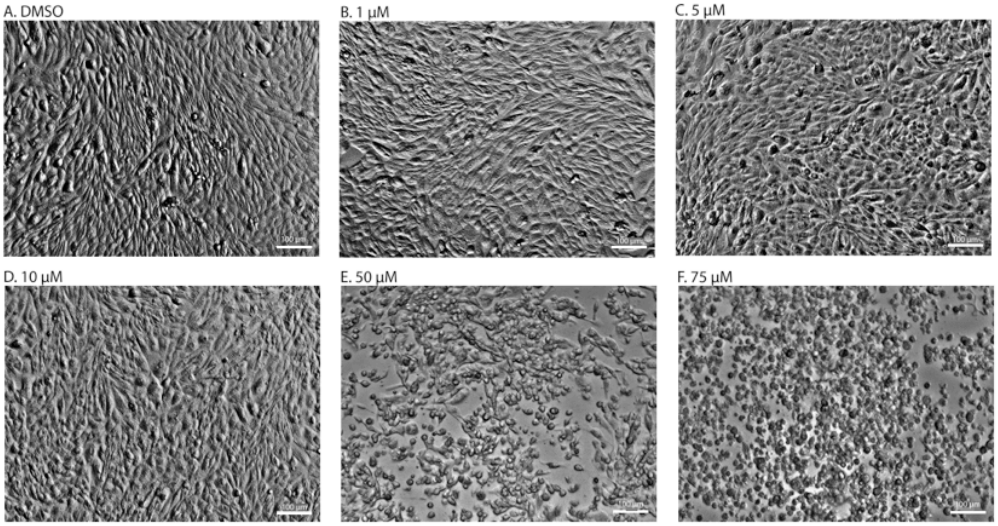


Figure S4.6. Cellular morphology of RPTECs upon 48 hr **37** exposure. RPTECs were exposed to A) DMSO or B) 1 µM, C) 5 µM, D) 10 µM, E) 50 µM, or F) 75 µM **37** for 48 hr. Scale bar: 100 µm.

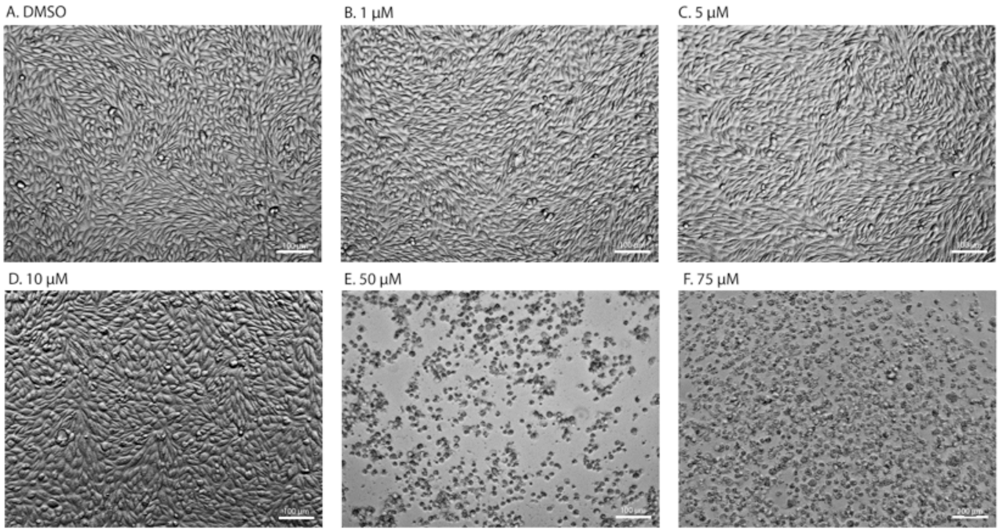


Figure S4.7: Cellular morphology of RPTECs upon 48 hr **38** exposure. RPTECs were exposed to A) DMSO or B) 1 μM, C) 5 μM, D) 10 μM, E) 50 μM, or F) 75 μM **38** for 48 hr. Scale bar: 100 μm.

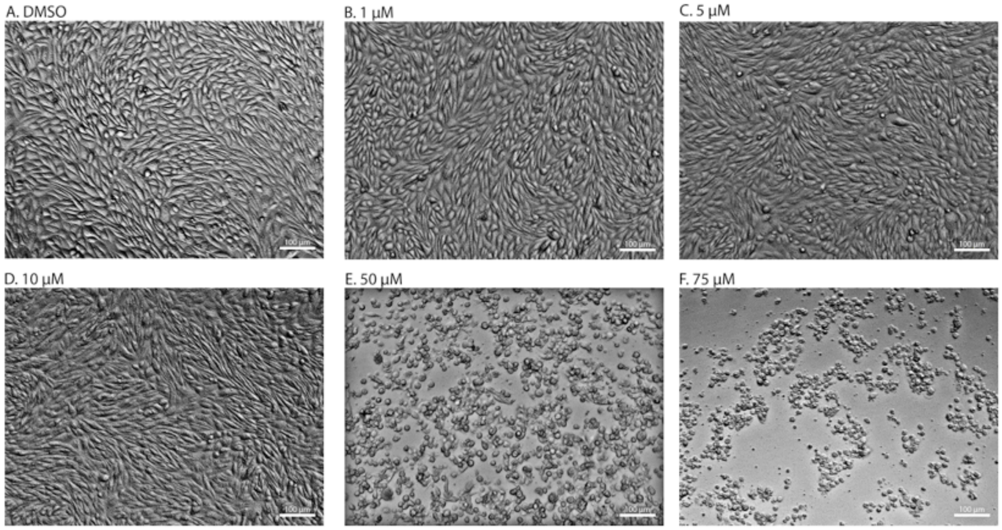


Figure S4.8. Cellular morphology of RPTECs upon 48 hr **31** exposure. RPTECs were exposed to A) DMSO or B) 1 μM, C) 5 μM, D) 10 μM, E) 50 μM, or F) 75 μM **31** for 48 hr. Scale bar: 100 μm.

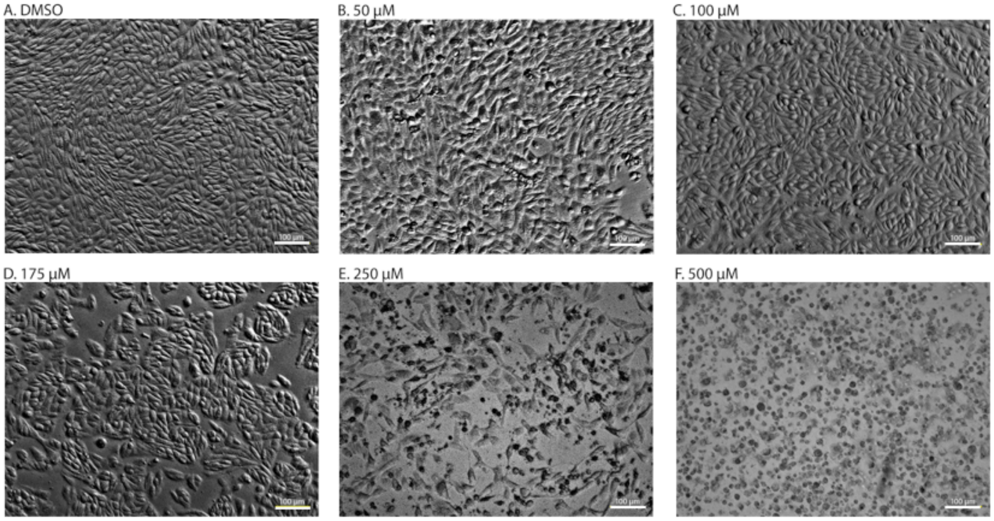


Figure S4.9. Cellular morphology of RPTECs upon 48 hr **GRL0617** exposure. RPTECs were exposed to A) DMSO or B) 50 μM, C) 100 μM, D) 175 μM, E) 250 μM, or F) 500 μM **GRL0617** for 48 hr. Scale bar: 100 μm.

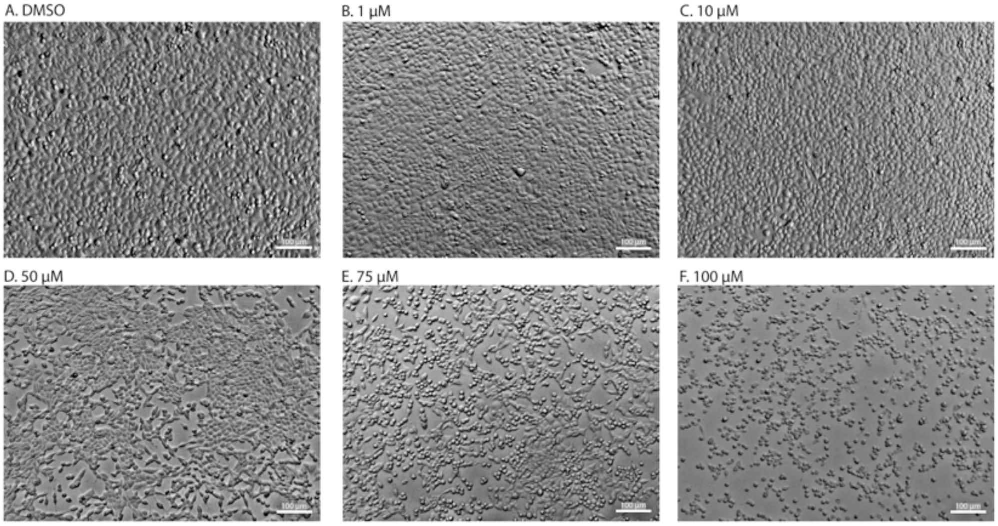


Figure S4.10. Cellular morphology of BEAS-2Bs upon 48 hr **37** exposure. BEAS-2Bs were exposed to A) DMSO or B) 1 μM, C) 10 μM, D) 50 μM, E) 75 μM, or F) 100 μM **37** for 48 hr. Scale bar: 100 μm.

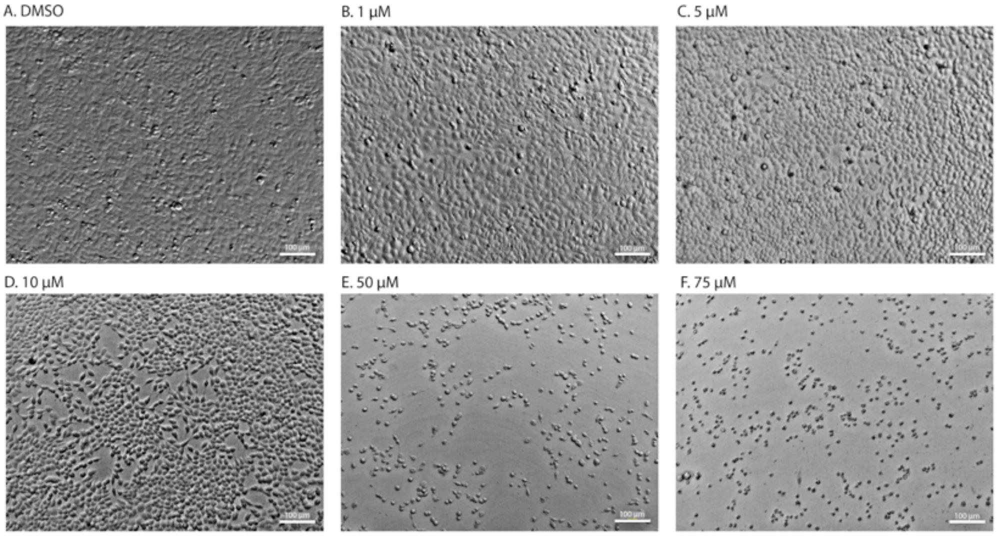


Figure S4.11. Cellular morphology of BEAS-2Bs upon 48 hr **38** exposure. BEAS-2Bs were exposed to A) DMSO or B) 1 μM, C) 5 μM, D) 10 μM, E) 50 μM, or F) 75 μM **38** for 48 hr. Scale bar: 100 μm.

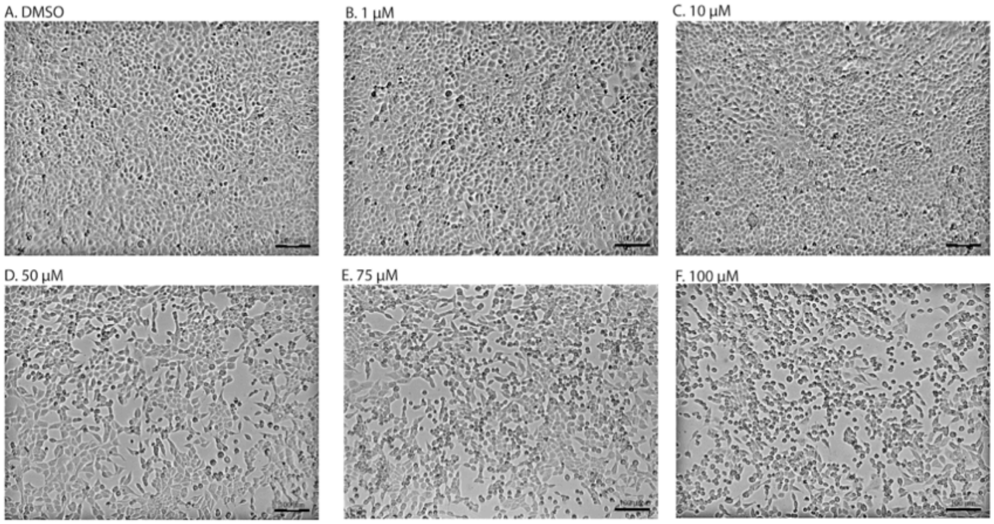


Figure S4.12. Cellular morphology of BEAS-2Bs upon 48 hr **31** exposure. BEAS-2Bs were exposed to A) DMSO or B) 1 μM, C) 10 μM, D) 50 μM, E) 75 μM, or F) 100 μM **31** for 48 hr. Scale bar: 100 μm.

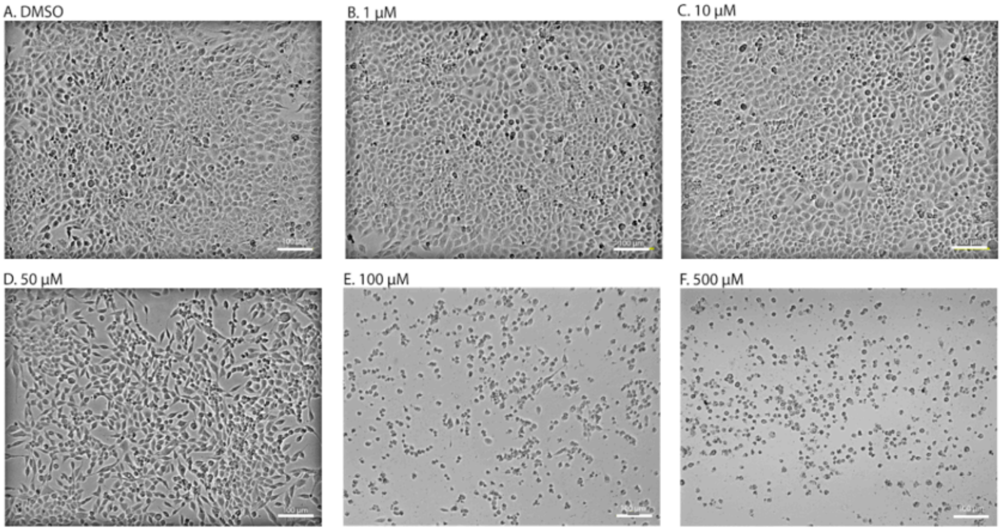


Figure S4.13. Cellular morphology of BEAS-2Bs upon 48 hr **1** exposure. BEAS-2Bs were exposed to A) DMSO or B) 1 µM, C) 10 µM, D) 50 µM, E) 100 µM, or F) 500 µM **1** for 48 hr. Scale bar: 100 µm.

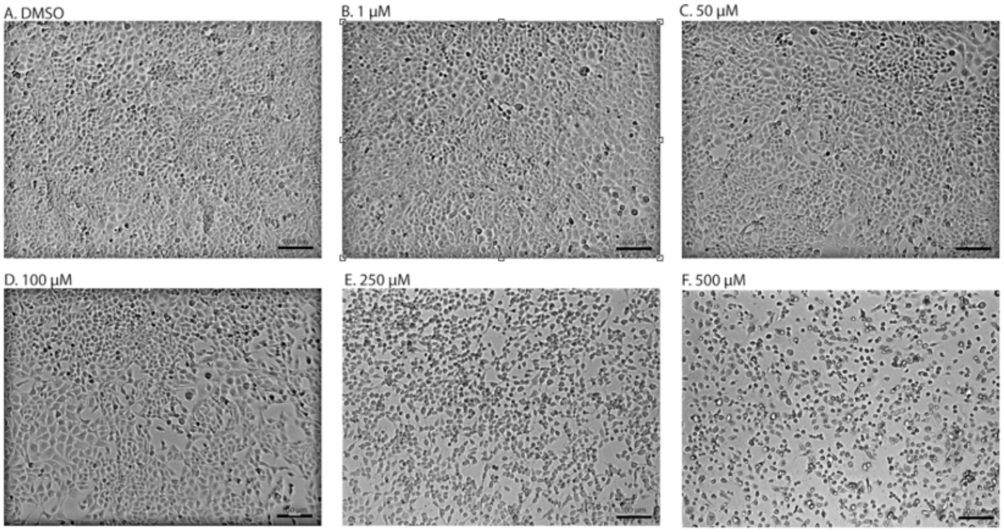


Figure S4.14. Cellular morphology of BEAS-2Bs upon 48 hr **GRL0617** exposure. BEAS-2Bs were exposed to A) DMSO or B) 1 µM, C) 50 µM, D) 100 µM, E) 250 µM, or F) 500 µM **GRL0617** for 48 hr. Scale bar: 100 µm.

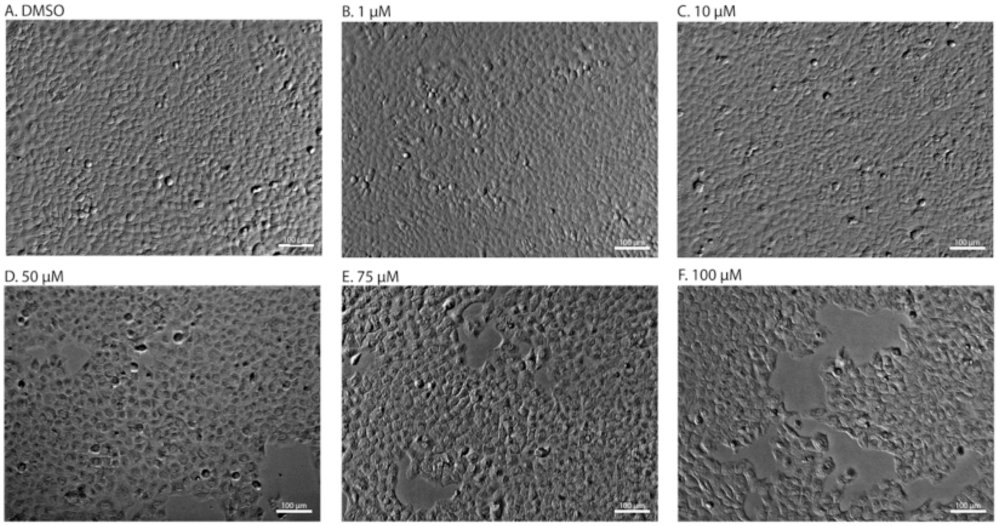


Figure S4.15. Cellular morphology of A549s upon 48 hr **37** exposure. A549s were exposed to A) DMSO or B) 1 μM, C) 10 μM, D) 50 μM, E) 75 μM, or F) 100 μM **37** for 48 hr. Scale bar: 100 μm.

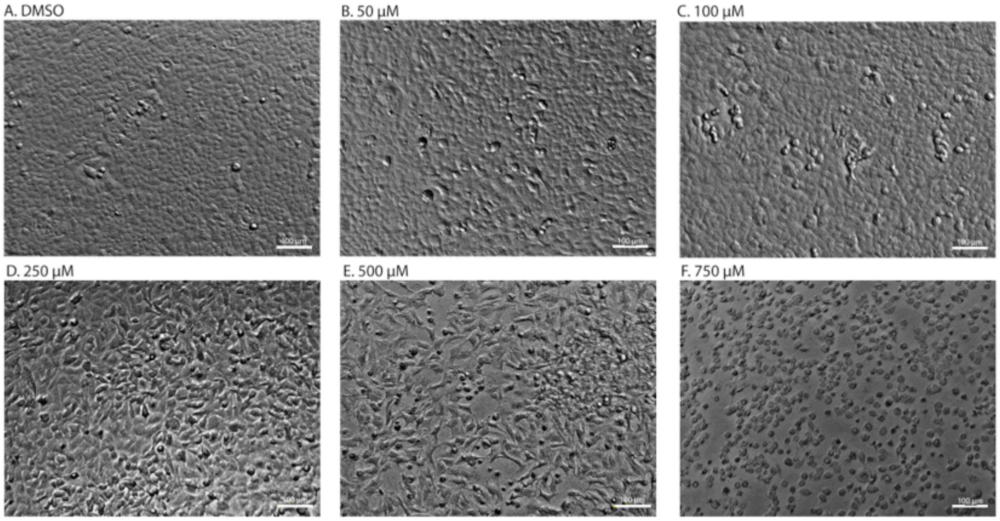


Figure S4.16. Cellular morphology of A549s upon 48 hr **GRL0617** exposure. A549s were exposed to A) DMSO or B) 50 μM, C) 100 μM, D) 250 μM, E) 500 μM, or F) 750 μM **GRL0617** for 48 hr. Scale bar: 100 μm.

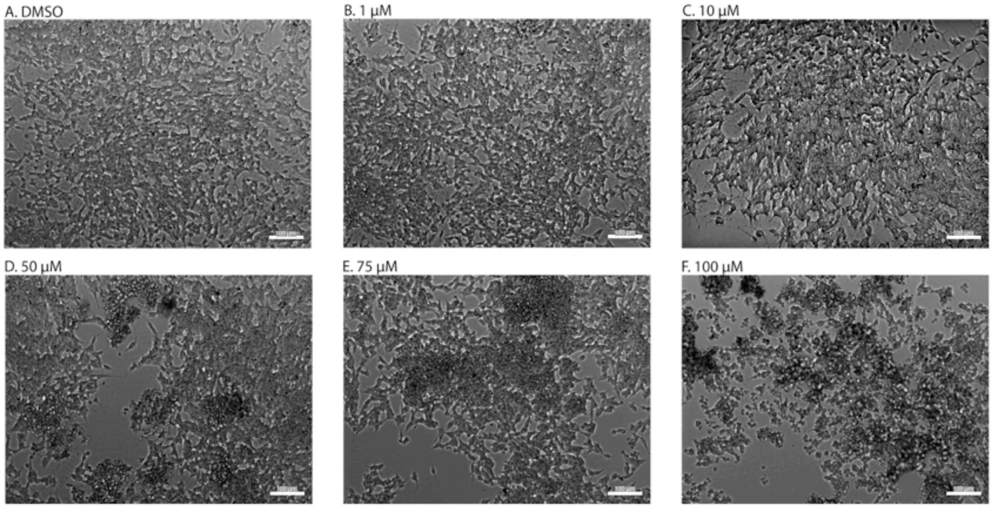


Figure S4.17. Cellular morphology of SH-SY5Ys upon 48 hr **37** exposure. A549s were exposed to A) DMSO or B) 1 μM, C) 10 μM, D) 50 μM, E) 75 μM, or F) 100 μM **37** for 48 hr. Scale bar: 100 μm.

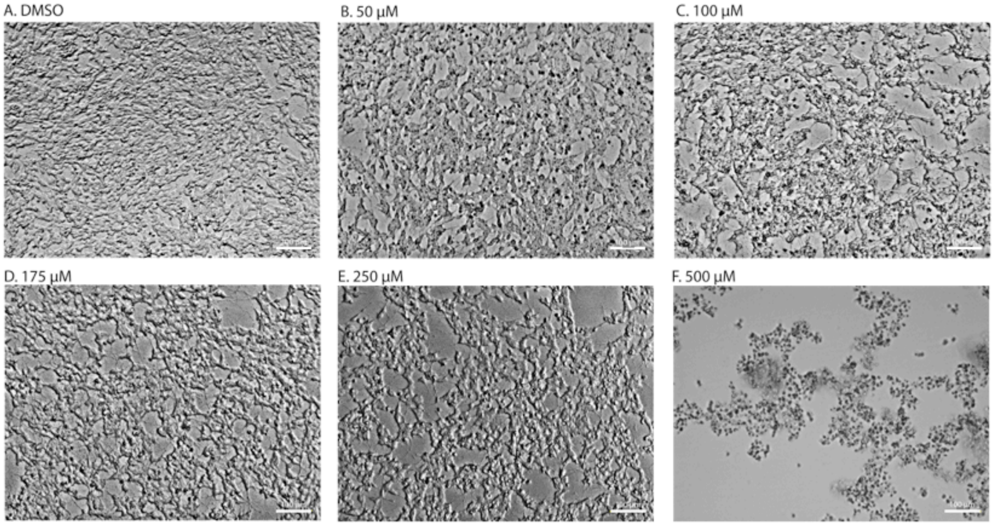
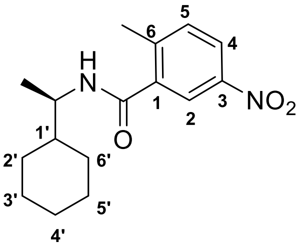


Figure S4.18. Cellular morphology of SH-SY5Ys upon 48 hr **GRL0617** exposure. RPTECs were exposed to A) DMSO or B) 50 μM, C) 100 μM, D) 175 μM, E) 250 μM, or F) 500 μM **GRL0617** for 48 hr. Scale bar: 100 μm.

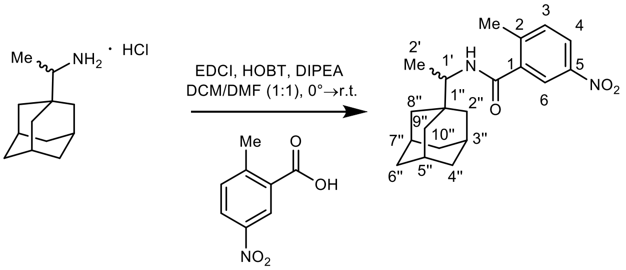
Synthesis of Naphthalene-based compounds.

(*R*)-*N*-[1-(1-Cyclohexyl)ethyl] 2-methyl-5-nitro-benzamide (6).



A stirred solution of 2-methyl-5-nitrobenzoic acid (0.39 g, 2.1 mmol), *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride (EDCI) (0.489 g, 2.5 mmol), and 1-hydroxybenzotriazole hydrate (HOBT) (0.344 g, 2.5 mmol) in DCM (3 mL) was treated with a solution of (*R*)-cyclohexylethylamine **3** (0.25 g, 1.9 mmol) and diisopropylethylamine (1.87 mL, 10.8 mmol) in DCM (2 mL) at 0°C and stirred for 12 hr at room temperature. The reaction mixture was quenched with water (20 mL) and extracted with DCM (2 x 15 mL). The organic layers were dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (EtOAC:Hexane (30:70) to furnish the title benzamide (0.495 g, 87%) as a white solid. $[\alpha]^{20}_{D}$ -21.4 (*c* = 2.2, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.56 – 7.82 (m, 2H, H2,H4), 7.38 (d, *J* = 8.3 Hz, 1H, H5), 5.68 (d, *J* = 9.1 Hz, 1H, NH), 4.35 – 3.84 (m, 1H, C*H*-NH-CO), 2.53 (d, *J* = 1.7 Hz, 3H, CH_3-C6), 1.88 – 1.72 (m, 4H, H2’,H3’,H5’,H6’), 1.72 – 1.52 (m, 1H, H4’), 1.52 – 1.32 (m, *J* = 5.7, 3.0 Hz, 1H, H1), 1.32 – 1.24 (m, 2H, H4, H5), 1.22 (dd, *J* = 6.7, 1.8 Hz, 3H, CH_3), 1.19 – 1.00 (m, 3H, H2, H3, H6); ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.0 (CO-NH), 145.8 (C3), 144.0 (C6), 138.0 (C2), 131.9 (C4), 124.2 (C5), 121.5 (C1), 50.2 (*C*H-NH-CO), 43.0 (C1’), 29.1 (C6), 29.1 (C2), 26.3 (C4), 26.1 (C3,C5), 20.0 (CH_3-C6), 18.0 (CH_3). HRMS (ESI) (*m*/*z*): Calcd for $C_{16}H_{23}N_2O_3Na$ $[M+H]^+$ 291.1697; Found 291.1703.

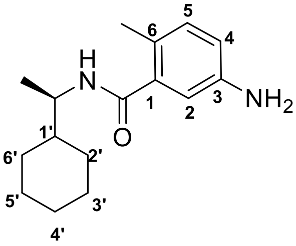
(±)-*N*-[1'-(Adamantan-1'-yl)ethyl]-2-methyl-5-nitrobenzamide (7).



To a stirred solution of 2-methyl-5-nitrobenzoic acid (168 mg, 0.93 mmol), EDCI (194 mg, 1.16 mmol), and HOBT (156 mg, 1.16 mmol) in anhydrous DCM/DMF (1:1) (3 mL) at 0°C was added (±)-1-(1-adamantyl)ethylamine hydrochloride **4** (200 mg, 0.93 mmol) followed by diisopropylethylamine (647 µL, 3.72 mmol). The reaction mixture was allowed to reach room temperature over a period of 10 min and stirred for 12 hr. After such time, the reaction mixture was quenched *via* addition of H_2O (4 mL) and the aqueous layer was re-extracted with DCM (10 mL, 3X). The organic layers were combined, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue obtained was purified by flash column chromatography on silica (eluent: 5:95 – 20:80 EtOAc:hexanes) to afford the title compound (259 mg, 0.76 mmol, 81 %) as a white solid. R_f = 0.10 (2:8 EtOAc:hexanes; CAM). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.18 – 8.11 (m, 2H, H-6, H-4), 7.39 (d, *J* = 8.3 Hz, 1H, H-3), 5.64 (d, *J* = 9.9 Hz, 1H, NH), 3.92 (m, 1H, H-1'), 2.54 (s, 3H, H-8), 2.02 (p, *J* = 3.1 Hz, 3H, H-3'', H-5'', H-7''), 1.73 (dt, *J* = 12.5, 2.7 Hz, 3H, H-4'', H-6'', H-10''), 1.68 – 1.58 (m, 6H, H-4'', H-6'', H-10'', H-2'', H-8'', H-9''), 1.54 (dq, *J* = 11.9, 2.5 Hz, 3H, H-2'', H-8'', H-9''), 1.16 (d, *J* = 6.9 Hz, 3H, H-2'). ^{13}C NMR (126 MHz, $CDCl_3$): δ 167.3 (C=O), 146.1 (C5), 144.2 (C2), 138.4 (C1), 132.1 (C3), 124.5 (C4), 121.7 (C6), 53.9 (C-

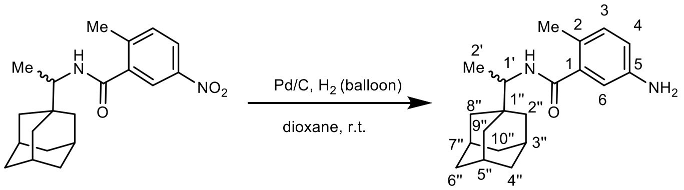
1'), 38.7 (3C) (C 2'',8'',9''), 37.1 (3C) (C4'', C6'', C10''), 36.0 (C1''), 28.4 (3C) (C3'', C5'', C7''), 20.3 (CH_3-aryl), 14.9 (C2'). HRMS (ESI) *m*/*z*: Calcd for $[C_{20}H_{27}O_3N_2]^+$ $[M+H]^+$ 343.2016; Found 343.2011.

(*R*)-*N*-[1-(1-Cyclohexyl)ethyl] 5-amino-2-methylbenzamide (8).



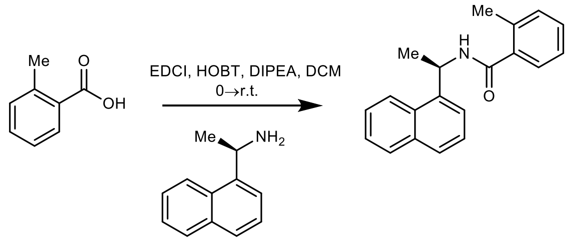
A stirred solution of (*R*)-*N*-[1-(1-cyclohexyl)ethyl] 2-methyl-5-nitro-benzamide **6** (0.440 g, 1.5 mmol) in 50 mL (33.8 M) MeOH:EtOAc (1:1) was hydrogenated over 10% Pd(C) (28 mg) for 12 hr. Upon completion, the reaction mixture was filtered on Celite and the filtrate was concentrated. The residue was purified by chromatography over silica using EtOAc:Hexane (30:70) as the eluent to afford the title amide (0.364 g, 92%) as a white solid. $[\alpha]^{20}_{D}$ -22.1 (*c* = l, $CHCl_3$); ^1H NMR (500 MHz, $CDCl_3$) δ 6.98 (d, *J* = 8.1 Hz, 1H, H5), 6.68 (d, *J* = 2.5 Hz, 1H, H2), 6.64 (dd, *J* = 8.1, 2.6 Hz, 1H, H4), 5.51 (d, *J* = 9.3 Hz, 1H, N*H*), 4.04 (dp, *J* = 9.2, 6.7 Hz, 1H, C*H*-NH-CO), 2.31 (s, 3H, CH_3-C6), 1.88 – 1.70 (m, 4H, H3'b,H5'b,H2'b,H6'b), 1.70 – 1.63 (m, 1H, H4'b), 1.39 (m, 1H, H1), 1.30 – 1.17 (m, 2H, H4,H5) 1.18-1.16 (d, *J* = 4.4 Hz, 3H, CH_3), 1.13 – 0.96 (m, 3H, H2, H3,H6,); ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.5 (CO-NH), 144.0 (C6), 137.8 (C5), 131.7 (C3), 125.1 (C1), 116.4 (C4), 113.4 (C2), 49.6 (*C*H-NH), 43.1 (C1'), 29.1 (C6), 29.0 (C2), 26.4 (C4), 26.1 (C3,C5), 18.7 (CH_3-C5), 18.0 (CH_3). HRMS (ESI) (*m*/*z*): Calcd for $C_{16}H_{25}N_2O$ $[M+H]^+$ 261.1961; Found: 261.1953.

(±)-*N*-(1’-(Adamantan-1’-yl)ethyl)-5-amino-2-methylbenzamide (9).



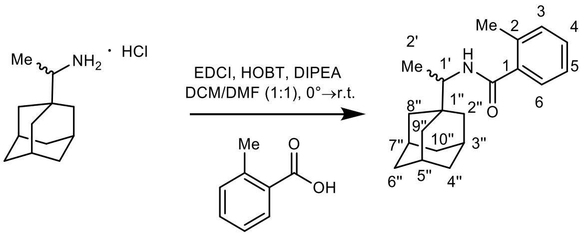
To a stirred solution of (±)-*N*-(1-adamantan-1-yl)ethyl)-2-methyl-5-nitrobenzamide **7** (60 mg, 0.175 mmol) in dioxane (3 mL) was added Pd/C (10% wt) (6 mg, 0.056 mmol). The vessel was purged with H_2 (3X), after which, H_2 was allowed to continuously flow into the stirred vessel for 12 hr. After such time, the reaction mixture was filtered over Celite and washed with MeOH (10 mL). The filtrate was concentrated *in vacuo* and the residue obtained was subjected to flash column chromatography on silica (10:90 MeOH:DCM – 100:0 MeOH:DCM) to afford the title compound (49 mg, 0.157 mmol, 90%) as a grey solid. R_f = 0.10 (1:9 MeOH:DCM; CAM). ^{1}H NMR (500 MHz, CD_3OD): δ 7.92 – 7.84 (m, 1H, NH), 6.96 (d, *J* = 7.9 Hz, 1H, H-6), 6.73 – 6.66 (m, 2H, H-3, H-4), 3.81 (m, 1H, H-1’), 2.25 (s, 3H, H-8), 2.00 (m, 3H, H-3’’, H-5’’, H-7’’), 1.77 (m, 3H, H-4’’, H-6’’, H-10’’), 1.74 – 1.66 (m, 3H, H-4’’, H-6’’, H-10’’), 1.65 (br s, 6H, H-2’’, H-8’’, H-9’’), 1.11 (d, *J* = 7.0 Hz, 3H, H-2’). ^{13}C NMR (126 MHz, CD_3OD): δ 173.3 (C7), 146.4 (C5), 139.2 (C1), 132.2 (C3), 125.2 (C2), 117.8 (C4), 115.0 (C6), 54.8 (C1’), 39.7 (3C) (C2’’, C9’’, C8’’), 38.2 (3C) (C4’’, C6’’, C10’’), 37.5 (C1’’), 29.9 (3C) (C3’’, C5’’, C7’’), 18.8 (C8), 14.4 (C2’). HRMS (ESI) *m*/*z*: Calcd for $[C_{20}H_{29}ON_2]^+$ $[M+H]^+$ 313.2274; Found 313.2268.

(*R*)-2-Methyl-*N*-(1'-(naphthalene-1'-yl)ethyl)benzamide (7724772)[1]



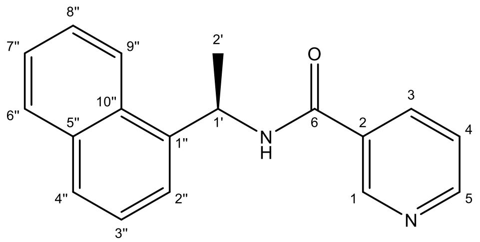
A stirred solution of *o*-toluic acid (49 mg, 0.36 mmol), *N*-(3-dimethylaminopropyl)-*N*'-ethylcarboiimide hydrochloride (EDCI) (87.3 mg, 0.45 mmol), hydroxybenzotriazole (HOBT) (62 mg, 0.45 mmol) in anhydrous DCM (1 mL) at 0°C was treated with *(R)*-1-(1-naphthyl)ethylamine (60 mg, 0.36 mmol) and then gradually brought to r.t. over a period of 10 min and stirred for 16 hr. After such time, the reaction mixture was quenched with H_2O (2 mL) and the aqueous layer re-extracted with EtOAc (5 mL, 3X). The organic layers were combined and dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue obtained was purified by flash column chromatography on silica (eluent: 15:85 EtOAc:hexanes) to afford the title compound (64 mg, 0.22 mmol, 62%) as a white solid, with spectral data in accord with that previously reported in the literature.[1] R_f: 0.40 (1:3 EtOAc:hexanes; CAM). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.24 (d, *J* = 8.5 Hz, 1H), 7.87 (d, *J* = 10 Hz, 1H), 7.81 (d, *J* = 8.2 Hz, 1H), 7.61 – 7.54 (m, 2H), 7.54 – 7.48 (m, 1H), 7.45 (t, *J* = 7.7 Hz, 1H), 7.27 (s, 2H), 7.17 (d, *J* = 7.5 Hz, 1H), 7.11 (t, *J* = 7.5 Hz, 1H), 6.13 (p, *J* = 7.0 Hz, 1H), 5.93 (d, *J* = 8.2 Hz, 1H), 2.43 (s, 3H), 1.79 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$): 169.1, 138.2, 136.6, 136.2, 134.1, 131.4, 131.1, 129.9, 129.0, 128.7, 126.8, 126.7, 126.1, 125.8, 125.3, 123.7, 122.7, 45.0, 20.7, 19.9. HRMS (ESI) *m*/*z*: Calcd for $[C_{20}H_{20}NO]^+$ $[M+H]^+$ 220.1539; Found 220.1528.

(±)-*N*-(1’-(Adamantan-1’-yl)ethyl)-2-methylbenzamide (10).



To a stirred solution of *o*-toluic acid (35 mg, 0.23 mmol), EDCI (45 mg, 0.29 mmol), HOBT (39 mg, 0.29 mmol) in anhydrous DCM/DMF (1:1) (1 mL) at 0°C was added (±)-1-(1-adamantyl)ethylamine hydrochloride **4** (50 mg, 0.23 mmol) followed by diisopropylethylamine (160 μL, 0.92 mmol). The reaction mixture was allowed to reach room temperature over a period of 10 min and stirred for 12 hr. After such time, the reaction mixture was quenched *via* addition of H_2O (3 mL) and the aqueous layer was re-extracted with DCM (5 mL, 3X). The organic layers were combined, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue obtained was purified by flash column chromatography on silica (eluent: 20:80 EtOAc:hexanes) to afford the title compound (60 mg, 0.20 mmol, 88%) as a white solid. $R_f = 0.60$ (2:8 EtOAc:hexanes; CAM). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.35 (dd, *J* = 7.3, 1.5 Hz, 1H, H-6), 7.30 (td, *J* = 7.4, 1.4 Hz, 1H, H-4), 7.24 – 7.17 (m, 2H, H-3, H-5), 5.57 (d, *J* = 9.9 Hz, 1H, NH), 3.93 (dq, *J* = 10.0, 6.9 Hz, 1H, H-1’), 2.46 (s, 3H, H-8), 2.01 (p, *J* = 3.1 Hz, 3H, H-3’’, H-5’’, H-8’’), 1.73 (m, 3H, H-4’’, H-6’’, H-10’’), 1.64 (m, 6H, H-4’’, H-6’’, H-10’’, H-2’’, H-8’’, H-9’’), 1.55 (m, 3H, H-2’’, H-8’’, H-9’’), 1.13 (m, 3H, H-2’). ^{13}C NMR (126 MHz, $CDCl_3$): δ 169.7 (C7), 137.5 (C1), 136.0 (C2), 131.1 (C4), 129.8 (C3), 126.6 (C5), 125.9 (C6), 53.4 (C1’), 38.7 (3C) (C2’’, C8’’, C9’’), 37.2 (3C) (C3’’, C5’’, C7’’), 36.1 (C1’’), 28.5 (3C) (C4’’, C6’’, C10’’), 20.0 (C8), 14.9 (C2’). HRMS (ESI) *m/z*: Calcd for $[C_{20}H_{28}ON]^+$ $[M+H]^+$ 298.2165; Found 298.2157.

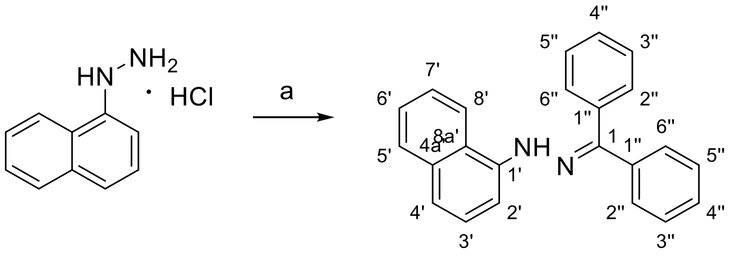
(*R*)-*N*-(1-(Naphthalen-1-yl)ethyl) nicotinamide (11).



A solution of DCC (0.253 g, 1.23 mmol) in CH_2Cl_2 (4.2 mL) was added to a stirred solution of nicotinic acid (0.102 g, 0.83 mmol), DMAP (24 mg, 0.20 mmol), and (*R*)-(+)-1-(1-naphthyl)ethylamine (0.20 mL, 1.25 mmol) in CH_2Cl_2 (4.0 mL) and stirred for 7.5 hr. The reaction mixture was then diluted with ethyl acetate and washed with 1 N HCl. Following concentration of the aqueous layer the resulting yellow oil was dissolved in ethyl acetate, washed with 10% NH_4OH solution and brine, and dried with Na_2SO_4 before concentration under vacuum. The resulting solid was recrystallized from CH_2Cl_2 and hexanes to give the title amide (0.193 mg, 0.70 mmol, 84%) as white crystals. $[\alpha]_D^{23}$ = -19.0 (*c* = 0.4, $CHCl_3$), MP = 155°C, ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.91 – 8.87 (m, 1H, H-1), 8.64 – 8.59 (m, 1H, H-5), 8.12 (d, *J* = 8.4 Hz, 1H, H-9'''), 8.06 (dt, *J* = 8.0, 2.0 Hz, 1H, H-3), 7.85 (dd, *J* = 8.0, 1.6 Hz, 1H, H-6'''), 7.79 (d, *J* = 8.2 Hz, 1H, H-4'''), 7.57 (d, *J* = 7.2 Hz, 1H, H-2'''), 7.55 – 7.46 (m, 2H, H-7''', H-8'''), 7.45 (dd, *J* = 8.2, 7.2 Hz, 1H, H-3'''), 7.30 (dd, *J* = 8.0, 4.8 Hz, 1H, H-4), 6.62 (d, *J* = 7.8 Hz, 1H, NH), 6.10 (p, *J* = 7.0 Hz, 1H, H-1'), 1.77 (d, *J* = 6.8 Hz, 3H, H-2'). ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.6 (C-6), 152.0 (C-5), 147.8 (C-1), 137.8 (C-1'''), 135.5 (C-3), 134.1 (C-5'''), 131.2 (C-10'''), 130.3 (C-2), 129.0 (C-6'''), 128.8 (C-4'''), 126.9 (C-8'''), 126.1 (C-7'''), 125.3 (C-3'''), 123.6 (C-

4), 123.3 (C-9'''), 122.9 (C-2'''), 45.5 (C-1'), 20.7 (C-2'). HRMS (ESI) *m/z*: Calcd for $C_{18}H_{17}N_2O$ $[M+H]^+$ 277.1335; Found 277.1326.

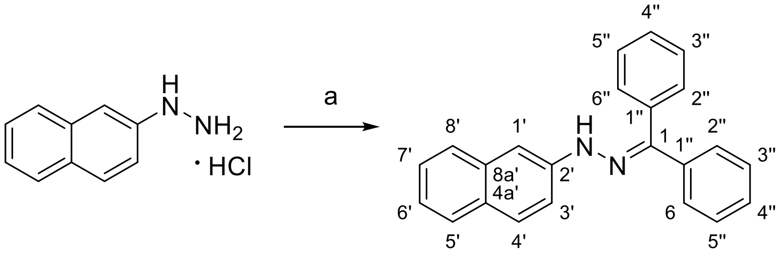
Benzophenone 1-naphthylhydrazone (14).



Reagents and conditions: (a) Benzophenone, H_2SO_4, MeOH, 50°C, 4 hr.

A stirred solution of benzophenone (0.94 g, 5.14 mmol) in MeOH (8 mL) was treated dropwise with concentrated H_2SO_4 (0.16 mL) at 20°C followed by 1-napthylhydrazine hydrochloride **12**[2] (1 g, 5.14 mmol). The reaction mixture was heated to 50°C and stirred for 4 hr then was cooled at 0°C for 20 min, and the precipitate was filtered, and washed with cold MeOH (6 mL) to afford the product as a yellow solid (1.4 g, 83%), which was directly used for the next step without further purification. ^1H NMR (500 MHz, $CDCl_3$) δ 8.31 (s, 1H, H-8′), 7.92 – 7.86 (m, 2H, H-5′, H-4′), 7.83 – 7.79 (m, 2H, Nap), 7.73 – 7.69 (m, 2H, Nap), 7.68 – 7.62 (m, 1H, Nap), 7.58 (t, *J* = 7.9 Hz, 1H, Ar), 7.53 (d, *J* = 1.7 Hz, 1H, Ar), 7.49 – 7.30 (m, 8H, Ar). ^{13}C NMR (126 MHz, $CDCl_3$) δ 146.5 (C-1), 139.5 (C-1′), 138.4, 134.4, 133.1, 130.3, 130.1, 129.7, 129.3, 129.0, 128.5, 128.5, 127.0, 125.8, 125.2, 122.1, 120.2, 119.2, 108.3. HRMS (ESI) *m*/*z*: $[M + H]^+$ Calcd for $C_{23}H_{19}N_2$ 323.1543; Found 323.1535.

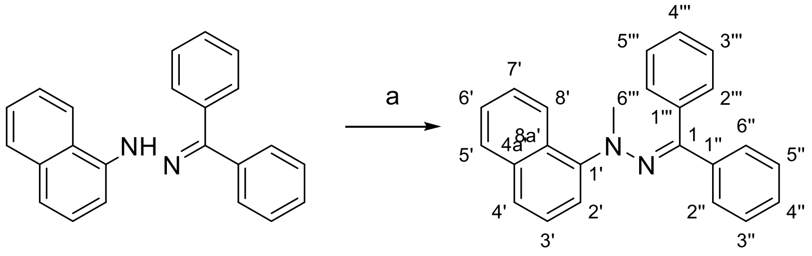
Benzophenone 2-naphthylhydrazone (15).



Reagents and conditions: (a) Benzophenone, H_2SO_4, MeOH, 50°C, 4 hr.

A stirred solution of benzophenone (0.94 g, 5.14 mmol) in MeOH (8 mL) was treated dropwise with concentrated H_2SO_4 (0.16 mL) at 20°C, followed by 2-napthylhydrazine hydrochloride **13** [2] (1 g, 5.14 mmol). The reaction mixture was heated to 50°C and stirred for 6 hr then was cooled on an ice bath for 20 min. The precipitate was filtered, washed with cold MeOH (6 mL) and dried to give the title product as a pale pink solid (1.5 g, 88%), which was directly used for the next step without further purification. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.75 – 7.68 (m, 3H, Nap), 7.68 – 7.63 (m, 3H, Nap), 7.63 – 7.58 (m, 2H, -NH, Nap), 7.57 – 7.51 (m, 1H, Ar), 7.45 (d, J = 2.2 Hz, 1H, Ar), 7.42 – 7.35 (m, 4H, Ar), 7.35 – 7.22 (m, 4H, Ar). ^{13}C NMR (126 MHz, $CDCl_3$) δ 144.9 (C-1), 142.4 (C-2′), 138.4, 134.9, 132.9, 129.8, 129.4, 129.3, 129.2, 129.1, 128.3, 128.2, 127.9, 126.7, 126.6, 126.5, 123.0, 115.7, 107.2. HRMS (ESI) *m/z*: $[M + H]^+$ Calcd for $C_{23}H_{19}N_2$ 323.1543; Found 323.1533.

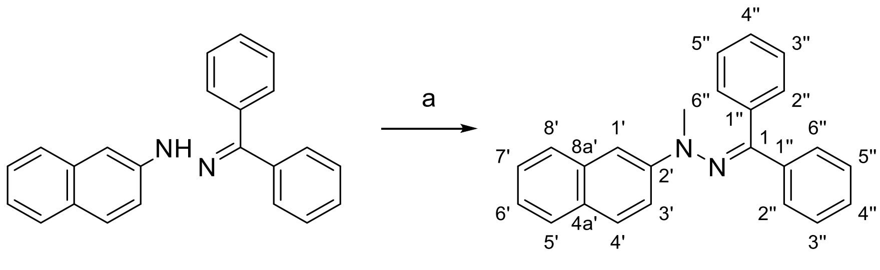
Benzophenone *N*-methyl-(1-naphthyl)hydrazone (16).



Reagents and conditions: (a) NaH, MeI, MeOH, DMF, 20°C, 3.5 hr.

A stirred solution of benzophenone 1-naphthylhydrazone **14** (2 g, 6.20 mmol) in anhydrous DMF (12 mL) was slowly treated with NaH (0.04 g, 9.30 mmol) at 0°C, and then stirred at 0°C for 0.5 h, before methyl iodide (0.58 mL, 9.30 mmol) was added dropwise over 5 min and the resulting mixture was warmed to 20°C and stirred for 3 hr. The reaction mixture was cooled to 0°C before it was quenched with water (20 mL) and extracted with ethyl acetate (3×20 mL), and the combined organic phase washed with 1N HCl (30 mL), saturated aqueous NaHCO$_3$ (30 mL), and brine (30 mL), dried over Na$_2$SO$_4$ and concentrated *in vacuo*. The residue was subjected to flash chromatography (0 to 10% ethyl acetate/hexane) to give the title product (1.8 g, 87%) as a light orange solid. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.99 – 7.91 (m, 1H, H-8′), 7.71 – 7.67 (m, 1H, H-5′), 7.64 – 7.59 (m, 2H, H-2′, H-7′), 7.47 (m, 2H, H-6‴, H-2″), 7.41 – 7.36 (m, 2H, H-4′, H-3′), 7.35 – 7.32 (m, 3H, H-6′, H-2‴, H-6″), 7.31 – 7.22 (m, 1H, H-5‴), 7.09 – 7.04 (m, 2H, H-3″, H-4‴), 7.03 – 6.99 (m, 2H, H-5″, H-3‴), 6.98 – 6.93 (m, 1H H-4″), 3.26 (s, 3H, -N*CH$_3$*). ^{13}C NMR (126 MHz, CDCl$_3$) δ 155.9 (C-1), 150.8 (C-1′), 139.6 (C-8a′), 137.2 (C-1″), 134.4 (C-1‴), 132.5 (C-4a′), 130.2, 129.0, 128.9, 128.8, 128.4, 128.2, 128.0, 127.8, 127.6, 127.4, 125.7, 125.3, 124.9, 123.9, 120.4, 48.7 (CH$_3$). HRMS (ESI) *m*/*z*: [M + Na]$^+$ Calcd for C$_{24}$H$_{20}$N$_2$Na 359.1524; Found 359.1530.

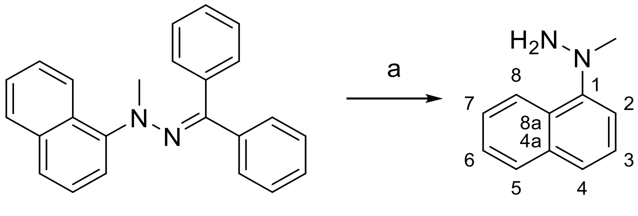
Benzophenone *N*-methyl-(2-naphthyl)hydrazone (17).



Reagents and conditions: (a) NaH, CH_3I, DMF, 20°C, 3.5 hr.

A stirred solution of benzophenone 2-naphthylhydrazone **15** (6.49 g, 20.1 mmol) in 40.3 mL of dry DMF was slowly treated with NaH (1.21 g, 30.2 mmol) at 0°C, and then stirred at this temperature for 0.5 hr. Then methyl iodide (1.88 ml, 30.19 mmol, 1.5 equiv.) was added dropwise over 5 min and the resulting mixture was warmed to 20°C and stirred for 3 hr when TLC showed complete conversion of the starting material. The reaction mixture was cooled to 0°C and quenched with water (100 mL). The mixture was extracted with ethyl acetate (3×100 mL) and the combined organic phase was washed with 1N HCl, saturated aqueous $NaHCO_3$, brine, dried over Na_2SO_4, and concentrated *in vacuo*. The residue was purified by flash column chromatography (0 - 10% ethyl acetate/hexane) to yield the title product (6 g, 89%) as an orange solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 – 7.71 (m, 3H, Nap), 7.71 – 7.61 (m, 3H, Nap), 7.47 – 7.37 (m, 9H, Nap, Ar), 7.30 (m, 2H, Ar), 3.03 (s, 3H, $-CH_3$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.9 (C-1), 148.2 (C-2′), 139.6, 137.4, 134.6, 129.4, 129.3, 128.8, 128.7, 128.6, 128.5, 128.2, 127.7, 126.9, 126.3, 123.1, 117.4, 108.7, 41.0 (CH_3). HRMS (ESI) *m/z*: [M + Na]$^+$ Calcd for $C_{24}H_{20}N_2Na$ 359.1524; Found 359.1536.

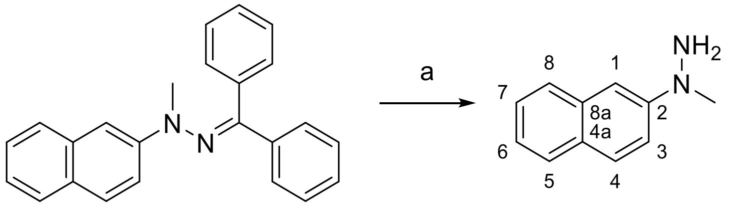
N-Methyl-N-(1-naphthyl)hydrazine (18).



Reagents and conditions: (a) HCl, THF, 20°C, 12 hr.

A solution of benzophenone *N*-methyl-(1-naphthyl)hydrazone **16** (2 g, 5.94 mmol) in THF (60 mL) was treated with 6 M HCl (18.2 mL, 109 mmol) and stirred at room temperature for 10 hr before concentration *in vacuo*. The residue was washed with diethyl ether (2× 20 mL) and the solid was dried under reduced pressure. The solid was dissolved in water (30 mL) with vigorous stirring and the pH was carefully adjusted to 7 with 20% aqueous NaOH. The cloudy mixture was extracted with diethyl ether (2× 20 mL) and the combined organic phase was dried over Na_2SO_4 and concentrated under reduced pressure to give the title product (0.8 g, 81%) as a dark brown oil. ^{1}H NMR (500 MHz, CD_3OD) δ 8.26 (d, J = 8.9 Hz, 1H, H-8), 7.92 (d, J = 9.7 Hz, 1H, H-2), 7.85 (d, J = 8.0 Hz, 1H, H-5), 7.63 – 7.58 (m, 2H, H-3, H-7), 7.55 (m, 2H, H-6, H-4) 3.20 (s, 3H, CH_3). ^{13}C NMR (126 MHz, CD_3OD) δ 145.7 (C-1), 136.1 (C-8a), 129.7 (C-2), 129.5 (C-8), 129.0 (C-3), 128.1 (C-7), 128.0 (C-4a), 126.5 (C-6), 123.4 (C-4), 117.6 (C-5), 46.3 (CH_3). HRMS (ESI) *m*/*z*: $[M + H]^+$ Calcd for $C_{11}H_{13}N_2$ 173.1079; Found 173.1070.

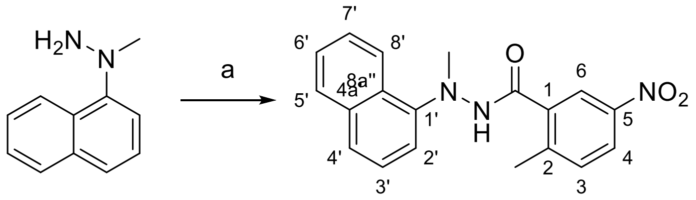
_N_-Methyl-_N_-(2-naphthyl)hydrazine (19).



Reagents and conditions: (a) HCl, THF, 20°C, 12 hr.

To a solution of benzophenone *N*-methyl-(2-naphthyl)hydrazone **17** (1.27 g, 3.77 mmol) in THF (38 mL) was added 6 M HCl (18.2 mL, 109 mmol). The resulting mixture was stirred at 20°C for 12 hr and then concentrated *in vacuo*. The residue was washed with diethyl ether (2×20 mL) and the solid was dried under reduced pressure and then was dissolved in water (20 mL) with vigorous stirring and carefully neutralized with 20% aqueous NaOH, after which the mixture was extracted three times with diethyl ether. The combined organic phase was dried over Na_2SO_4 and concentrated to give the product (0.6 g, 85%) as a dark brown oil. ^{1}H NMR (500 MHz, CD_3OD) δ 7.91 (d, *J* = 9.0 Hz, 1H, H-3), 7.83 (m, 2H, H-1, H-4), 7.53 – 7.47 (m, 2H, H-8, H-5), 7.43 (m, 1H, H-6), 7.39 (dd, *J* = 9.1, 2.5 Hz, 1H, H-7), 3.28 (s, 3H, CH_3). ^{13}C NMR (126 MHz, CD_3OD) δ 145.5 (C-2), 133.7 (C-1), 130.9 (C-3), 129.6 (C-4), 127.4 (C-8a), 127.3 (C-4a), 126.9 (C-8,) 125.5 (C-7), 118.3 (C-5), 113.2 (C-6), 42.8 (CH_3). HRMS (ESI) *m*/*z*: $[M + H]^+$ Calcd for $C_{11}H_{13}N_2$ 173.1079; Found 173.1085.

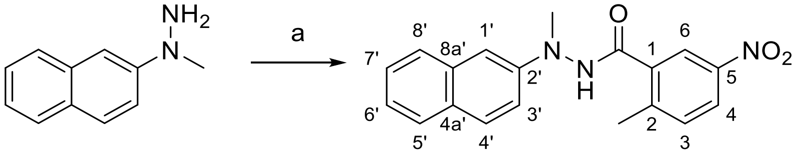
N'-Methyl-N'-(naphthalen-1-yl) 2-methyl-5-nitrobenzhydrazide (20).



Reagents and conditions: (a) 2-Methyl-5-nitrobenzoic acid, EDCI, DIPEA, CH_2Cl_2, 20°C, 3 hr.

To a mixture of *N*-methyl-*N*-(1-naphthyl)hydrazine **18** (70 mg, 0.41 mmol) and anhydrous dichloromethane (2.3 mL), were added 2-methyl-5-nitrobenzoic acid (110 mg, 0.61 mmol), 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (95 mg, 0.61 mmol) and diisopropylethylamine (0.14 mL, 0.81 mmol). The reaction mixture was stirred at 20°C for 3 hr before it was washed with 1N HCl (2×2 mL), saturated aqueous $NaHCO_3$ (2× 2 mL), brine (2 mL), dried over Na_2SO_4, filtered, concentrated, and then purified by flash chromatography (0 to 50% ethyl acetate/hexane) to give the title product (0.1 g, 84%) as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.30 (dd, *J* = 6.3, 3.5 Hz, 1H), 8.01 (d, *J* = 2.5 Hz, 1H), 7.80 (dd, *J* = 6.2, 3.3 Hz, 1H), 7.65 (s, 1H), 7.61 (d, *J* = 8.5 Hz, 1H), 7.48 (dd, *J* = 6.4, 3.3 Hz, 2H), 7.36 (d, *J* = 7.7 Hz, 1H), 7.28 (d, *J* = 7.3 Hz, 1H), 7.26 – 7.22 (m, 1H), 7.18 – 7.12 (m, 1H), 3.28 (s, 3H), 2.43 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 165.6, 145.7, 145.6, 144.8, 135.5, 134.7, 132.0, 131.2, 128.3, 126.5, 126.2, 125.9, 125.2, 124.8, 123.8, 122.0, 114.8, 43.0, 20.0. HRMS (ESI) *m/z*: Calc for $C_{19}H_{17}N_3O_3Na$ $[M+Na]^+$ 358.1168; Found 358.1160.

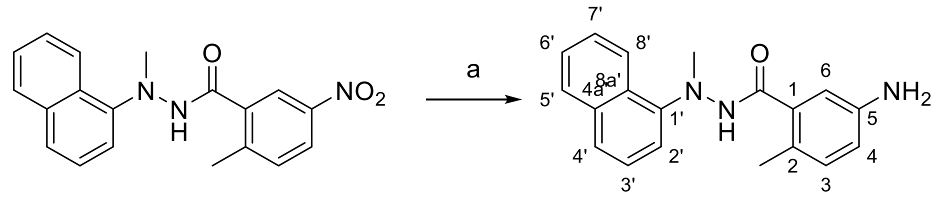
N'-Methyl-N'-(naphthalen-2-yl) 2-methyl-5-nitrobenzhydrazide (21).



Reagents and conditions: (a) 2-Methyl-5-nitrobenzoic acid, EDCI, DIPEA, CH_2Cl_2, 20°C, 4 hr.

To a stirred mixture of *N*-methyl-*N*-(2-naphthyl)hydrazine **19** (100 mg, 0.58 mmol) and dry dichloromethane (3 mL), were added 2-methyl-5-nitrobenzoic acid (158 mg, 0.87 mmol), 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (135 mg, 0.87 mmol) and diisopropylethylamine (0.20 ml, 1.16 mmol). The reaction mixture was stirred at 20°C for 4 hr before it was washed with 1N HCl (2×2 mL), saturated aqueous $NaHCO_3$ (2×2 mL), brine (2 mL), dried over Na_2SO_4, filtered, concentrated, and then subjected to flash column chromatographic purification (0 to 50% ethyl acetate/hexane) to give the title product (0.2 g, 80%) as a white solid. ^1H NMR (500 MHz, $CDCl_3$) δ 8.50 (s, 1H), 8.08 (d, *J* = 2.4 Hz, 1H), 7.93 (dd, *J* = 8.4, 2.5 Hz, 1H), 7.64 (d, *J* = 8.1 Hz, 1H), 7.55 (dd, *J* = 8.6, 6.6 Hz, 2H), 7.36 (ddd, *J* = 8.2, 6.7, 1.3 Hz, 1H), 7.26 (ddd, *J* = 8.0, 6.7, 1.3 Hz, 1H), 7.12 (d, *J* = 8.4 Hz, 1H), 6.96 (dd, *J* = 9.0, 2.5 Hz, 1H), 3.09 (s, 3H), 2.32 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.9 (-CONH), 146.7 (C-2′), 145.6 (C-5), 144.9 (C-1), 134.3 (C-4a′), 132.1, 129.1, 128.5, 127.6, 126.8, 126.6, 124.9, 123.6, 122.0, 115.8, 107.7, 60.5, 40.6 (-NCH_3), 20.0 (CH_3). HRMS (ESI) *m/z*: $[M + Na]^+$ Calcd for $C_{19}H_{17}N_3O_3Na$ 358.1168; Found 358.1156.

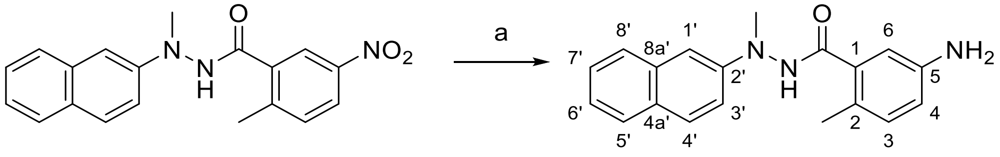
N'-Methyl-N'-(naphthalen-1-yl) 5-amino-2-methylbenzhydrazide (22).



Reagents and conditions: (a) Pd/C, H_2, 1,4-dioxane, 20°C, 5 hr.

A stirred solution of N'-methyl-N'-(naphthalen-1-yl) 2-methyl-5-nitrobenzhydrazide **20** (110 mg, 0.33 mmol) in 1,4-dioxane (3.3 mL) was treated with Pd/C (17.42 mg, 0.02 mmol, 0.05 equiv) and stirred at 20°C under hydrogen (1 atm) for 5 hr. The reaction mixture was filtered, and the filtrate was evaporated to give a residue, which was purified by silica gel column chromatography (20 to 60% ethyl acetate/hexane) to yield the title product (79.1 mg, 79%) as a white solid. ^{1}H NMR (500 MHz, CD_3OD) δ 7.7 (d, J = 9.0 Hz, 1H, H-8′), 7.7 – 7.7 (m, 2H, H-5′, H-4′), 7.4 (m, 1H, H-3′), 7.3 – 7.2 (m, 2H, H-7′, H-2′), 7.2 (d, J = 2.5 Hz, 1H, H- H-6), 7.0 (d, J = 8.2 Hz, 1H, H-6′), 6.9 (d, J = 2.5 Hz, 1H, H-3), 6.7 (dd, J = 8.1, 2.5 Hz, 1H, H-4), 3.3 (s, 3H, CH_3), 2.3 (s, 3H, -N*CH_3*). ^{13}C NMR (126 MHz, CD_3OD) δ 171.4 (-CONH), 147.5 (C-1), 145.4 (C-1′), 134.7, 131.3, 128.7, 128.6, 127.1, 126.5, 126.0, 124.7, 122.9, 117.3, 116.0, 113.9, 107.4, 39.8 (CH_3), 17.5 (-N*CH_3*). HRMS (ESI) *m/z*: $[M + H]^+$ Calcd for $C_{19}H_{20}N_3O$ 306.1601; Found 306.1610.

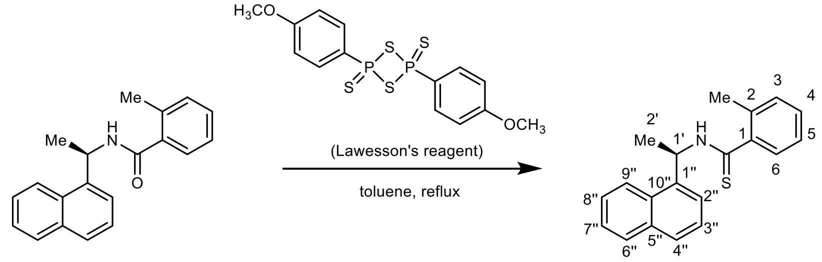
N'-Methyl-N'-(naphthalen-2-yl) 5-amino-2-methylbenzhydrazide (23).



Reagents and conditions: (a) Pd/C, H_2, 1,4-dioxane, 20°C, 5 hr.

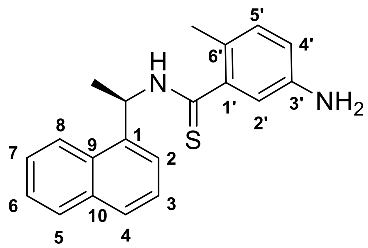
A stirred solution of N'-methyl-N'-(naphthalen-2-yl) 2-methyl-5-nitrobenzhydrazide **21** (50 mg, 0.15 mmol) in 1,4-dioxane (1.5 mL) was treated with Pd/C (7.93 mg, 0.01 mmol) then was stirred at 20°C under hydrogen (1 atm) for 5 hr. Then the mixture was filtered and the filtrate was evaporated. The residue was purified by silica gel column chromatography (20 - 60% ethyl acetate/hexane) to give the title product (38.7 mg, 85%) as a white solid. ^{1}H NMR (500 MHz, CD_3OD) δ 8.37 (d, J = 8.2, 1H, H-3′), 7.83 – 7.80 (m, 1H, H-1′), 7.62 (dt, J = 7.8, 1.1 Hz, 1H), 7.51 – 7.44 (m, 2H), 7.43 – 7.35 (m, 2H), 6.92 (d, J = 8.2 Hz, 1H, H-4′), 6.66 (dd, J = 8.1, 2.5 Hz, 1H, H-4), 6.58 (d, J = 2.5 Hz, 1H, H-3), 3.24 (s, 3H, $-CH_3$), 2.20 (s, 3H, -NCH_3). ^{13}C NMR (126 MHz, CD_3OD) δ 170.2 (-CONH), 146.3 (C-2′), 145.2 (C-5), 134.9 (C-1), 134.8 (C-4a′), 131.1, 128.6, 127.9, 125.8, 125.3, 125.0, 124.8, 124.7, 123.7, 117.1, 114.5, 113.9, 42.1 ($-CH_3$), 17.2 (-NCH_3). HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{19}H_{20}N_3O$ 306.1601; Found 306.1615.

(*R*)-2-Methyl-*N*-(1’-(naphthalene-1’-yl)ethyl)benzothioamide (24).



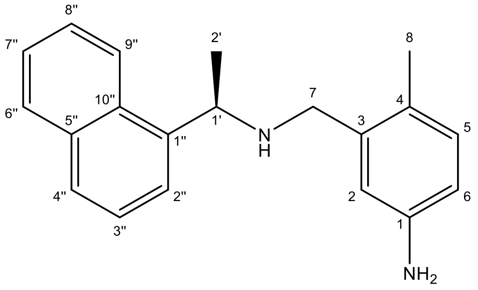
To a stirred solution of (*R*)-2-methyl-*N*-(1-(naphthalene-1-yl)ethyl)benzamide (**7724772**) (29 mg, 0.1 mmol) in anhydrous toluene (0.5 mL) was added Lawesson’s reagent (49 mg, 0.12 mmol) under an argon atmosphere. The solution was gradually brought to reflux and stirred for 2 hr. After such time, the solvent was removed *in vacuo* and the residue obtained was subjected to flash column chromatography on silica (eluent: 2:6 EtOAc:hexanes) affording the title compound (12.2 mg, 0.35 mmol) as a white solid. R_f = 0.50 (2:6 EtOAc:hexanes; CAM). $[\alpha]_D^{21}$ = -44.0 ° (c = 1.0, $CHCl_3$). ^1H NMR (500 MHz, $CDCl_3$): δ 8.21 (d, *J* = 8.5 Hz, 1H, H-9’’), 7.88 (dd, *J* = 8.1, 1.4 Hz, 1H, H-6’’), 7.84 (d, *J* = 8.2 Hz, 1H, H-4’’), 7.65 – 7.57 (m, 2H, H-8’’, H-2’’), 7.54 (dd, *J* = 8.1, 6.8 Hz, 1H, H-7’’), 7.46 (dd, *J* = 8.2, 7.2 Hz, 1H, H-3’’), 7.38 – 7.32 (m, 1H, NH), 7.20 – 7.13 (m, 2H, H-6, H-4), 7.13 – 7.06 (m, 2H, H-3, H-5), 6.66 – 6.57 (m, 1H, H-1’), 2.31 (s, 3H, CH_3-H-8), 1.89 (d, *J* = 6.7 Hz, 3H, CH_3-H-2’). ^{13}C NMR (126 MHz, $CDCl_3$): 199.9 (C=S), 143.9 (C-1), 136.4 (C-1’’), 134.1 (C10’’), 133.2 (C2), 131.7 (C5’’), 130.9 (C3), 129.3 (C4’’), 129.01 (C6’’), 128.98 (C6), 127.1 (C8’’), 126.4 (2C) (C7’’, C4), 126.0 (C5), 125.3 (C3’’), 123.8 (C9’’), 123.4 (C2’’), 50.7 (C-1’), 19.5 (CH_3-aryl), 17.8 (CH_3 – C2’). HRMS (ESI) *m*/*z*: Calcd for $[C_{20}H_{20}NS]^+$ $[M+H]^+$ 306.1311; Found 306.1309.

(*R*)-*N*-[1-(1-Naphthyl)ethyl] 5-amino-2-methylthiobenzamide (25)



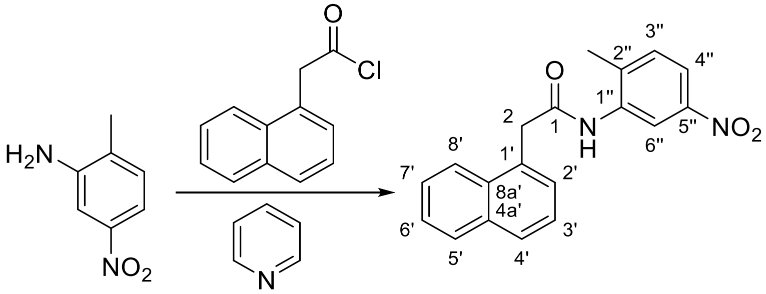
A mixture of (*R*)-*N*-[1-(1-naphthyl)ethyl] 5-amino-2-methylbenzamide (**GRL0617**) (0.035 g, 0.11 mmol) and Lawesson's reagent (0.046 g, 0.11 mmol) were dissolved in 2 mL toluene and heated to 80°C with stirring for 2 hr. The solvents were evaporated under reduced pressure and the residue dissolved in EtOAc (5 mL) and washed with water (10 x 2 mL). The organic layer was separated, dried over Na_2SO_4, filtered, and concentrated. The residue was subjected to column chromatography over silica gel eluting with EtOAc:Hexane (30:70) to afford the title thiobenzamide (0.027 g, 75 %) as a brown solid. $[\alpha]^{20}_{D}$ +66.8 (*c* = l, $CHCl_3$); ^{1}H NMR (500 MHz, MeOH-d_4) δ 8.33 (d, *J* = 8.5 Hz, 1H, 8H), 7.94 (d, *J* = 8.4 Hz, 1H, 5H), 7.87 (d, *J* = 8.2 Hz, 1H, 4H), 7.67 (d, *J* = 7.2 Hz, 1H, 2H), 7.64 – 7.59 (m, 7H), 7.59 – 7.48 (m, 2H, 3H, 6H), 6.91 (d, *J* = 8.1 Hz, 1H, H5'), 6.68 – 6.57 (m, 3H, C*H*-NH, H2', H4'), 2.15 (s, 3H, C*H*$_3$-C6'), 1.81 (d, *J* = 6.9 Hz, 3H, C*H*$_3$); ^{13}C NMR (126 MHz, MeOH-d_4) δ 201.4 (*C*S-NH), 146.3 (C3'), 146.0 (C6'), 138.7 (C1), 135.4 (C1'), 133.0 (C9), 132.0 (C10), 129.9 (C5'), 129.4 (C5), 127.4 (C4), 126.9 (C7), 126.4 (C6), 124.9 (C3), 124.6 (C8), 123.4 (C2), 117.1 (C2'), 115.0 (C4'), 51.9 (*C*H-NH), 19.1 (C*H*$_3$-C6), 18.5 (*C*H$_3$); HRMS (ESI) (*m*/*z*): Calcd for $C_{20}H_{21}N_2S$ [M+H]+, 321.1420; Found 321.1406.

(*R*)-4-Methyl-3-(((1-(naphthalen-1-yl)ethyl)amino)methyl)aniline (26).



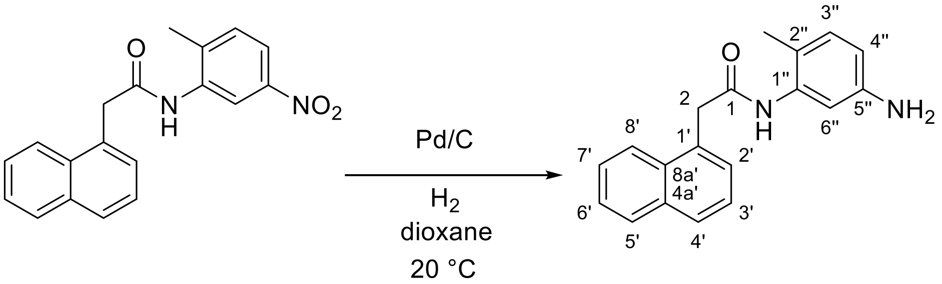
TMSCl (0.04 mL, 0.32 mmol) was added to a stirred solution of (*R*)-5-amino-2-methyl-*N*-(1-(naphthalen-1-yl)ethyl)benzamide (**GRL0617**) (41.6 mg, 0.14 mmol) in THF (1.5 mL) at 0°C. After 20 min $LiAlH_4$ (24.9 mg, 0.66 mmol) was added, and the reaction mixture was heated to reflux for 16 hr. The reaction mixture was then diluted with ethyl acetate, washed with 1N NaOH and brine, dried over Na_2SO_4, filtered, and concentrated. The crude product was adsorbed on Celite® and purified over silica gel eluting with 0% to 10% methanol in CH_2Cl_2 to give the title amine (22.6 mg, 0.08 mmol, 57%) as a yellow film. $[\alpha]_D^{23}$ = -12.28 (*c* = 0.9, $CHCl_3$), ^1H NMR (500 MHz, $CDCl_3$) δ 8.18 (dd, *J* = 7.7, 2.0 Hz, 1H, H-9’’), 7.89 (dd, *J* = 7.0, 2.7 Hz, 1H, H-6"), 7.80 – 7.74 (m, 2H, H-2", H-4"), 7.55 – 7.45 (m, 3H, H-3", H-7", H-8"), 6.93 (d, *J* = 7.9 Hz, 1H, H-5), 6.69 (d, *J* = 2.5 Hz, 1H, H-2), 6.51 (dd, *J* = 7.9, 2.6 Hz, 1H, H-6), 4.72 (q, *J* = 6.6 Hz, 1H, H-1'), 3.66 (d, *J* = 13.2 Hz, 1H, H-7"), 3.62 (d, *J* = 13.2 Hz, 1H, H-7'), 2.15 (s, 3H, H-8), 1.54 (d, *J* = 6.6 Hz, 3H, H-2'). ^{13}C NMR (126 MHz, $CDCl_3$) δ 144.5 (C-1), 141.2 (C-1"), 139.5 (C-3), 134.1 (C-5"), 131.5 (C-10"), 131.1 (C-5), 129.1 (C-6"), 127.3 (C-4"), 126.2 (C-4), 125.9 (C-3"), 125.8 (C-8"), 125.4 (C-7"), 123.2 (C-9"), 123.0 (C-2"), 115.9 (C-2), 113.8 (C-6), 53.5 (C-1'), 49.8 (C-7), 23.8 (C-2'), 18.1 (C-8). HRMS (ESI) *m/z*: Calcd for $C_{20}H_{23}N_2$ $[M+H]^+$ 291.1856; Found 291.1854.

N-(2-Methyl-5-nitrophenyl)-2-(naphthalen-1-yl)acetamide (27).



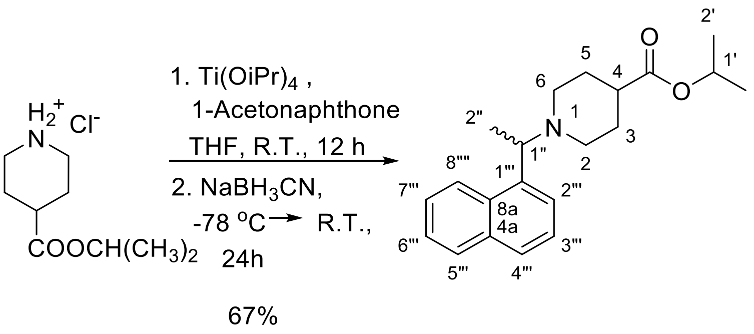
2-Methyl-5-nitroaniline (3.0 g, 19.7 mmol) was dissolved in pyridine (4 mL) and anhydrous CH_2Cl_2 (4 mL) and reaction mixture was cooled down to 0°C before 2-(naphthalen-1-yl)acetyl chloride[3] (3.6 mL, 21.7 mmol) was added dropwise. The reaction mixture then was allowed to warm up to 20°C and was stirred for 20 min before MeOH (10 mL) was added and the precipitate was filtered off. The precipitate was washed with MeOH (30 mL) and suspended hot EtOH (30 mL) giving a slightly yellow suspension. The white precipitate was filtered off, washed with EtOH (3×30 mL), and dried *in vacuo* to give the title compound as a white solid (4.5 g, 70%). R_f 0.38 (hexanes : EtOAc 7 : 3 (v/v); UV). ^{1}H NMR (500 MHz, DMSO-*d*6) δ 9.89 (s, 1H, CONH), 8.46 (d, J = 2.5 Hz, 1H, H-6’’), 8.17 (d, J = 8.3 Hz, 1H, H-8’’), 8.02 – 7.77 (m, 3H, H-2’, H-4’’, H-7’), 7.61 – 7.25 (m, 5H, H-3’’, H-3’, H-4’, H-5’, H-6’), 4.28 (s, 2H, -C<u>H</u>$_2$CO), 2.35 (s, 3H, -CH_3). ^{13}C NMR (126 MHz, DMSO-*d*6) δ 169.8 (C-1), 145.7 (C-5’’), 138.9 (C-1’’), 137.2 (C-4a’), 133.4 (C-1’), 132.2 (C-8a’), 132.0 (C-3’’), 131.4 (C-2’’), 128.5 (C-4’), 127.3 (C-5’), 126.1 (C-2’), 125.7 (C-7’), 125.6 (C-6’), 124.1 (C-8’), 119.3 (C-6’’), 118.3 (C-4’’), 40.1 (C-2), 18.1 (CH_3). HRMS (ESI) *m*/*z*: Calcd for $C_{19}H_{16}N_2NaO_3$ $[M+Na]^+$ 343.1059; Found 343.1044.

N-(5-Amino-2-methylphenyl)-2-(naphthalen-1-yl)acetamide (28).



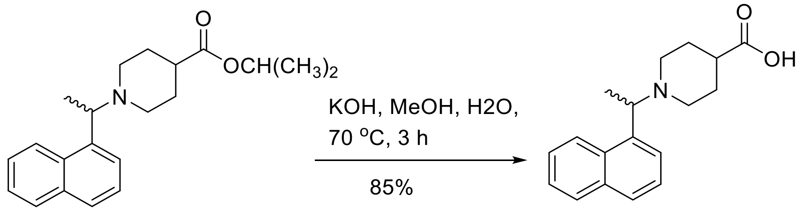
N-(2-Methyl-5-nitrophenyl)-2-(naphthalen-1-yl)acetamide **27** (100 mg, 0.31 mmol) was dissolved in 1,4-dioxane (6 mL) and Pd/C (10% wt) (19.9 mg, 0.037 mmol) was added. The reaction mixture then was stirred for 4h under H_2 (balloon) at 20°C. After completion of the reaction, detected by LCMS, the reaction mixture was filtered on Celite and the filter cake was washed with Et_2O (3×20 mL). The filtrate was concentrated to dryness and the crude product was purified by flash column chromatography over silica gel (eluent: hexanes : acetone (acetone 0 → 30 %)) to give the title compound as a white solid (75 mg, 82 %). R_f 0 (hexanes : EtOAc 7 : 3 (v/v); UV). ^{1}H NMR (500 MHz, DMSO-*d6*) δ 9.31 (s, 1H, CONH), 8.20 (d, *J* = 8.2 Hz, 1H, H-8’), 7.94 (d, *J* = 7.5 Hz, 1H, H-2’), 7.84 (d, *J* = 8.0 Hz, 1H, H-4’), 7.60 – 7.45 (m, 4H, H-7’, H-6’, H-5’, H-3’), 6.80 (d, *J* = 8.0 Hz, 1H, H-3’’), 6.68 (d, *J* = 2.5 Hz, 1H, H-6’’), 6.30 (dd, *J* = 8.0, 2.5 Hz, 1H, H-4’’), 4.83 (s, 2H, NH_2), 4.13 (s, 1H, CH_2CO), 1.99 (s, 3H, CH_3). ^{13}C NMR (126 MHz, DMSO-*d6*) δ 168.7 (C-1), 146.7 (C-5’’), 136.5 (C-1’’), 133.4 (C-4a’), 132.9 (C-1’), 132.0 (C-8a’), 130.3 (C-3’’), 128.4 (C-4’), 127.8 (C-2’’), 127.1 (C-5’), 126.0 (C-2’), 125.7 (C-7’), 125.5 (C-6’), 124.3 (C-3’), 118.3 (C-8’), 111.3 (C-4’’), 110.9 (C-6’’), 40.3 (C-2), 16.8 (CH_3). HRMS (ESI) *m*/*z*: Calcd for $C_{19}H_{19}N_2O$ $[M+H]^+$ 313.1317; Found 313.1300.

(±)-Isopropyl 1-(1-(naphthalen-1-yl)ethyl)piperidine-4-carboxylate (29).



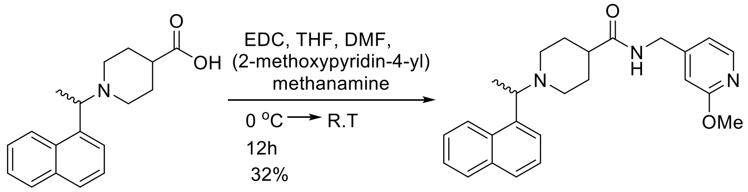
Isopropyl piperidine-4-carboxylate hydrochloride[4] (2.07 g, 10 mmol) was added into the solution of acetonaphthone (1.70 g, 10 mmol) in anhydrous tetrahydrofuran (THF) (300 mL) under argon atmosphere. Then titanium (IV) isopropoxide (3.55 g, 12.5 mmol) was added into the reaction mixture. The reaction mixture was stirred for 12 hr at room temperature, cooled down to -78°C, and sodium cyanoborohydride (1.35 g, 20 mmol) was added. The temperature was slowly raised to room temperature, and the reaction mixture was stirred for another 24 hr. The solvent was removed under reduced pressure to obtain a yellowish viscous crude product. The desired product was obtained as a colorless liquid (2.18 g, 67%) after the silica gel column purification using 100% hexane to 15% EtOAc in hexane as eluent. ^{1}H NMR (500 MHz, $CDCl_3$) ^{1}H NMR δ 8.44 (d, J = 7.2 Hz, 1H, $H_{8'''}$), 7.87 – 7.81 (m, 1H, $H_{5'''}$), 7.73 (d, J = 8.1 Hz, 1H, $H_{4'''}$), 7.57 (d, J = 7.2 Hz, 1H, $H_{2'''}$), 7.50 – 7.39 (m, 3H, $H_{3'''}$, $H_{6'''}$, $H_{7'''}$), 4.99 (hept, J = 6.3 Hz, 1H, $H_{1'}$), 4.09 (q, J = 6.8 Hz, 1H, $H_{1''}$), 3.17 – 3.09 (m, 1H, H_2), 2.82 (m, 1H, H_6), 2.23 (tt, J = 11.2, 4.1 Hz, 1H, H_4), 2.13 – 2.01 (m, 2H, H_2, H_6), 1.95 – 1.86 (m, 1H, H_5), 1.81 – 1.64 (m, 3H, H_3, H_5), 1.46 (d, J = 6.7 Hz, 3H, $H_{2''}$), 1.20 (d, J = 6.4 Hz, 6H, $H_{2'}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 175.1 (C=O), 141.0 ($C_{1'''}$), 134.2 ($C_{4a'''}$), 131.8 ($C_{8a'''}$), 128.8 ($C_{5'''}$), 127.4 ($C_{4'''}$), 125.5 ($C_{3'''}$), 125.5 ($C_{2'''}$), 125.4 ($C_{6'''}$), 124.6 ($C_{7'''}$), 124.4 ($C_{8'''}$), 67.4 ($C_{1'}$), 61.7 ($C_{1''}$), 51.6 (C_6), 49.3 (C_2), 41.7 (C_4), 28.7 (2C, C_3, C_5), 21.9 (2C, $C_{2'}$), 18.8 ($C_{2''}$). HRMS (ESI) *m/z*: Calcd for $C_{21}H_{28}NO_2$ $[M+H]^+$ 326.2115; Found 326.2111.

(±)-1-(1-(Naphthalen-1-yl)ethyl)piperidine-4-carboxylic acid (30).



Isopropyl 1-(1-(naphthalen-1-yl)ethyl)piperidine-4-carboxylate **29** (0.325 g, 1 mmol) was dissolved in methanol (5 mL) and aqueous KOH (2 mL, 5N), and refluxed at 70°C for 3 hours with continuous stirring. The solvent was evaporated under reduced pressure to obtain white solid, which was slurried with methanol (10 mL) for 15 minutes, after which the suspension was filtered, and the filtrate dried over anhydrous sodium sulfate and concentrated under the reduced pressure to afford a white solid. The solid was dissolved in water (2 mL), and the pH of the solution was adjusted to four with concentrated hydrochloric acid. The precipitate formed at pH four was collected and dried to obtain of the desired product (0.261 g, 85%) as a white solid. The compound was used in the subsequent step without further purification.

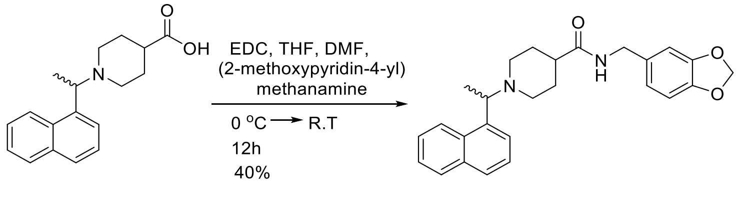
(±)-*N*-((2-Methoxypyridin-4-yl)methyl)-1-(1-(naphthalen-1-yl)ethyl)piperidine-4-carboxamide (31).



1-(1-(Naphthalen-1-yl)ethyl)piperidine-4-carboxylic acid **30** (0.141 g, 0.5 mmol) and (2-methoxypyridin-4-yl)methanamine (0.069 g, 0.5 mmol) was dissolved in anhydrous THF (3 mL) and dimethylformamide (DMF) (1 mL). Then reaction mixture was cooled to the 0°C and 1-ethyl-3-(-3-dimethylaminopropyl) carbodiimide hydrochloride (0.115 g, 0.6 mmol) was added. The temperature was raised to room temperature and the reaction mixture was stirred for 12 hr. After completion, the solvent was removed under reduced pressure to give a yellowish white solid, which was dissolved in dichloromethane (10 mL) and washed with water (5 mL), saturated sodium bicarbonate (5 mL), and then water (5 mL) and dried over anhydrous sodium sulfate. The solvent was evaporated under reduced pressure to obtain a white solid. The solid was subjected to silica gel column chromatography using 100% dichloromethane to 5% methanol in dichloromethane as eluent to afford the title product (0.065 g, 32%) as white solid. The ^{1}H and ^{13}C NMR of the product match with the literature.[5] ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.40 (d, *J* = 8.3 Hz, 1H), 8.08 (d, *J* = 5.3 Hz, 1H), 7.86 – 7.80 (m, 1H), 7.72 (d, *J* = 8.1 Hz, 1H), 7.55 (d, *J* = 7.1 Hz, 1H), 7.50 – 7.35 (m, 3H), 6.73 (dd, *J* = 5.3, 1.5 Hz, 1H), 6.61 (d, *J* = 1.5 Hz, 1H), 5.98 (t, *J* = 5.9 Hz, 1H), 5.66 (m, 1H), 5.10 – 4.93 (m, 2H), 4.37 (d, *J* = 5.8 Hz, 2H), 4.11 – 4.05 (m, 1H), 2.84 (m, 1H), 2.58 (m, 1H), 2.42 – 2.16 (m, 4H), 2.05 (m, 1H), 1.91 (m, *J* = 13.7, 5.3, 2.5 Hz, 1H), 1.59 (m, 2H), 1.44 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ

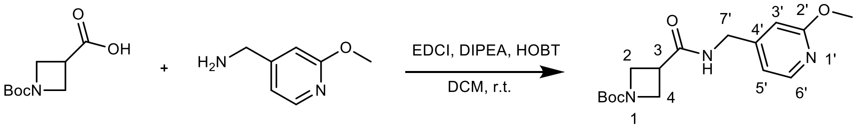
175.6, 164.8, 150.5, 147.2, 140.9, 134.1, 133.3, 131.7, 128.8, 127.4, 125.6, 125.5, 125.4, 124.6, 124.2, 118.6, 115.9, 109.3, 61.7, 53.5, 48.6, 47.4, 45.0, 43.7, 42.4, 33.8, 33.7, 18.9.

(±)-*N*-(Benzo[*d*][1,3]dioxol-5-ylmethyl)-1-(1-(naphthalen-1-yl)ethyl)piperidine-4-carboxamide (1).



1-(1-(Naphthalen-1-yl)ethyl)piperidine-4-carboxylic acid **30** (0.0705 g, 0.25 mmol) and 1,3-benzodioxole-5-methylamine (0.057 g, 0.375 mmol) was dissolved in anhydrous THF (2 mL) and DMF (1 mL). Then the reaction mixture was cooled to 0°C and 1-ethyl-3-(-3-dimethylaminopropyl) carbodiimide hydrochloride (0.071 g, 0.375 mmol) was added. The temperature was raised to room temperature and the reaction mixture was stirred for 12 hr. After completion, the solvent was removed under reduced pressure to obtain a yellowish white solid, which was dissolved in dichloromethane (5 mL) and washed with water (2.5 mL), saturated sodium bicarbonate (2.5 mL), and then water (2.5 mL) and dried over anhydrous sodium sulfate. The solvent was removed under reduced pressure to afford a white solid. The white solid was subject silica gel column chromatography using 100% dichloromethane to 5% methanol in dichloromethane as eluent to afford the title product (0.042 g, 40%) as a white solid. The ^{1}H and ^{13}C NMR of the product match with the literature.[6] ^{1}H NMR (500 MHz, MeOH-*D4*) δ 8.39 (d, *J* = 8.1 Hz, 1H), 7.82 (dd, *J* = 7.7, 1.9 Hz, 1H), 7.72 (d, *J* = 8.2 Hz, 1H), 7.55 (dd, *J* = 7.2, 1.2 Hz, 1H), 7.49 – 7.35 (m, 3H), 6.73 – 6.65 (m, 3H), 5.86 (s, 2H), 4.22 – 4.14 (m, 3H), 3.27 – 3.20 (m, 1H), 2.88 – 2.76 (m, 1H), 2.25 – 2.11 (m, 1H), 2.11 – 1.93 (m, 2H), 1.81 – 1.72 (m, 2H), 1.72 – 1.57 (m, 2H), 1.45 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (126 MHz, MeOH-*D4*) δ 176.7, 147.9, 146.8, 140.0, 134.3, 132.8, 131.8, 128.4, 127.2, 125.30, 125.1, 125.0, 124.3, 123.7, 120.5, 107.7, 101.0, 61.0, 51.3, 49.4, 48.2, 48.0, 47.9, 47.8, 47.7, 47.5, 47.3, 47.2, 43.1, 42.4, 28.8, 17.8.

***tert*-Butyl-3-(((2'-methoxypyridin-4'-yl)methyl)carbamoyl)azetidine-1-carboxylate (32).**

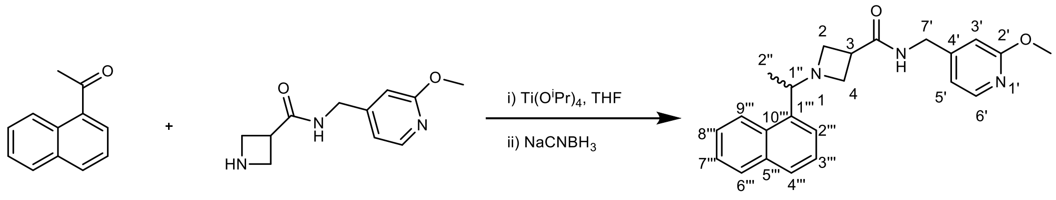


To a stirred solution of 1-(*tert*-butoxycarbonyl)azetidine-3-carboxylic acid (500 mg, 2.5 mmol), EDCI (485 mg, 3.13 mmol) and HOBT (422 mg, 3.13 mmol) in anhydrous DCM (10 mL) was added (2-methoxypyridin-4-yl)methanamine (345 mg, 2.5 mmol) followed by DIPEA (1.7 mL, 9.75 mmol) at 0°C. The reaction mixture was gradually brought up to r.t. over a period of 15 min, and stirred for 16 hr. After such time, the reaction mixture was quenched *via* addition of H_2O (20 mL) and the aqueous layer re-extracted with DCM (20 mL, 3X). The organic layers were combined, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue obtained was subjected to flash column chromatography on silica (eluent: 5:95 MeOH:DCM) to afford the title compound (610 mg, 1.90 mmol, 75%) as a white solid. $R_f = 0.20$ (5:95 MeOH:DCM; CAM). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.09 (d, *J* = 5.3 Hz, 1H, H-6'), 6.75 (dd, *J* = 5.2, 1.7 Hz, 1H, H-5'), 6.59 (s, 1H, H-3'), 5.91 (s, 1H, NH), 4.41 (d, *J* = 6.0 Hz, 2H, CH_2-aryl), 4.15 – 4.08 (m, 2H, azetidine CH_2), 4.07 (q, *J* = 9.8, 8.4 Hz, 2H, azetidine CH_2), 3.91 (d, *J* = 1.4 Hz, 3H, OCH_3-aryl), 3.21 (dddd, *J* = 14.6, 8.3, 6.1, 1.3 Hz, 1H, azetidine CH), 1.42 (s, 9H, 3 x CH_3-Boc). ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.3 (C(O)NHR), 164.6 (C2'), 156.3 (C(O)Boc), 150.3 (C4'), 146.9 (C6'), 115.7 (C3’), 108.8 (C4’), 79.9 (C(R)$_3$Boc), 53.4 (OCH$_3$-aryl)), 51.6 (2C) (azetidine CH$_2$), 42.2 (CH$_2$-aryl), 32.9 (azetidine CH), 28.3 (3C) (3 x CH$_3$ (Boc)). HRMS (ESI) *m/z*: Calcd for $[C_{16}H_{24}O_4N_4]^+$ $[M+H]^+$ 322.1761; Found 322.1748.

***N*-((2’-Methoxypyridin-4’-yl)methyl)azetidine-3-carboxamide (33).**

To a stirred solution of *tert*-butyl-3-(((2-methoxypyridin-4-yl)methyl)carbamoyl)azetidine-1-carboxylate **32** (610 mg, 1.9 mmol) in anhydrous DCM (2.83 mL) was added TFA (2.83 mL, 37 mmol) dropwise at r.t. The solution was stirred for 45 min, after such time, the reaction mixture was washed with NaHCO$_3$ (15 mL, 1X) and the organic layers concentrated *in vacuo*. The solution was then co-concentrated with toluene (20 mL, 3X) to afford the title compound (400 mg, 1.8 mmol, 95%) as a yellow oil. $R_f = 0.10$ (10:90 MeOH:DCM; CAM). ^{1}H NMR (500 MHz, CD$_3$OD): δ 8.11 (dd, *J* = 5.8, 1.3 Hz, 1H, H-6’), 7.07 (dt, *J* = 5.9, 1.5 Hz, 1H, H-5’), 6.99 (s, 1H, H-3’), 4.47 (s, 2H, CH$_2$-aryl), 4.27 – 4.17 (m, 4H, 2 x azetidine CH$_2$), 4.00 (d, *J* = 1.4 Hz, 3H, OCH$_3$-aryl), 3.78 – 3.67 (m, 1H, azetidine CH). ^{13}C NMR (126 MHz, CD$_3$OD): δ 172.6 (C(O)NHR), 164.7 (C2’), 157.1 (C4’), 144.7 (C6’), 117.4 (C3’), 109.7 (C4’), 56.0 (OCH$_3$-aryl), 49.6 (2 x azetidine CH$_2$), 43.4 (CH$_2$-aryl), 37.0 (azetidine CH). HRMS (ESI) *m/z*: Calcd for $[C_{11}H_{16}O_2N_3]^+$ $[M+H]^+$ 222.1237; Found 222.1229.

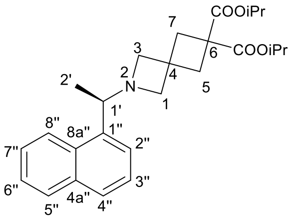
(±)-*N*-((2’-Methoxypyridin-4’-yl)methyl)-1-(1’’-(naphthalene-1’’-yl)ethyl)azetidine-3-carboxamide (34).



To a stirred solution of 1-acetonaphthone (230 mg, 1.5 mmol) in anhydrous THF (5 mL) was added *N*-((2’-methoxypyridin-4’-yl)methyl)azetidine-3-caboxamide **33** (400 mg, 1.8 mmol), followed by Ti(O^{i}Pr)$_4$ (532 µL, 1.8 mmol) at r.t. The reaction mixture was stirred for 1 h, after which, it was cooled down to -78°C, and NaCNBH$_3$ (762 mg, 3.6 mmol) was added portion-wise. The mixture was stirred for 15 min at -78°C, after which, it was slowly brought to r.t. over 20 min and stirred for an additional 1 hr. After such time, the mixture was quenched *via* addition of H$_2$O (5 mL) and diluted with EtOAc (10 mL). The aqueous layer was re-extracted with EtOAc (10 mL, 3X). The organic layers were combined and dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. The residue obtained was subjected to flash column chromatography on silica (eluent: 98:2 EtOAc:Et$_3$N) to afford the title compound (169 mg, 0.45 mmol, 30%) as a white foam. R$_f$ = 0.40 (98:2 EtOAc:Et$_3$N; CAM). ^{1}H NMR (500 MHz, CD$_3$OD): δ 8.22 (d, *J* = 8.4 Hz, 1H, H-9’’’), 7.99 (d, *J* = 5.4 Hz, 1H, H-6’), 7.88 – 7.75 (m, 1H, H-6’’’), 7.71 (d, *J* = 8.2 Hz, 1H, H-4’’’), 7.55 (dd, *J* = 7.3, 1.3 Hz, 1H, H-2’’’), 7.48 (ddd, *J* = 8.6, 6.8, 1.6 Hz, 1H, H-8’’’), 7.46 – 7.37 (m, 2H, H-7’’’, H-3’’’), 6.80 (dd, *J* = 5.2, 1.4 Hz, 1H, H-5’), 6.69 – 6.61 (m, 1H, H-3’), 4.30 (dd, *J* = 8.9, 3.2 Hz, 3H, CH$_2$-aryl, H-1’’), 3.83 (s, 3H, OCH$_3$-aryl), 3.66 (td, *J* = 7.6, 1.7 Hz, 1H, azetidine CH$_2$), 3.47 – 3.37 (m, 2H, azetidine CH$_2$), 3.35 – 3.19 (m, 1H, azetidine CH), 3.14 (t, *J* = 7.4 Hz, 1H, azetidine CH$_2$), 1.29 (d, *J* = 6.5 Hz, 3H, CH$_3$-H2’’). ^{13}C NMR (126

MHz, CD_3OD): 174.9 (C(O)NHR), 166.1 (C2’), 152.8 (C4’), 147.8 (C6’), 139.5 (C1’’’), 135.5 (C10’’’), 132.6 (C5’’’), 129.9 (C6’’’), 128.6 (C4’’’), 126.9 (C8’’’), 126.6 (2C) (C7’’’, C3’’’), 124.6 (C2’’’), 124.1 (C9’’’), 116.9 (C5’), 109.6 (C3’), 64.9 (C1’’), 57.0 (azetidine CH_2), 56.7 (azetidine CH_2), 54.1 (OCH_3-aryl), 42.9 (CH_2-aryl), 35.3 (azetidine CH), 20.5 (C2’’). HRMS (ESI) *m/z*: Calcd for $[C_{23}H_{26}O_2N_3]^+$ $[M+H]^+$ 376.2019; Found 376.2003.

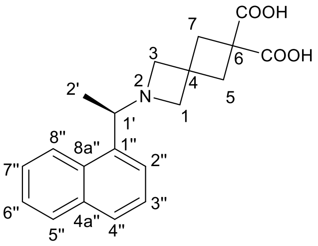
Diisopropyl (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6,6-dicarboxylate (35).



Diisopropyl 3,3-bis(hydroxymethyl)cyclobutane-1,1-dicarboxylate[7] (0.5 g, 1.73 mmol) was dissolved in dry CH_3CN (4 mL) and the solution was cooled to -20°C. Triflic anhydride (0.62 mL, 3.64 mmol) was added in a dropwise manner while keeping temperature below -10°C. DIPEA (0.76 mL, 4.34 mmol) was then added slowly over 15 mins to the reaction mixture. After formation of the bistriflate (by mass spectral analysis), additional DIPEA (0.76 mL, 4.34 mmol) was added in a dropwise manner followed by (*R*)-(+)-1-(1-naphthyl)ethylamine (0.28 mL, 1.73 mmol), and the resulting mixture was heated to 70°C for 2 hr. After completion, the reaction mixture was cooled to room temperature and diluted with 25 mL of toluene and washed with 3 times with 20 mL of water. The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a crude residue which was purified by silica gel column chromatography (1:2, Ethyl

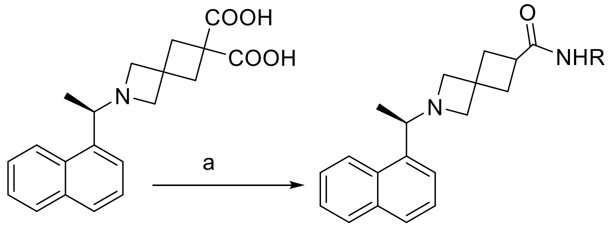
Acetate:Hexane) to furnish the title compound (0.58g, 78%) as a colorless syrup. $[\alpha]_D^{22}$ = +25.5 (c = 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.18 (d, *J* = 8.2 Hz, 1H, $H_{8''}$), 7.83 (dd, *J* = 8.1, 1.6 Hz, 1H, $H_{5''}$), 7.70 (d, *J* = 8.2 Hz, 1H, $H_{4''}$), 7.62 (d, *J* = 7.1 Hz, 1H, $H_{2''}$), 7.52 – 7.38 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 5.02 (hept, *J* = 6.3 Hz, 2H, -*CH*(CH_3)$_2$), 4.02 (q, *J* = 6.5 Hz, 1H, $H_{1'}$), 3.24 (q, *J* = 7.7 Hz, 4H, H_1, H_3), 2.67 (qd, *J* = 11.5, 1.8 Hz, 4H, H_5, H_7), 1.30 (d, *J* = 6.5 Hz, 3H, $H_{2'}$), 1.21 (d, *J* = 2.4 Hz, 6H, -CH*(CH_3)$_2$*), 1.20 (d, *J* = 2.4 Hz, 6H, -CH*(CH_3)$_2$*). ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.2 (CO), 139.2 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.2 ($C_{8a''}$), 129.0 ($C_{5''}$), 127.2 ($C_{4''}$), 125.8 ($C_{3''}$), 125.7 ($C_{2''}$), 125.2 ($C_{6''}$), 123.9 ($C_{7''}$), 123.3 ($C_{8''}$), 68.8 (-*CH*(CH_3)$_2$), 65.8 (C_1, C_3), 64.7 ($C_{1'}$) 49.1, (C_4), 39.5 (C_5, C_7), 32.7 (C_6), 21.6 (-CH*(CH_3)$_2$*), 20.6 ($C_{2'}$). HRMS (ESI) m/z: Calcd for $C_{26}H_{34}NO_4$ $[M + H]^+$ 424.2482; Found 424.2473.

(*R*)-2-(1-(Naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6,6-dicarboxylic acid (36).



To a solution of diisopropyl (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6,6-dicarboxylate **35** (0.50g, 1.15 mmol) in ethanol (5 mL), a solution of NaOH (47mg, 3.90 mmol) in ethanol (1 mL) was added. The resulting mixture was refluxed for 2 hr and cooled down to room temperature. A white salt formed and was filtered, washed with EtOH (2×10 mL) and dried under high vacuum to afford the disodium salt of **35**, which was dissolved in water (5 mL) and stirred for 10 min until homogeneity was achieved. The resulting solution was acidified with 2N HCl to pH = 2, leading to the formation of a slurry, which was cooled to 0°C and stirred for 1 hr. The so-obtained solid was filtered and washed with MeOH (2×10 mL) and dried under high vacuum to furnish the title compound (0.34g, 86%) as a white solid. m.p.: 252-256°C; $[\alpha]_D^{22}$ = +28.7 (c = 0.7, DMSO); ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.21 (d, *J* = 8.2 Hz, 1H, $H_{8''}$), 7.88 (dd, *J* = 7.7, 1.8 Hz, 1H, $H_{5''}$), 7.75 (d, *J* = 8.0 Hz, 1H, $H_{4''}$), 7.61 (d, *J* = 7.0 Hz, 1H, $H_{2''}$), 7.57 – 7.41 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 4.48 (q, *J* = 6.6 Hz, 1H, $H_{1'}$), 4.32 (d, *J* = 9.9 Hz, 1H, H_3), 4.25 (d, *J* = 9.7 Hz, 1H, H_3), 2.50 (d, *J* = 12.5 Hz, 1H, H_1), 2.37 (d, *J* = 12.5 Hz, 1H, H_1), 2.29 (dd, *J* = 15.4, 8.9 Hz, 2H, H_5, H_7), 2.03 – 1.87 (m, 2H, H_5, H_7), 1.35 (d, *J* = 6.6 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, DMSO-D_6) δ 172.0 (COOH), 171.1 (COOH), 141.7 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.4 ($C_{8a''}$), 129.2 ($C_{5''}$), 127.3 ($C_{4''}$), 126.3 ($C_{3''}$), 126.2 ($C_{2''}$), 125.9 ($C_{6''}$), 123.6 ($C_{7''}$), 123.3 ($C_{8''}$), 75.9 (C_3), 54.2 ($C_{1'}$), 50.5 (C_1), 49.2 (C_4), 38.3 (C_6), 34.8 (C_7), 34.7 (C_5), 24.2 ($C_{2'}$). HRMS (ESI) *m*/*z*: Calcd for $C_{20}H_{22}NO_4$ $[M + H]^+$ 340.1543; Found 340.1528.

General procedure for synthesis of amides from (*R*)-2-(1-(Naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6,6-dicarboxylic acid 36.

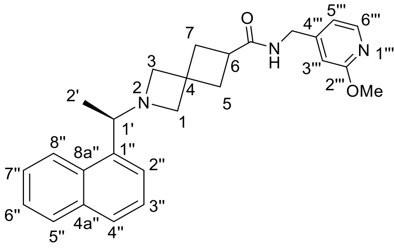


3 **4a-f**

^{a}Reagents and conditions: (a) RNH_2, carbonyl diimidazole, THF, rt to reflux

To a stirred suspension of the diacid **36** (50 mg, 0.147 mmol) in dry THF (2 mL). was added 1.1'-carbonyl diimidazole (26 mg, 0.22 mmol) in two portions. The resulting mixture was stirred at rt for 2 hr under an argon atmosphere. The amine (0.16 mmol) was added in one portion and the mixture refluxed (70°C) for 0.5 hr. The reaction mixture was cooled down to room temperature and the solvent was removed *in vacuo* to give a crude residue, which was dissolved in 20 mL of ethyl acetate and washed with saturated $NaHCO_3$ solution (10 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a crude residue which was purified by silica gel column chromatography (1-10% MeOH/CH_2Cl_2) to afford the desired amide.

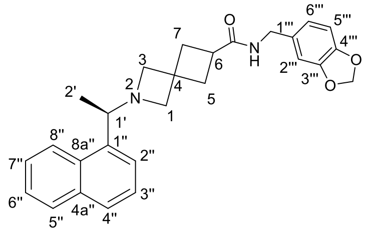
_N_-((2-Methoxypyridin-4-yl)methyl) (_R_)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxamide (37).



This colorless oil was prepared according to the general procedure (44 mg, 72%); $[\alpha]_D^{22}$ = +8.9 (c = 0.5, $CHCl_3$); ^1H NMR (500 MHz, Chloroform-*d*) δ 8.15 (d, *J* = 8.1 Hz, 1H, $H_{8''}$), 8.06 (d, *J* = 5.3 Hz, 1H, $H_{6'''}$), 7.84 (dd, *J* = 8.1, 1.6 Hz, 1H, $H_{5''}$), 7.71 (d, *J* = 8.2 Hz, 1H, $H_{4''}$), 7.64 (d, *J* = 7.2 Hz, 1H, $H_{2''}$), 7.53 – 7.41 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 6.72 (dd, *J* = 5.4, 1.5 Hz, 1H, H5'''), 6.56 (s, 1H, H2'''), 5.76 (t, *J* = 6.0 Hz, 1H, -NH), 4.35 (d, *J* = 6.0 Hz, 2H, -NH*CH₂*), 4.04 (q, *J* = 6.3 Hz, 1H, $H_{1'}$), 3.89 (s, 3H, -OCH_3), 3.38 – 3.19 (m, 4H, H_1, H_3), 2.85 (p, *J* = 8.1 Hz, 1H, H_6), 2.49 – 2.41 (m, 2H, H_5, H_7), 2.40 – 2.29 (m, 2H, H_5, H_7), 1.33 (d, *J* = 6.6 Hz, 3H, $H_{2'}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.6 (-CONH), 164.7 ($C_{2'''}$), 150.2 ($C_{4'''}$), 147.2 ($C_{6'''}$), 139.1 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.2 ($C_{8a''}$), 129.0 ($C_{5''}$), 127.2 ($C_{4''}$), 125.8 ($C_{3''}$), 125.7 ($C_{2''}$), 125.2 ($C_{6''}$), 123.9 ($C_{7''}$), 123.2 ($C_{8''}$), 115.7 ($C_{3'''}$), 109.0 ($C_{4'''}$), 66.1 (C_3), 65.2 (C_1), 64.7 ($C_{1'}$), 53.5 (OCH_3), 42.3 (-$NHCH_2$), 36.2 (C_5), 35.8 (C_7), 35.0 (C_6), 34.5

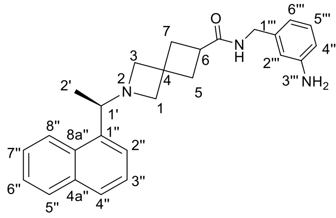
(C_4), 20.6 ($C_{2'}$). HRMS (ESI) *m/z*: Calcd for $C_{26}H_{30}N_3O_2$ $[M + H]^+$ 416.2332; Found 416.2317.

N-(Benzo[d][1,3]dioxol-5-ylmethyl) (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxamide (38).



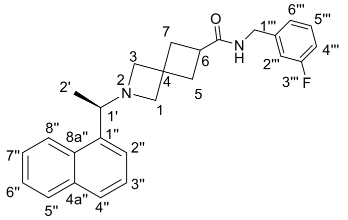
This colorless syrup was prepared according to the general procedure (51 mg, 81%); $[\alpha]_D^{22}$ = +9.7 (c = 0.7, $CHCl_3$); ^1H NMR (500 MHz, $CDCl_3$) δ 8.17 (d, *J* = 8.4 Hz, 1H, $H_{8''}$), 7.83 (dd, *J* = 7.8, 1.7 Hz, 1H, $H_{5''}$), 7.71 (d, *J* = 8.1 Hz, 1H, $H_{4''}$), 7.62 (d, *J* = 7.1 Hz, 1H, $H_{2''}$), 7.55 – 7.38 (m, 3H. $H_{3''}$, $H_{6''}$, $H_{7''}$), 6.80 – 6.68 (m, 2H, $H_{3'''}$, $H_{6'''}$), 6.68 (dd, *J* = 7.9, 1.7 Hz, 1H. $H_{4'''}$), 5.90 (s, 2H, O*CH₂*O), 5.67 (t, *J* = 5.7 Hz, 1H. -NH), δ 4.28 (d, *J* = 5.7 Hz, 1H, -NH*CH₂*), 4.03 (q, *J* = 6.5 Hz, 1H, $H_{1'}$), 3.25 (dd, *J* = 15.6, 7.6 Hz, 2H, H_3), 3.18 (t, *J* = 7.0 Hz, 2H, H_1), 2.79 (p, *J* = 8.3 Hz, 1H, H_6), 2.42 (dt, *J* = 11.6, 8.0 Hz, 2H, H_5, H_7), 2.38 – 2.21 (m, 2H, H_5, H_7), 1.31 (d, *J* = 6.5 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.3 (-CONH), 148.0 ($C_{4'''}$), 147.0 ($C_{3'''}$), 139.2 ($C_{1''}$), 134.0 ($C_{4a''}$), 132.3 ($C_{1'''}$), 131.2 ($C_{8a''}$), 129.0 ($C_{5''}$), 127.2 ($C_{4''}$), 125.8 ($C_{3''}$), 125.7 ($C_{2''}$), 125.2 ($C_{6''}$), 123.9 ($C_{7''}$), 123.3 ($C_{8''}$), 121.1 ($C_{6'''}$), 108.4 ($C_{2'''}$), 108.3 ($C_{5'''}$), 101.1 ($-OCH_2O-$), 66.2 (C_3), 65.2 (C_1), 64.7 ($C_{1'}$), 43.5 ($-NHCH_2$), 36.2 (C_5), 35.9 (C_7), 35.1 (C_6), 34.4 (C_4), 20.6 ($C_{2'}$). HRMS (ESI) *m/z*: Calcd for $C_{27}H_{29}N_2O_3$ $[M + H]^+$ 429.2172; Found 429.2167.

N-(3-Aminobenzyl) (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxamide (39).



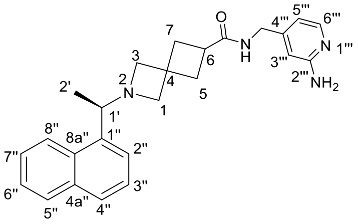
This colorless syrup was prepared according to the general procedure (40 mg, 68%); $[\alpha]_D^{22}$ = +24.5 (c = 0.4, $CHCl_3$); ^1H NMR (500 MHz, $CDCl_3$) δ 8.19 (d, *J* = 8.5 Hz, 1H, $H_{8''}$), 7.84 (d, *J* = 8.0 Hz, 1H, $H_{5''}$), 7.71 (d, *J* = 8.2 Hz, 1H, $H_{4''}$), 7.63 (d, *J* = 7.2 Hz, 1H, $H_{2''}$), 7.46 (dq, *J* = 21.2, 7.3 Hz, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 7.07 (t, *J* = 7.6 Hz, 1H, $H_{5'''}$), 6.60 (d, *J* = 7.6 Hz, 1H, $H_{2'''}$), 6.54 (d, *J* = 8.2 Hz, 2H, $H_{4'''}$, $H_{6'''}$), 5.72 (t, *J* = 5.8 Hz, 1H, -NH), 4.28 (d, *J* = 5.7 Hz, 2H, -NH*CH₂*), 4.03 (q, *J* = 6.5 Hz, 1H, $H_{1'}$), 3.65 (s, 2H, NH_2), 3.36 – 3.07 (m, 4H, H_1, H_3), 2.79 (p, *J* = 8.3 Hz, 1H, H_6), 2.43 (dt, *J* = 14.2, 7.2 Hz, 2H, H_5, H_7), 2.38 – 2.21 (m, 2H, H_5, H_7), 1.31 (d, *J* = 6.5 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.4 (-CONH), 146.9 ($C_{3'''}$), 139.6 ($C_{1'''}$), 139.3 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.2 ($C_{8a''}$), 129.7 ($C_{5'''}$), 129.0 ($C_{5''}$), 127.2 ($C_{4''}$), 125.8 ($C_{3''}$), 125.7 ($C_{2''}$), 125.2 ($C_{6''}$), 123.9 ($C_{7''}$), 123.3 ($C_{8''}$), 117.8 ($C_{6'''}$), 114.4 ($C_{2'''}$), 114.3 ($C_{4'''}$), 66.2 (C_3), 65.2 (C_1). 64.7 ($C_{1'}$), 43.7 (-$NHCH_2$), 36.3 (C_5), 35.9 (C_7), 35.1 (C_6), 34.4 (C_4), 20.7 ($C_{2'}$). HRMS (ESI) *m/z*: Calcd for $C_{26}H_{30}N_3O$ $[M + H]^+$ 400.2383; Found 400.2372.

N-(3-Fluorobenzyl) (R)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxamide (40).



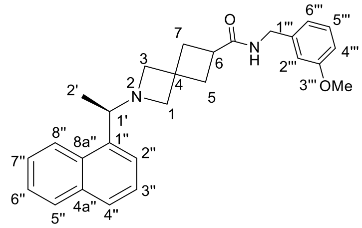
This colorless syrup was prepared according to the general procedure (52 mg, 88%); $[\alpha]_D^{22}$ = +7.4 (c = 0.6, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.15 (d, *J* = 8.3 Hz, 1H, $H_{8''}$), 7.84 (dd, *J* = 8.0, 1.6 Hz, 1H, $H_{5''}$), 7.72 (d, *J* = 8.2 Hz, 1H, $H_{4''}$), 7.65 (d, *J* = 7.1 Hz, 1H, $H_{2''}$), 7.57 – 7.39 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 7.31 – 7.20 (m, 1H, $H_{5'''}$), 7.00 (d, *J* = 7.7 Hz, 1H, $H_{2'''}$), 6.98 – 6.87 (m, 2H, $H_{4'''}$, $H_{6'''}$), 5.70 (t, *J* = 6.0 Hz, 1H, -NH), 4.39 (d, *J* = 5.8 Hz, 2H, -NH*CH₂*-), 4.11 (s, 1H, $H_{1'}$), 3.30 (dd, *J* = 33.5, 7.9 Hz, 4H, H_1, H_3), 2.83 (p, *J* = 8.2 Hz, 1H, H_6), 2.45 (dt, *J* = 11.7, 7.2 Hz, 2H, H_5, H_7), 2.42 – 2.25 (m, 2H, H_5, H_7), 1.34 (d, *J* = 6.5 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.4, 163.0 (d, *J* = 247.2, $C_{3'''}$), 141.07 ($C_{1'''}$), 141.01 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.1 ($C_{8a''}$), 130.3 (d, *J* = 8.3, $C_{5'''}$), 129.0 ($C_{5''}$), 127.4 ($C_{4''}$), 125.8 ($C_{3''}$, $C_{2''}$), 125.3 ($C_{6''}$), 124.0 ($C_{7''}$), 123.2 (d, *J* = 2.7, $C_{6'''}$), 123.1 ($C_{8''}$), 114.6 (d, *J* = 19.8, $C_{2'''}$), 114.4 (d, *J* = 19.6, $C_{4'''}$), 65.9 (C_3), 65.1 (C_1), 64.4 ($C_{1'}$), 43.1 (-NHCH₂), 36.2 (C_5), 35.8 (C_7), 35.0 (C_6), 34.5 (C_4), 20.4 ($C_{2'}$). HRMS (ESI) *m/z*: Calcd for $C_{26}H_{28}N_2OF$ $[M + H]^+$ 403.2180; Found 403.2169.

N-((2-Aminopyridin-4-yl)methyl) (R)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxamide (41).



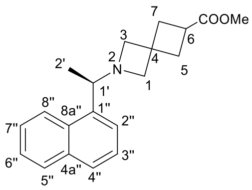
This white foam was prepared according to the general procedure (43 mg, 73%); $[\alpha]_D^{22}$ = +13.0 (c = 0.7, CHCl$_3$); ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.18 (s, 1H, H$_{8''}$), 7.98 (d, *J* = 5.6 Hz, 1H, H$_{6'''}$), 7.88 – 7.80 (m, 1H, H$_{5''}$), 7.71 (d, *J* = 8.0 Hz, 1H, H$_{4''}$), 7.61 (d, *J* = 7.2 Hz, 1H, H$_{2''}$), 7.53 – 7.38 (m, 3H, H$_{3''}$, H$_{6''}$, H$_{7''}$), 6.53 – 6.46 (m, 1H, H$_{5'''}$), 6.34 (s, 1H, H$_{2'''}$), 5.62 (t, *J* = 5.9 Hz, 1H, -NH), 4.38 (s, 2H. NH$_2$), 4.29 (d, *J* = 6.0 Hz, 2H, -NH$\mathit{CH_2}$), 4.03 (q, *J* = 6.5 Hz, 1H, H$_{1'}$), 3.34 – 3.15 (m, 4H, H1, H$_3$), 2.85 (p, *J* = 8.4 Hz, 1H, H$_6$), 2.51 – 2.24 (m, 4H, H$_5$, H$_7$), 1.30 (d, *J* = 6.6 Hz, 3H, H$_{2'}$); ^{13}C NMR (126 MHz, CDCl$_3$) δ 174.6 (CONH-), 158.8 (C$_{2'''}$), 149.2 (C$_{4'''}$), 148.6 (C$_{6'''}$), 139.2 (C$_{1''}$), 134.0 (C$_{4a''}$), 131.2 (C$_{8a''}$), 129.0 (C$_{5''}$), 127.2 (C$_{4''}$), 125.7 (C$_{3''}$), 125.7 (C$_{2''}$), 125.2 (C$_{6''}$), 123.9 (C$_{7''}$), 123.2 (C$_{8''}$) 113.0 (C$_{5'''}$), 106.8 (C$_{3'''}$), 66.1 (C$_3$), 65.2 (C$_1$), 64.7 (C$_{1'}$), 42.5 (-NHCH$_2$), 36.3 (C$_5$), 35.9 (C$_7$), 35.1 (C$_6$), 34.5 (C$_4$), 20.6 (C$_{2'}$). HRMS (ESI) *m/z*: Calcd for $C_{25}H_{29}N_4O$ $[M + H]^+$ 401.2335; Found 401.2320.

***N*-(3-Methoxybenzyl) (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxamide (42).**



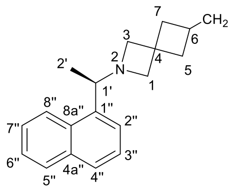
This colorless syrup was prepared according to the general procedure (51 mg, 84%); $[\alpha]_D^{22}$ = +5.3 (c = 1.0, $CHCl_3$); ^1H NMR (500 MHz, $CDCl_3$) δ 8.17 (d, *J* = 8.1 Hz, 1H, $H_{8''}$), 7.84 (dd, *J* = 8.0, 1.6 Hz, 1H, $H_{5''}$), 7.71 (d, *J* = 8.1 Hz, 1H, $H_{4''}$), 7.62 (d, *J* = 7.1 Hz, 1H, $H_{2''}$), 7.53 – 7.38 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 7.22 (t, *J* = 7.8 Hz, 1H, $H_{5'''}$), δ 6.84 – 6.77 (m, 3H, $H_{2'''}$, $H_{4'''}$, $H_{6'''}$), 5.62 (t, *J* = 5.8 Hz, 1H, -NH), 4.37 (d, *J* = 5.7 Hz, 2H, -NHCH_2), 4.04 (q, *J* = 6.6 Hz, 1H, $H_{1'}$), 3.77 (s, 3H, -OCH_3), 3.27 (dd, *J* = 15.6, 7.7 Hz, 2H, H_3), 3.22 – 3.14 (m, 2H, H_1), 2.81 (p, *J* = 8.3 Hz, 1H, H_6), 2.45 (dt, *J* = 11.6, 7.6 Hz, 2H, H_5, H_7), 2.39 – 2.24 (m, 2H, H_5, H_7), 1.31 (d, *J* = 6.4 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.3 (-CONH), 159.9 ($C_{3'''}$), 140.0 ($C_{1'''}$), 139.2 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.2 ($C_{8a''}$), 129.8 ($C_{5'''}$), 129.0 ($C_{5''}$), 127.2 ($C_{4''}$), 125.8 ($C_{3''}$), 125.7 ($C_{2''}$), 125.2 ($C_{6''}$), 123.9 ($C_{7''}$), 123.2 ($C_{8''}$), 120.0 ($C_{6'''}$), 113.4 ($C_{4'''}$), 113.0 ($C_{2'''}$), 66.1 (C_3), 65.2 (C_1), 64.7 ($C_{1'}$), 55.3 (-OCH_3), 43.6 (-$NHCH_2$), 36.3 (C_5), 35.9 (C_7), 35.1 (C_6), 34.4 (C_4), 20.6 ($C_{2'}$). HRMS (ESI) *m*/*z*: Calcd for $C_{27}H_{31}N_2O_2$ $[M + H]^+$ 415.2380; Found 415.2373.

Methyl (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-carboxylate (43).



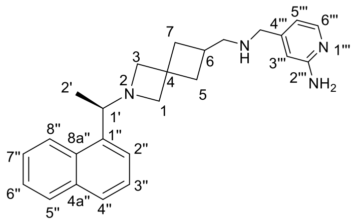
To a stirred suspension of (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6,6-dicarboxylic acid **36** (0.50g, 1.47 mmol) in dry THF (5 mL). was added 1.1’-carbonyl diimidazole (0.36g, 2.21 mmol) in two portions. The resulting mixture was stirred at rt for 2h under an argon atmosphere. Dry MeOH (2 mL) was added in one portion and the mixture refluxed (70°C) for 0.5 hr. The reaction mixture was cooled down to room temperature and solvent was removed *in vacuo* to give a crude residue. This residue was dissolved in 20 mL of ethyl acetate and washed with saturated $NaHCO_3$ solution (10 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a residue which was purified by neutral alumina column chromatography (5-10% EtOAc/Hexane) to furnish the title compound (0.42g, 92%) as a colorless syrup; $[\alpha]_D^{22}$ = +36.8 (c = 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.19 (d, *J* = 8.4 Hz, 1H, $H_{8''}$), 7.84 (dd, *J* = 8.0, 1.6 Hz, 1H, $H_{5''}$), 7.72 (d, *J* = 8.1 Hz, 1H, $H_{4''}$), 7.64 (d, *J* = 7.1 Hz, 1H, $H_{2''}$), 7.54 – 7.40 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 4.03 (q, *J* = 6.5 Hz, 1H, $H_{1'}$), 3.65 (s, 3H, COOMe), 3.27 (d, *J* = 7.4 Hz, 1H, H_3), 3.26 – 3.16 (m, 3H, H_3, H_1), 2.98 (p, *J* = 8.4 Hz, 1H, H_6), 2.48 – 2.30 (m, 4H, H_5, H_7), 1.32 (d, *J* = 6.5 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 175.8 (-CO), 139.3 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.2 ($C_{8a''}$), 129.0 ($C_{5''}$), 127.2 ($C_{4''}$), 125.8 ($C_{3''}$), 125.7 ($C_{2''}$), 125.2 ($C_{6''}$), 123.9 ($C_{7''}$), 123.3 ($C_{8''}$), 66.1 (C_3), 65.4 (C_1), 64.7 ($C_{1'}$) 51.7 (OMe), 36.2 (C_5), 35.9 (C_7), 34.7 (C_6), 33.3 (C_4), 20.7 ($C_{2'}$). HRMS (ESI) *m*/*z*: Calcd for $C_{20}H_{24}NO_2$ $[M + H]^+$ 310.1801; Found 310.1796.

(*R*)-2-(1-(Naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-methanol (44).



To a suspension of LAH (27.6 mg, 0.73 mmol) in THF (1 mL) was added a solution of **43** (150 mg, 0.49 mmol) in THF (1 mL) in a dropwise manner at 0ºC under an argon atmosphere. The reaction was then brought to rt, stirred until completion (monitored by TLC and mass spectral analysis), then quenched with ice cold water and diluted with CH_2Cl_2 (15 mL). The organic layer was separated, collected, dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a residue which was purified by neutral alumina column chromatography (20-30% EtOAc/Hexane) to furnish the title compound (0.12g, 87%) as white solid; m.p.: 128-132ºC; $[\alpha]_D^{22}$ = +32.1 (c = 0.6, DMSO); ^{1}H NMR (500 MHz, DMSO-*d*$_6$) δ 8.23 (d, *J* = 8.7 Hz, 1H, $H_{8''}$), 7.90 – 7.83 (m, 1H, $H_{5''}$), 7.73 (d, *J* = 8.1 Hz, 1H, $H_{4''}$), 7.55 (d, *J* = 7.1 Hz, 1H, $H_{2''}$), 7.45 (m, 3H, $H_{3''}$, $H_{6''}$, $H_{7''}$), 4.38 (t, *J* = 5.3 Hz, 1H, -OH), 3.96 (q, *J* = 6.6 Hz, 1H, $H_{1'}$), 3.27 – 3.23 (m, 2H, -*CH*$_2$OH), 3.12 – 3.03 (m, 2H, H_3), 3.03 – 2.95 (m, 2H, H_1), 2.15 (h, *J* = 6.9, 6.4 Hz, 1H, H_6), 2.07 – 1.99 (m, 2H, H_5, H_7), 1.82 – 1.72 (m, 2H, H_5, H_7), 1.16 (d, *J* = 6.4 Hz, 3H, $H_{2'}$); ^{13}C NMR (126 MHz, DMSO-*D*$_6$) δ 140.0 ($C_{1''}$), 134.0 ($C_{4a''}$), 131.2 ($C_{8a''}$), 129.2 ($C_{5''}$), 127.3 ($C_{4''}$), 126.2 ($C_{3''}$), 126.2 ($C_{2''}$), 125.8 ($C_{6''}$), 124.3 ($C_{7''}$), 123.8 ($C_{8''}$), 66.4 (C_3), 66.0 (C1), 65.5 ($C_{1'}$), 35.5 (-CH_2OH), 34.6 (C_6), 32.3 (C_4), 21.1 ($C_{2'}$). HRMS (ESI) *m/z*: Calcd for $C_{19}H_{24}NO$ [M + H]; 282.1852; Found 282.1845.

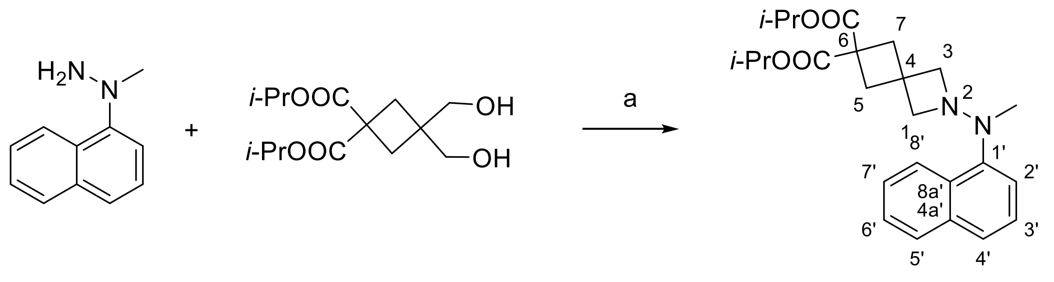
(*R*)-*N*-((2-Methoxypyridin-4-yl)methyl)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-methanamine (45).



A stirred solution of (*R*)-2-(1-(naphthalen-1-yl)ethyl)-2-azaspiro[3.3]heptane-6-methanol **44** (0.11g, 0.39 mmol) in CH_2Cl_2 (3 mL) was treated with Dess Martin Periodinane (0.33g, 0.78 mmol), and then stirred at rt for 2 hr. After complete conversion, the reaction mixture was diluted with 5 mL CH_2Cl_2, 5 mL of sat. aqueous $NaHCO_3$, 5 mL of aqueous sodium thiosulfate and stirred for 0.5 hr at room temperature. The organic layer was separated, collected, dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a residue which was used immediately for the next reaction without purification. The above residue was dissolved in dry MeOH (2 mL), treated with (2-methoxypyridin-4-yl)methanamine (0.14 mL, 0.8 mmol), and stirred at room temperature for 2 hr. $NaCNBH_3$ (74.2 mg, 1.18 mmol) and AcOH (0.09 mL, 1.57 mmol) were then added, and the mixture was stirred overnight at room temperature under an argon atmosphere. The reaction was quenched with 2 mL Et_3N, and the volatiles evaporated under high vacuum to give a residue, which was dissolved in CH_2Cl_2 (20 mL), and diluted with saturated aqueous $NaHCO_3$. The organic layer was separated, collected, dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give a residue, which was purified by silica gel column chromatography (5-10% MeOH/CH_2Cl_2) to furnish the title compound (91mg, 58%) as a colorless syrup; $[\alpha]_D^{22} = +10$ (c = 0.4, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ

8.20 – 8.13 (m, 1H, H$_{8''}$), 8.07 (d, *J* = 5.2 Hz, 1H, H$_{6'''}$), 7.87 – 7.82 (m, 1H, H$_{5''}$), 7.72 (d, *J* = 8.2 Hz, 1H, H$_{4''}$), 7.60 (d, *J* = 7.0 Hz, 1H, H$_{2''}$), 7.45 (m, 3H, H$_{3''}$, H$_{6''}$, H$_{7''}$), 6.87 (dd, *J* = 5.3, 1.3 Hz, 1H, H$_{5'''}$), 6.73 (s, 1H, H$_{2'''}$), 4.52 (q, *J* = 6.6 Hz, 1H, H$_{1'}$), 3.93 (s, 3H, OCH$_3$), 3.61 (s, 2H, -NH*CH$_2$*), 2.83 – 2.73 (m, 3H, H3, -*CH$_2$*NH), 2.71 (d, *J* = 9.4 Hz, 1H, H$_1$, H$_3$), 2.47 (d, *J* = 11.7 Hz, 1H, H$_1$), 2.37 (d, *J* = 11.8 Hz, 1H, H$_1$), 2.30 – 2.24 (m, 1H, H$_6$), 1.89 (s, 1H, -NH), 1.71 (dt, *J* = 10.9, 7.1 Hz, 2H, H$_5$, H$_7$), 1.57 (dt, *J* = 22.9, 7.9 Hz, 2H, H$_5$, H$_7$), 1.45 (d, *J* = 6.6 Hz, 3H, H$_{2'}$); ^{13}C NMR (126 MHz, CDCl$_3$) δ 164.6 (C$_{2'''}$), 151.7 (C$_{4'''}$), 146.6 (C$_{6'''}$), 141.3 (C$_{1''}$), 134.0 (C$_{4a''}$), 131.4 (C$_{8a''}$), 129.0 (C$_{5''}$), 127.2 (C$_{4''}$), 125.7 (C$_{3''}$), 125.7 (C$_{2''}$), 125.3 (C$_{6''}$), 123.1 (C$_{7''}$), 122.9 (C$_{8''}$), 117.3 (C$_{3'''}$), 110.3 (C$_{4'''}$), 60.2 (C$_3$), 59.9 (C$_1$), 56.4 (C$_6$), 55.1 (C$_{1'}$), 54.3 (C$_4$), 53.4 (OCH$_3$), 42.9 (-NHCH$_2$), 35.6 (C$_5$), 35.4 (C$_7$), 30.2 (-CH$_2$NH), 23.6 (C$_{2'}$). HRMS (ESI) *m/z*: Calcd for C$_{26}$H$_{32}$N$_3$O [M + H]$^+$ 402.2539; Found 402.2530.

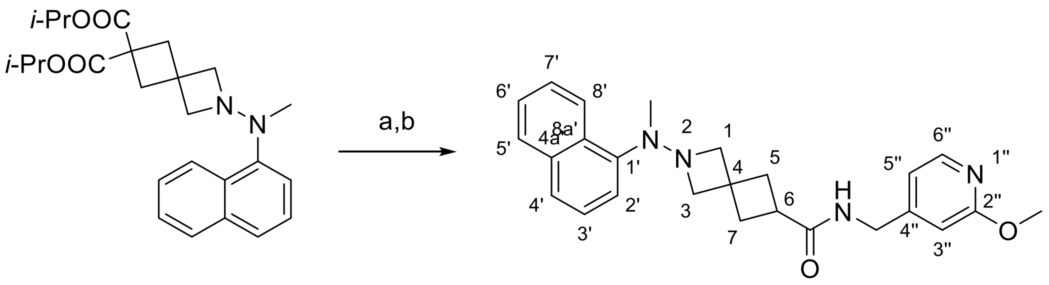
Diisopropyl *N*-[(naphthalen-1-yl)methylamino]-2-azaspiro[3.3]heptane-6,6-dicarboxylate (46).



Reagents and conditions: (a) Trifluoromethanesulfonic anhydride, DIPEA, CH$_3$CN, -20°C to 70°C, 2 hr.

To a stirred mixture of diisopropyl 3,3-bis(hydroxymethyl)cyclobutane-1,1-dicarboxylate[7] (167 mg, 0.58 mmol) and acetonitrile (1 mL) at -20°C, was carefully added trifluoromethanesulfonic anhydride (0.2 mL, 1.22 mmol) while maintaining the temperature below -10ºC. DIPEA (0.25 mL, 1.45 mmol) was then added slowly to the reaction mixture over 10 min. After formation of the bistriflate was complete (by mass spectral analysis), a second aliquot of DIPEA (0.25 mL, 1.45 mmol) was added dropwise followed by addition of *N*-methyl-*N*-(1-naphthyl)hydrazine **18** (100 mg, 0.58 mmol) and heating of the resulting mixture to 70ºC for 2 hr. The reaction mixture was cooled to room temperature and diluted with ethyl acetate (6 mL), washed with water (3×5 mL) and brine (3 mL). The combined organic layer was dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude residue was purified by silica gel column chromatography (0 to 3% methanol/dichloromethane) to yield the title product (185 mg, 75%) as a brown foam. ^1H NMR (500 MHz, $CDCl_3$) δ 8.34 (br s, 1H, H-8′), 8.00 (m, 2H, H-4′, H-5′), 7.79 (t, *J* = 7.8 Hz, 1H, H-2′), 7.62 (t, *J* = 7.6 Hz, 1H, H-3′), 7.54 (t, *J* = 8.1 Hz, 1H, H-7′), 7.37 (t, *J* = 8.2 Hz, 1H, H-6′), 5.05 (d, *J* = 11.7 Hz, 1H, H-1a), 4.94 (h, 6.3 Hz, 2H, $-CH(CH_3)_2$), 4.71 (d, *J* = 11.7 Hz, 1H, H-1b), 4.08 (s, 3H, $-CH_3$), 3.85 (dd, *J* = 12.3, 7.0 Hz, 1H, H-3a), 3.36 (dd, *J* = 12.3, 9.7 Hz, 1H, H-3b), 2.98 – 2.87 (m, 2H, H-5), 2.15 (d, *J* = 12.4 Hz, 2H, H-7), 1.19 – 1.10 (m, 9H, $-CH(CH_3)_2$), 1.07 (d, *J* = 6.3 Hz, 3H, $CH(CH_3)_2$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.5 (CO), 170.4 (CO), 140.2 (C-1′), 135.9 (C-4a′), 133.1 (C-4′), 130.6 (C-4′), 129.0 (C-2′), 127.2 (C-3′), 124.8 (C-7′), 123.6 (C-6′), 122.4 (C-8a′), 122.0 (C-4a′), 120.5 (C-8a′), 119.4, 79.5 (C-1), 69.6 ($-CH(CH_3)_2$), 58.4 (C-3), 57.1 (CH_3), 47.8 (C-4), 42.0 (C-6), 39.0 (C-7), 38.0 (C-5), 21.5 ($-CH(CH_3)_2$), 21.4 ($-CH(CH_3)_2$), 21.4 ($-CH(CH_3)_2$). HRMS (ESI) *m*/*z*: $[M + H]^+$ Calcd for $C_{25}H_{33}N_2O_4$ 425.2435; Found 425.2445.

_N_-[(2-Methoxypyridin-4-yl)methyl] _N'_-[(naphthalen-1-yl)methylamino]-2-azaspiro[3.3]heptane-6-carboxamide (47).

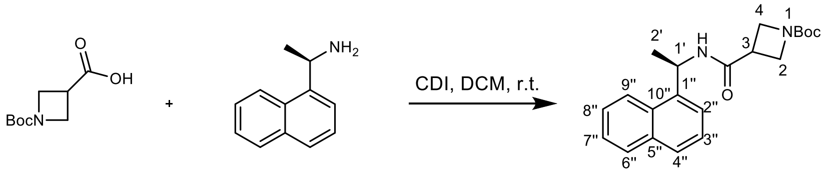


Reagents and conditions: (a) NaOH, H_3O^+, rt, 12 hr (b) (2-Methoxypyridin-4-yl)methanamine, CDI, THF, rt to 70°C, 2.5 hr.

Diisopropyl _N_-[(naphthalen-1-yl)methylamino]-2-azaspiro[3.3]heptane-6,6-dicarboxylate **46** (60 mg, 0.14 mmol) and methanol (0.2 mL), and 2 M aqueous NaOH (0.6 mL) were stirred at 20°C for 12 hr. The reaction mixture was then acidified by 3 N HCl to pH 3 and then extracted with diethyl ether (4×1 mL). The combined organic layer was dried over Na_2SO_4, filtered, and concentrated _in vacuo_. The crude mixture was dissolved in dry THF (1.6 mL) and treated with 1,1′-carbonyl diimidazole (30.5 mg, 0.19 mmol) in two portions. The resulting mixture was stirred at 20°C for 2 h, then was treated with (2-methoxypyridin-4-yl)methylamine (16.2 μL, 0.13 mmol) and the temperature increased to 70°C for 0.5 hr. After cooling to 20°C the mixture was concentrated _in vacuo_, dissolved in ethyl acetate (2 mL), washed with saturated aqueous $NaHCO_3$ (1 mL), dried over Na_2SO_4, filtered, concentrated under reduced pressure to give a residue, which was purified by column chromatography on neutral alumina (50 to 100% ethyl acetate/hexane) to furnish the title product (43 mg, 73%, 2 steps) as a colorless syrup. ^1H NMR (500 MHz, CD_3OD) δ 8.63 (br s, 1H, H-8′), 8.37 (d, J = 7.9 Hz, 1H, H-6″), 8.14 (d, J = 8.2 Hz, 1H, H-5′), 8.10 (d, J =

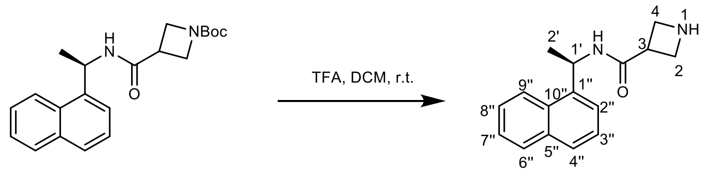
11.9 Hz, 1H, H-4′), 8.00 (dd, *J* = 5.4, 3.3 Hz, 1H, H-3′), 7.79 (q, *J* = 7.3 Hz, 1H, H-6′), 7.72 – 7.65 (m, 2H), 7.62 (td, *J* = 8.1, 3.3 Hz, 1H), 7.03 (s, 1H, H-3″), 6.81 (t, *J* = 5.1 Hz, 1H), 6.63 (s, 1H), 5.03 (d, *J* = 11.4 Hz, 1H, H-1a), 4.62 (t, *J* = 11.5 Hz, 1H, H-1b), 4.31 (s, 2H, -NH*CH₂*), 3.95 – 3.83 (m, 6H, -N*CH₃*, -O*CH₃*), 3.78 (dd, *J* = 17.7, 12.5 Hz, 1H, H-3a), 3.48 (dd, *J* = 21.5, 12.5 Hz, 1H, H-3b), 3.22 – 3.08 (m, 1H, H-6), 2.73 (td, *J* = 10.0, 9.1, 5.3 Hz, 1H, H-5), 2.68 – 2.59 (m, 1H, H-5), 2.41 – 1.99 (m, 2H, H-7). ^{13}C NMR (126 MHz, CD_3OD) δ 175.5 (-CONH), 164.8 (C-2″), 151.5 (C-4″), 146.5 (C-6″), 142.4 (C-1′), 136.1 (C-4a′), 132.7 (C-8a′), 130.0, 128.2, 126.9, 124.6, 123.0, 121.7 (C-8′), 119.9 (C-6″), 115.5, 108.3, 80.4 (C-1), 80.3, 58.0 (C-3), 57.6, 52.7 (-O*CH₃*), 41.6 (-NH*CH₂*), 35.6 (C-5), 33.6 (C-6), 33.2 (C-7). HRMS (ESI) *m*/z: $[M + H]^+$ Calcd for $C_{25}H_{29}N_4O_2$ 417.2285; Found 417.2292.

tert-Butyl (_R_)-3-[(1’-(Naphthalen-1’-yl)ethyl]carbamoyl)azetidine-1-carboxylate(48).[8]



A stirred solution of 1-(*tert*-butoxycarbonyl)azetidine-3-carboxylic acid[8] (1.57 g, 7.8 mmol) in anhydrous DCM (15 mL) was treated with 1,1’-carbonyl diimidazole (2.53 g, 15.6 mmol) and stirred for 1 hr at r.t. After such time, *(R)*-1-(1-naphthyl)ethylamine (3.75 mL, 23.4 mmol) was added and stirring continued for 1 hr at r.t. The reaction mixture was then filtered and the filtrate diluted with EtOAc (60 mL) and subsequently quenched *via* addition of H_2O (30 mL). The layers were separated and the aqueous layer re-extracted with EtOAc (20 mL, 3X). The organic layers were combined and dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue obtained was subjected to flash column chromatography on silica (eluent: 40:60 EtOAc:Hexanes) to afford the title compound (2.15 g, 6.07 mmol, 78%) as a yellow oil. $R_f = 0.70$ (60:40 EtOAc:Hexanes; CAM). $[\alpha]_D^{21} = +46.0$ ° (c = 1.0, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.04 (d, *J* = 8.3 Hz, 1H, H-9''), 7.88 – 7.82 (m, 1H, H-6''), 7.79 (d, *J* = 7.7 Hz, 1H, H-4''), 7.57 – 7.43 (overlapping m, 4H, H-8'', H-7'', H-2'', H-3''), 5.96 – 5.90 (m, 1H, H1'), 5.83 (s, 1H, NH), 4.13-3.89 (overlapping m, 4H, azetidine CH_2's), 3.07 (dt, *J* = 8.5, 6.0 Hz, 1H, azetidine CH), 1.68 – 1.63 (m, 3H, CH_3-H-2'), 1.41 (s, 9H, CH_3-Boc). ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.5 (C=O), 156.3 (C=O-Boc), 137.9 (C-1''), 134.1 (C10''), 131.2 (C5''), 129.0 (C4''), 128.7 (C6''), 126.8 (C8''), 126.1 (C7''), 125.3 (C3''), 123.4 (C9''), 122.7 (C2''), 79.9 (C-Boc), 44.9 (2C) (azetidine CH_2), 33.5 (C-1'), 28.5 (3C) (CH_3 (3)-Boc), 28.0 (azetidine CH), 20.7 (CH_3-C2'). HRMS (ESI) *m*/*z*: Calcd for $[C_{21}H_{26}O_3N_2Na]^+$ $[M+Na]^+$ 377.1836; Found 377.1832.

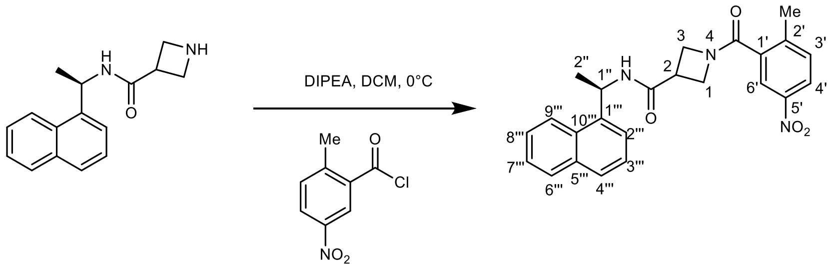
(*R*)-*N*-(1'-(Naphthalene-1'-yl)ethyl)azetidine-3-carboxamide (49).



A stirred solution of *tert*-butyl (*R*)-3-[(1-(naphthalen-1-yl)ethyl]carbamoyl)azetidine-1-carboxylate **48** (2.1 g, 5.9 mmol) in anhydrous DCM (9 mL) was treated with trifluoroacetic acid (TFA) (9 mL, 118 mmol) and stirred at r.t. for 45 min. After such time, the reaction mixture was washed with $NaHCO_3$ (15 mL, 1X) and the organic layers concentrated *in vacuo*. The solution was then co-concentrated with toluene (20 mL, 3X) to afford the title compound (1.05 g, 4.13 mmol, 70%) as a white solid. $R_f = 0.05$ (1:9

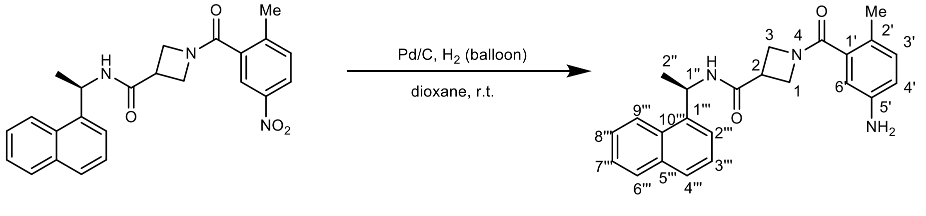
MeOH:DCM; CAM). $[\alpha]_D^{21}$ = +32.0 ° (c = 1.0, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.74 (s, 1H, NH), 7.90 (dd, *J* = 8.5, 3.2 Hz, 1H, H-9’’), 7.78 (dd, *J* = 8.2, 3.1 Hz, 1H, H-6’’), 7.72 (dd, *J* = 8.4, 3.2 Hz, 1H, H-4’’), 7.51 – 7.34 (m, 4H, H-8’’, H-7’’, H-2’’, H-3’’), 7.23 (dd, *J* = 7.3, 3.4 Hz, 1H, NH), 5.65 (dt, *J* = 9.6, 6.3 Hz, 1H, H-1’), 4.21 (s, 1H, azetidine CH_2), 4.07 (s, 1H, azetidine CH_2), 3.82 (d, *J* = 31.8 Hz, 2H, azetidine CH_2), 3.64 (q, *J* = 8.2 Hz, 1H, azetidine CH), 1.55 (dd, *J* = 7.3, 3.8 Hz, 3H, CH_3-H-2’). ^{13}C NMR (126 MHz, $CDCl_3$): δ 169.9 (C=O), 138.0 (C1’’), 133.9 (C10’’), 130.4 (C5’’), 129.3 (C4’’), 128.5 (C6’’), 126.9 (C8’’), 126.2 (C7’’), 125.7 (C3’’), 122.3 (2C) (C9’’, C2’’), 48.9 (azetidine CH_2), 46.4 (azetidine CH_2), 35.9 (C1’), 27.7 (azetidine CH), 21.1 (CH_3-C2’). HRMS (ESI) *m*/*z*: Calcd for $[C_{16}H_{19}ON_2]^+$ $[M+H]^+$ 255.1492; Found 255.1488.

(*R*)-1’-(2’-Methyl-5’-nitrobenzoyl)-*N*-(1’’-(naphthalene-1’’-yl)ethyl)azetidine-3-carboxamide (50).



A stirred solution of (*R)*-*N*-(1’-(naphthalene-1’-yl)ethyl)azetidine-3-carboxamide **49** (700 mg, 2.75 mmol) in anhydrous DCM (8 mL) was treated with DIPEA (1.9 mL, 10 mmol) at 0 ° and stirred for 10 min. After such time, 2-methyl-5-nitrobenzoyl chloride (547 mg, 2.75 mmol) was added portion-wise and the solution was stirred for 20 min. After such time, the reaction mixture was filtered, and the solids washed with DCM (10 mL, 3X). The filtrate was then concentrated *in vacuo* and the residue obtained was subjected to flash column chromatography on silica (1:9 MeOH:DCM) to afford the title compound (688 mg, 1.65 mmol, 60%) as an orange solid. $R_f = 0.30$ (1:9 MeOH:DCM; CAM). $[\alpha]_D^{21} = +69.0$ ° (c = 1.0, $CHCl_3$). ^1H NMR (500 MHz, CD_3OD): δ 8.21 – 8.13 (m, 2H, H-6', H-4'), 8.07 (dd, *J* = 12.1, 8.4 Hz, 1H, H-9'''), 7.83 (dd, *J* = 8.0, 1.5 Hz, 1H, H-6'''), 7.76 (dd, *J* = 17.2, 8.2 Hz, 1H, H-4'''), 7.54-7.38 (m, 5H, H-2''', H-3''', H-7''', H-8''', H-3'), 5.83 (qd, *J* = 7.0, 2.3 Hz, 1H, H-1''), 4.37 (m, 1.5, azetidine CH_2), 4.14 (dd, *J* = 5 Hz, 0.5H, azetidine CH_2), 4.10 (d, *J* = 7.2 Hz, 1H, azetidine CH_2), 4.04 (t, *J* = 8.8 Hz, 0.5 H, azetidine CH_2), 3.94 (dd, *J* = 9.0, 5.6 Hz, 0.5 H, azetidine CH_2), 3.45 (dddd, *J* = 11.6, 8.6, 5.6, 3.6 Hz, 1H, azetidine CH), 2.44 (s, 3H, CH_3-aryl), 1.58 (dd, *J* = 11.7, 6.9 Hz, 3H, CH_3-H-2'').* ^{13}C NMR (126 MHz, CD_3OD): δ 172.7 (C(O)NR_2), 170.6 (C(O)NHR), 147.5 (C5'), 144.9 (C2'), 140.0 (C1'), 135.9 (C-1'''), 135.5 (C10'''), 133.2 (naphthalene C), 132.2 (C5'''), 129.9 (C6'''), 129.1 (C4'''), 127.3 (naphthalene C), 126.8 (naphthalene C), 126.4 (naphthalene C), 125.5 (C6'), 124.0 (C9'''), 123.5 (naphthalene C), 122.8 (C4'), 54.8 (azetidine CH_2), 52.3 (azetidine CH_2), 46.2 (C1''), 34.1 (azetidine CH), 21.3 (C2''), 19.5 (CH_3-aryl).* HRMS (ESI) *m/z*: Calcd for $[C_{24}H_{23}O_4N_3Na]^+$ $[M+Na]^+$ 440.1581; Found 440.1571. *Where doubling of ^{13}C resonances is seen in the azetidine only one signal is reported.[9, 10]

(*R*)-1'-(5'-Amino-2'-methylbenzoyl)-*N*-(1''-(naphthalene-1''-yl)ethyl)azetidine-3-caboxamide (51).

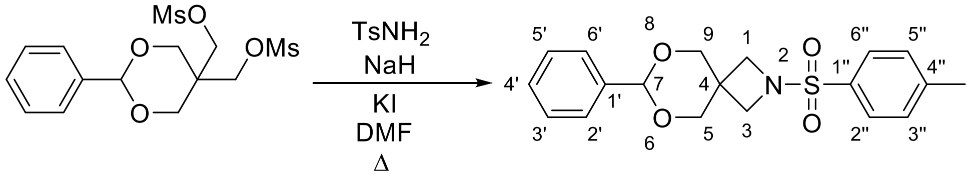


To a stirred solution of (*R*)-1'-(2'-methyl-5'-nitrobenzoyl)-*N*-(1''-(naphthalene-1''-yl)ethyl)azetidine-3-carboxamide **50** (120 mg, 0.29 mmol) in dioxane (6 mL) was added

Pd/C (10% wt) (12 mg, 0.11 mmol). The vessel was purged with H_2 (3X), after which, H_2 was allowed to continuously flow into the stirred vessel for 12 hr. After such time, the reaction mixture was filtered over Celite and washed with MeOH (10 mL). The filtrate was concentrated *in vacuo* and the residue obtained was subjected to flash column chromatography on silica (eluent: 10:90 MeOH:DCM – 100:0 MeOH:DCM) to afford the title compound (92 mg, 0.24 mmol, 82%) as a pink solid. $R_f = 0.05$ (1:9 MeOH:DCM). $[\alpha]_D^{21} = +60.0$ ° (c = 1.0, $CHCl_3$). ^1H NMR (500 MHz, CD_3OD): δ 8.07 (t, *J* = 10.0 Hz, 1H, H-9'''), 7.91 – 7.83 (m, 1H, H-6'''), 7.78 (dd, *J* = 11.3, 8.2 Hz, 1H, H-4'''), 7.56-7.39 (m, 4H, H-2''', H-3''', H-8''', H-7'''), 6.97 (t, *J* = 8.8 Hz, 1H, H-3'), 6.68 (tt, *J* = 8.1, 2.4 Hz, 1H, H-4'), 6.59 (dd, *J* = 16.6, 2.3 Hz, 1H, H-6'), 5.83 (q, *J* = 7.0 Hz, 1H, H-1''), 4.23 (dd, *J* = 30.3, 8.3 Hz, 1.5H, azetidine CH_2), 4.11 (t, *J* = 6.2 Hz, 1H, azetidine CH_2), 4.04 (t, *J* = 8.9 Hz, 0.5H, azetidine CH_2), 3.95 (d, *J* = 7.2 Hz, 1H, azetidine CH_2), 3.40 (q, *J* = 7.5, 6.7 Hz, 1H, azetidine CH), 2.18 (dd, *J* = 12.4, 2.0 Hz, 3H, CH_3-aryl), 1.62 – 1.55 (m, 3H, CH_3-H2'').* ^{13}C NMR (126 MHz, CD_3OD): δ 173.9 (C(O)NR_2), 172.8 (C(O)NHR), 146.9 (C5'), 140.1 (C1'), 135.5 (C-1'''), 134.8 (C10'''), 132.6 (C3'), 132.2 (C-5'''), 129.9 (C6'''), 129.0 (C4'''), 127.3 (naphthalene C), 126.8 (naphthalene C), 126.4 (naphthalene C), 124.9 (C2'), 124.0 (C9'''), 123.5 (naphthalene C), 118.2 (C4'), 114.1 (C6'), 54.8 (azetidine CH_2), 52.0 (azetidine CH_2), 46.2 (C1''), 33.9 (azetidine CH), 21.4 (C2''), 18.2 (CH_3-aryl).* HRMS (ESI) *m/z*: Calcd for $[C_{24}H_{26}O_2N_3]^+$ $[M+H]^+$ 388.2019; Found 388.2015.

*As a result of the dynamic conformational assembly about the azetidine ring due to pyramidalization, doubling of ^{13}C resonances and splitting of azetidine protons was observed. Where doubling of ^{13}C resonances appear, one signal is reported.[9, 10]

7-Phenyl-2-tosyl-6,8-dioxa-2-azaspiro[3.5]nonane (53).

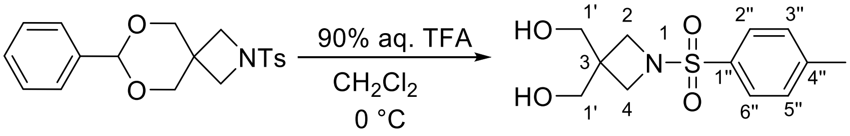


A stirred suspension of NaH (60 % dispersion in mineral oil) (262 mg, 6.57 mmol) in anhydrous DMF (20 mL) was treated with KI (43.6 mg, 0.263 mmol) followed portion-wise by $TsNH_2$ (495 mg, 2.89 mmol) and then was stirred for 0.5 hr at 20°C. 2-Phenyl-1,3-dioxane-5,5-diyl)bis(methylene) dimethanesulfonate **52**[11] (1g, 2.63 mmol) was then added and the reaction mixture stirred overnight (16 h) at 140°C. After cooling to room temperature, the reaction then was quenched with brine (20 mL) and the reaction mixture was extracted with Et_2O (3×30 mL). The combined organic layers were washed with brine (60 mL), dried over $MgSO_4$, and concentrated to dryness. The crude product was purified by flash column chromatography over silica gel (eluent: hexanes : EtOAc (EtOAc, 0 → 30 %)) to give the title compound as a white solid (550 mg, 58 %). R_f 0.36 (hexanes : EtOAc 7 : 3 (v/v); UV). ^1H NMR (500 MHz, DMSO-*d6*) δ 7.76 (d, *J* = 8.2 Hz, 2H, H-6'', H-2''), 7.52 (d, *J* = 8.0 Hz, 2H, H-3'', H-5''), 7.33 (s, 5H, $\underline{C_6H_5}CH$), 5.38 (s, 1H, $C_6H_5C\underline{H}$), 3.74 (d, *J* = 8.8 Hz, 4H, 2×H-9, 2×H-5), 3.66 (d, *J* = 11.2 Hz, 2H, H-1, H-3), 3.36 (s, 2H, H-1, H-3)[1], 2.44 (s, 3H, CH_3). ^{13}C NMR (126 MHz, DMSO-*d6*) δ 144.8 (C-4''), 138.4 (C-1'), 131.3 (C-1''), 130.6 (C-3'', C-5''), 129.3 (C-5', C-3'), 128.8 (C-4'), 128.5 (C-6'', C-2''), 126.6 (C-6', C-2'), 100.7 ($Ph\underline{C}H$), 72.5 (C-9), 60.2 (C-3), 53.8 (C-1), 32.1 (C-4), 21.6 (CH_3). HRMS (ESI) *m/z*: Calcd for $C_{19}H_{21}NNaO_4S$ $[M+Na]^+$ 382.1089; Found 382.1072.

---

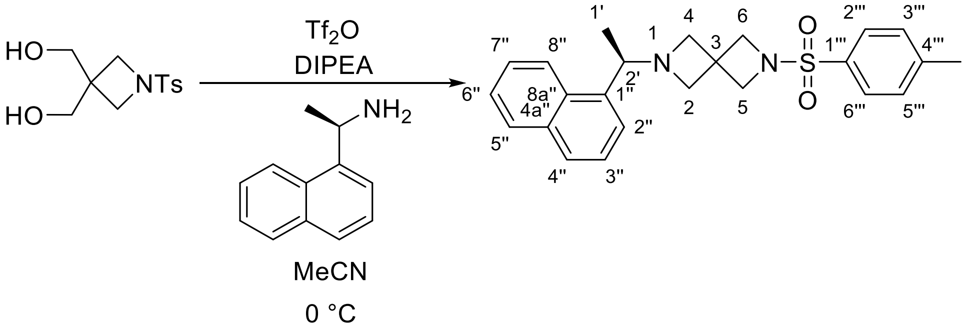
[1] Peak overlaps with residual signal of DMSO-*d6*

(1-Tosylazetidine-3,3-diyl)dimethanol (54).



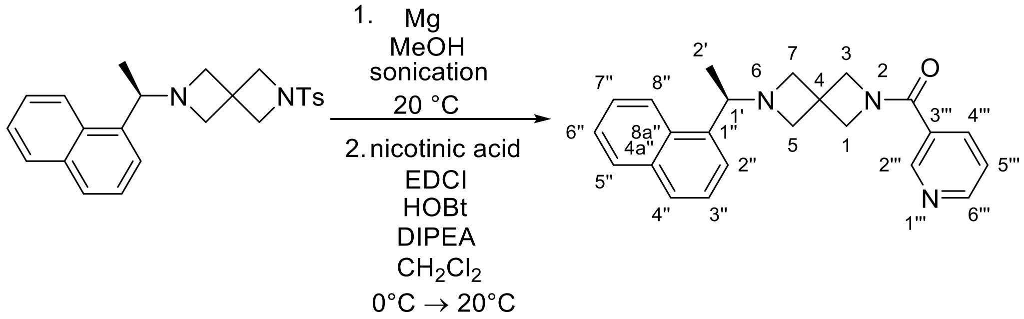
7-Phenyl-2-tosyl-6,8-dioxa-2-azaspiro[3.5]nonane **53** (500 mg, 1.39 mmol) was dissolved in CH_2Cl_2 (20 mL) and the reaction mixture was cooled down to 0°C before TFA (90% aq. solution) (4 mL) was added and the reaction mixture was stirred for 1.5 hr at 0°C. After completion of the reaction, detected by LCMS, the reaction mixture was concentrated to dryness and the crude product was purified by flash column chromatography over silica gel (eluents: hexanes : EtOAc (EtOAc, 0 → 30 %); CH_2Cl_2 : MeOH (MeOH 0 → 20 %)) to give the title compound as a colorless syrup (270 mg, 71 %). R_f 0 (hexanes : EtOAc 7 : 3 (v/v); CAM). ^{1}H NMR (500 MHz, CD_3OD) δ 7.77 – 7.71 (m, 1H, H-2''), 7.71 – 7.65 (m, 1H, H-6''), 7.43 (d, J = 8.0 Hz, 1H, H-3''), 7.30 (d, J = 8.1 Hz, 1H, H-5''), 3.49 (s, 2H, H-1'), 3.38 (s, 2H, H-1'), 2.43 (s, 2H, H-2), 2.37 (s, 2H, H-4).^{13}C NMR (126 MHz, CD_3OD) δ 144.4 (C-1''), 131.6 (C-4''), 129.7 (C-3'', C-5''), 128.3 (C-2'', C-4''), 62.8 (C-4), 54.2 (C-2), 39.7 (C-3), 35.7 (C-1'), 30.4 (C-1'), 20.3 (CH_3). HRMS (ESI) *m*/*z*: Calcd for $C_{12}H_{17}NNaO_4S$ $[M+Na]^+$ 294.0776; Found 294.0763.

(*R*)-2-(1-(Naphthalen-1-yl)ethyl)-6-tosyl-2,6-diazaspiro[3.3]heptane (55).



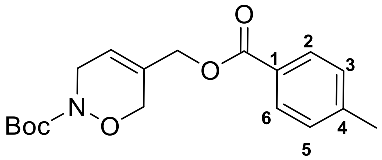
(1-Tosylazetidine-3,3-diyl)dimethanol **54** (250 mg, 0.921 mmol) was dissolved in anhydrous MeCN (3 mL) and the reaction mixture was cooled down to 0°C with stirring. Subsequently, Tf_2O (324 µL, 1.93 mmol) was added followed dropwise by DIPEA (400 µL, 2,39 mmol). After formation of the ditriflate (15 min), detected by LCMS, further DIPEA (400 µL, 2,39 mmol) was added dropwise followed by slow addition of (*R*)-1-(naphthalen-1-yl)ethan-1-amine (147 µL, 0.921 mmol). The reaction mixture then was stirred for 10 min at 0°C and then refluxed for 3 hr. After completion, detected by LCMS, the reaction mixture was quenched with sat. aq. $NaHCO_3$ (10 mL) and extracted with EtOAc (3×20 mL). The combined organic layers were washed with brine (50 mL), dried over $MgSO_4$, and concentrated to dryness. The crude product was purified by flash column chromatography over silica gel (eluent: hexanes : EtOAc (EtOAc, 0 → 30 %)) to give the title compound as white solid 260 mg (70 %). R_f 0.56 (hexanes : EtOAc 7 : 3 (v/v); UV). $[\alpha]_D^{20\text{-}22}$ – 9.6 (*c* 7.5, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.64 – 8.58 (m, 1H, H-8’’), 8.09 – 7.94 (m, 1H, H-5’’), 7.88 (d, *J* = 8.0 Hz, 1H, H-4’’), 7.81 (d, *J* = 8.2 Hz, 1H, H-7’’), 7.68 – 7.60 (m, 2H, H-2’’’, H-6’’’), 7.58 – 7.44 (m, 2H, H-3’’, H-2’’), 7.40 – 7.28 (m, 3H, H-3’’’, H-5’’’, H-6’’), 3.25 (s, 2H, H-5, H-6), 3.16 (s, 2H, H-5, H-6), 2.95 (d, *J* = 1.1 Hz, 2H, H-4, H-2), 2.87 (d, *J* = 1.1 Hz, 2H, H-4, H-2), 2.40 (s, CH_3), 1.50 (d, *J* = 6.5 Hz, 3H, H-1’), 1.35-1.31 (m, 1H, H-2’). ^{13}C NMR (126 MHz, $CDCl_3$) δ 149.1 (C-1’’’), 144.5 (C-4’’’), 133.4 (C-4a’’), 131.1 (C-8a’’), 130.8 (C-1’’), 129.9 (C-3’’’, C-5’’’), 129.6 (C-4’’), 129.4 (C-5’’), 128.4 (C-2’’’, C-6’’’), 128.2 (C-7’’), 126.0 (C-6’’), 125.9 (C-3’’), 124.0 (C-8’’), 121.8 (C-2’’), 61.7 (C-4, C-2), 59.7 (C-2’), 45.9 (C-6,C-5), 31.8 (C-3), 21.6 (C-2’), 20.7 (C-1’). HRMS (ESI) *m*/*z*: Calcd for $C_{24}H_{27}N_2O_2S^+$ $[M+H]^+$ 407.1788; Found 407.1775.

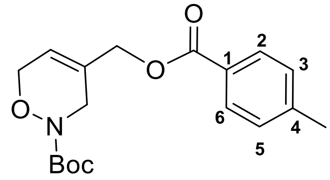
(*R*)-(6-(1-(Naphthalen-1-yl)ethyl)-2,6-diazaspiro[3.3]heptan-2-yl)(pyridin-3-yl)methanone (56).



A stirred solution of (*R*)-2-(1-(naphthalen-1-yl)ethyl)-6-tosyl-2,6-diazaspiro[3.3]heptane **55** (110 mg, 0.271 mmol) in MeOH (10 mL) was treated with Mg turnings (52.5 mg, 2.17 mmol) and sonicated (40 kHz) for 1.5 hr at 20°C. After completion, detected by LCMS, the reaction mixture was diluted with Et_2O (20 mL) and filtered on Celite. The filter cake was additionally washed with Et_2O (3×20 mL), and the filtrate was dried over $MgSO_4$ and concentrated to dryness. The residue (80 mg, 0.317 mmol) was taken up in anhydrous CH_2Cl_2 (2 mL) and added at 0°C to a stirred solution of nicotinic acid (39 mg, 0.317 mmol), EDCI (76 mg, 0.396 mmol) and HOBt (53.5 mg, 0.396 mmol) in anhydrous CH_2Cl_2 (4 mL), followed by addition of DIPEA (220 μL, 1.26 mmol). The reaction mixture then was allowed to warm up to 20°C and was stirred for 16 hr before it was quenched with H_2O (5 mL) and extracted with CH_2Cl_2 (3×10 mL). The combined organic layers were dried over $MgSO_4$ and concentrated to dryness. The crude material was purified by flash column chromatography over silica gel (eluents: hexanes : EtOAc (EtOAc, 0 → 30 %); CH_2Cl_2 : MeOH (MeOH 0 → 10 %)) to give the title compound as a colorless syrup (35 mg, 36 % over 2 steps). R_f 0.50 (CH_2Cl_2 : MeOH 99 : 1 (v/v); UV). $[\alpha]_D^{20-22}$ + 4.5 (*c* 5, MeOH). ^1H NMR (500 MHz, CD_3OD) δ 8.80 (dd, *J* = 2.2, 0.9 Hz, 1H, H-2'''), 8.66 (dd, *J* = 4.9, 1.6 Hz, 1H, H-6'''), 8.24 – 8.18 (m, 1H, H-4'''), 8.08 (dt, *J* = 7.9, 1.9 Hz, 1H, H-8''), 7.86 (dd, *J* = 8.2, 1.5 Hz, 1H, H-4''), 7.75 (d, *J* = 8.1 Hz, 1H, H-5''), 7.59 – 7.41 (m, 5H, H-2'', H-7'', H-5''', H-3'', H-6''), 4.52 – 4.44 (m, 1H, 2×H-3'), 4.26 (d, *J* = 13.9 Hz, 2H, 2×H-1), 3.49 – 3.39 (m, 4H, 2×H-5, 2×H-7), 2.01 (s, 1H, H-1'), 1.36 – 1.27 (m, 3H, H-2'). ^{13}C NMR (126 MHz, CD_3OD) δ 179.0 (CO), 167.8 (C-6'''), 151.1 (C-2'''), 148.1 (C-3'''), 138.2 (C-4a''), 136.1 (C-8a''), 134.2 (C-1''), 131.2 (C-4'''), 129.4 (C-4''), 128.7 (C-5''), 127.3 (C-7''), 125.7 (C-6''), 125.3 (C-3''), 125.2 (C-5'''), 123.8 (C-8''), 122.7 (C-2''), 62.8 (C-1), 62.7 (C-3), 58.8 (C-5, C-7), 32.9 (C-4), 22.9 (H-1'), 19.4 (H-2'). HRMS (ESI) *m*/*z*: Calcd for $C_{23}H_{24}N_3O$ $[M+H]^+$ 358.1914; Found 358.1902.

(2-*tert*-Butyloxycarbonyl)-3,6-dihydro-2*H*-1,2-oxazin-4-methanyl 4-methylbenzoate (60) and (2-*tert*-Butyloxycarbonyl)-3,6-dihydro-2*H*-1,2-oxazin-5-methanyl 4-methylbenzoate (61).





A mixture of 2-(*tert*-butyldimethylsilylmethylene-3-buten-1-ol) **57**[12] (0.080 g, 0.4 mmol) and *p*-toluenesulfonic acid (13.8 mg, 0.08 mmol) were dissolved in moist acetonitrile (2 mL) and heated to reflux until completion (1 hr). After cooling to room temperature, saturated aqueous $NaHCO_3$ was added dropwise until the effervescence stopped. The resultant suspension was diluted with DCM (3 mL) and extracted with saturated aqueous $NaHCO_3$ (3 mL). The organic layer was separated, dried on Na_2SO_4 and filtered. The filtrate was then immediately transferred into a reaction flask containing *N*-Boc-hydroxylamine (0.079 g, 0.60 mmol) and the resulting stirred solution was cooled to 0ºC before a solution of Bu_4NIO_4 (0.2 g, 0.047 mmol) in DCM (1 mL) was added dropwise over 1 hr at 0ºC. The reaction mixture was allowed to stir at 0ºC for 2 hr and then at room temperature for an additional 2 hr before saturated aqueous $Na_2S_2O_3$ (3 mL) was added. The aqueous phase was extracted with DCM (2 × 5 mL), the combined organic phases were washed with brine (10 mL), dried over Na_2SO_4, filtered, and concentrated, and the residue was purified by column chromatography over silica gel with EtOAc:Hexane (30:70) as the eluent to give an inseparable (1:6.3) mixture of two regiosomeric oxazines **58** and **59** (0.045 g, 48 % combined yield). HRMS (ESI) (*m*/*z*): Calcd for $C_{10}H_{17}NO_4Na$

$[M+Na]^+$ 238.1049; Found 238.1044. The major isomer was 2-*tert*-butyloxycarbonyl-5-(hydroxymethyl)-3,6-dihydro-2*H*-1,2-oxazine **59** and had ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.85 – 5.78 (m, 1H, C*H*=C), 4.45 (q, *J* = 2.4 Hz, 2H, C*H*$_2$-O-N), 4.11 (d, *J* = 3.4 Hz, 4H, C*H*$_2$-OH, C*H*$_2$-N), 1.50 (s, 9H, (C(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.8 (N-*C*O), 135.7 (*C*=CH), 118.2 (*C*H=C), 81.7 (*C*(CH_3)$_3$), 68.5 (*C*H_2-N-O), 63.1 (*C*H_2-OH), 44.6 (O-N-*C*H_2), 28.3 (*C*H_3)$_3$.

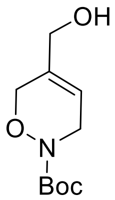
The above (6.3:1) mixture of regioisomeric oxazines **58** and **59** (40 mg, 0.18 mmol) and pyridine (29 mg, 0.37 mmol) was dissolved in DCM (2 mL), cooled to 0°C and a solution of *p*-toluoyl chloride (31 mg, 0.20 mmol) in DCM (1 mL) was added dropwise with stirring over 5 min. The reaction mixture was stirred at room temperature for 1 hr and then was diluted with 5 mL of DCM and washed with saturated aqueous $NaHCO_3$, brine and H_2O. The organic layer was dried on Na_2SO_4 and evaporated under reduced pressure and the residue was subjected to column purification over silica gel eluting with EtOAc:hexane (20:80) to afford (2-*tert*-butyloxycarbonyl)-3,6-dihydro-2*H*-1,2-oxazin-4-methanyl 4-methylbenzoate **60** (7 mg, 11%) and (2-*tert*-butyloxycarbonyl)-3,6-dihydro-2*H*-1,2-oxazin-5-methanyl 4-methylbenzoate **61** (49 mg, 82%) both as colorless oils.

(2-*tert*-Butyloxycarbonyl)-3,6-dihydro-2*H*-1,2-oxazin-4-methanol 4-methylbenzoate 60: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92 (d, *J* = 8.2 Hz, 2H, H2,H6), 7.23 (d, *J* = 8.0 Hz, 2H, H3, H5), 5.99 – 5.94 (m, 1H, C*H*=C)), 4.78 (d, *J* = 1.8 Hz, 2H, C*H*$_2$-O-N), 4.45 (dq, *J* = 2.6, 1.3 Hz, 2H, C*H*$_2$-N-O), 4.15 (q, *J* = 2.3 Hz, 2H, C*H*$_2$-O-CO-Tol), 2.40 (s, 3H, (CH_3-Tol), 1.48 (s, 9H, C(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.3 (O-*C*O), 155.0 (N-*C*O), 144.1 (C4), 130.2 (C1), 129.8 (C2), 129.3, 127.1 (C3), 123.0, 82.0 (*C*(CH_3)$_3$), 67.7 (*C*H_2-

O-N), 65.1 (*C*H$_2$-O-CO), 46.0 (*C*H$_2$-N-O), 29.8, 28.4 (*C*H$_3$-C4), 21.8 (C(*C*H$_3$)$_3$; HRMS (ESI) (*m*/*z*): Calcd for $C_{18}H_{23}NO_5Na$ [M+Na]$^+$ 356.1468; Found 356.1465.

(2-*tert*-butyloxycarbonyl)-3,6-dihydro-2*H*-1,2-oxazin-5-methanol 4-methylbenzoate 61: ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.90 (dd, *J* = 7.9, 1.7 Hz, 2H, H2,H6), 7.29 – 7.17 (m, 2H, H3, H5), 5.95 (td, *J* = 3.3, 1.7 Hz, 1H, C*H*=C), 4.75 (s, 2H, C*H*$_2$-O-N), 4.48 (t, *J* = 2.1 Hz, 2H, C*H*$_2$-O-CO), 4.16 – 4.07 (m, 2H, C*H*$_2$-N-O), 2.40 (d, *J* = 1.4 Hz, 3H, C*H*$_3$-C4), 1.49 (d, *J* = 1.7 Hz, 9H, C(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.2 (O-*C*O), 154.9 (N-*C*O), 143.9 (C4), 131.5 (C1), 129.6 (C2), 129.1, 126.9 (C3), 121.6, 81.8 (*C*(CH$_3$)$_3$), 68.6 (*C*H$_2$-O-N), 64.3 (*C*H$_2$-O-CO), 44.7 (*C*H$_2$-N-O), 28.2 (*C*H$_3$-C4), 21.6 (C(*C*H$_3$)$_3$; HRMS (ESI) (*m*/*z*): Calcd for $C_{18}H_{23}NO_5Na$ [M+Na]$^+$ 356.1468; Found 356.1456.

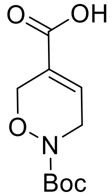
***tert*-Butyl 5-(hydroxymethyl)-3,6-dihydro-2H-1,2-oxazine-2-carboxylate (59).**



A solution of **61** (0.4 g, 1.2 mmol) in MeOH (2.5 mL), was treated with K_2CO_3 (3.3 mg, 0.024 mmol), and stirred at room temperature until completion (4 hr). The solvents were evaporated under vacuum and the residue was dissolved in DCM (5 mL), filtered and concentrated under reduced pressure. The residue was purified by column chromatography over silica gel eluting with EtOAc:Hexane (40:60) to give the title compound (0.220 g, 85%) as an oil. ^{1}H NMR (600 MHz, CDCl$_3$) δ 5.79 (s, 1H, C*H*=C), 4.43 (d, *J* = 2.2 Hz, 2H, C*H*$_2$O), 4.08 (s, 4H, C*H*$_2$-N, C*H*$_2$OH) 1.49 (s, 9H, C(C*H*$_3$)$_3$); ^{13}C NMR (151 MHz, CDCl$_3$) δ 155.0 (CO), 135.8 (*C*=CH), 117.9 (C=*C*H), 81.9 (*C*(CH$_3$)$_3$), 68.7 (*C*H$_2$-OH), 62.9 (*C*H$_2$-

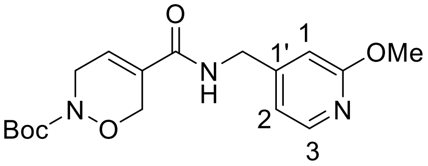
O), 44.6 (*C*H$_2$-N), 28.3 (C(*C*H$_3$)$_3$); HRMS (ESI) (*m*/*z*): Calcd for $C_{10}H_{17}NO_4Na$ [M+Na]$^+$ 238.1049; Found: 238.1046.

2-(*tert*-Butoxycarbonyl)-3,6-dihydro-2H-1,2-oxazine-5-carboxylic acid (62).



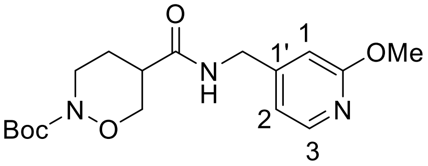
A solution of **59** (0.120 g, 0.55 mmol) in DCM (5 mL) was treated with Dess-Martin periodinane (0.354 g, 0.83 mmol) and stirred at room temperature for 2 hr before a 1:1 solution (12 mL) of saturated aqueous $NaHCO_3$ and saturated aqueous $Na_2S_2O_3$ was added. The resulting suspension was diluted with DCM (10 mL) and stirred for 30 min before the organic layer was separated, dried on Na_2SO_4 and concentrated under vacuum to afford a crude aldehyde, which was dissolved in a mixture of THF/*t*-BuOH (1:1, 3 mL) and cooled to 0°C. Water (1.5 mL) and 2-methyl-2-butene (1 mL) were added, followed by NaH_2PO_4 (0.472 g, 3.9 mmol) and $NaClO_2$ (0.5 g, 5.6 mmol) and the mixture was stirred at 0°C for 2 hr before it was diluted with EtOAc (30 mL) and acidified with HCl (1N, 10 mL). The organic phase was washed with brine, dried over Na_2SO_4, and purified over silica gel eluting with 2:98 MeOH:DCM to afford the title compound (0.105 g, 82%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.12 (t, *J* = 1.8 Hz, 1H, C*H*=C), 4.61 (d, *J* = 2.2 Hz, 2H, C*H*$_2$O), 4.27 (d, *J* = 3.0 Hz, 2H, C*H*$_2$-N), 1.51 (s, 9H, C(C*H*$_3$)$_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ 167.7 (*C*OOH), 154.6 (CO), 136.4 (*C*=CH), 127.9 (C=*C*H), 82.4 (*C*(CH$_3$)$_3$), 67.2 (*C*H$_2$-O), 45.1 (*C*H$_2$-N), 28.3 (C(*C*H$_3$)$_3$); HRMS (ESI) (*m*/*z*): Calcd for $C_{10}H_{15}NO_5Na$ [M+Na]$^+$ 252.0842; Found: 252.0839.

_N_-((2-Methoxypyridin-4-yl)methyl) 2-_tert_-butyloxycarbonyl-3,6-dihydro-2H-1,2-oxazine-2-carboxamide (63).



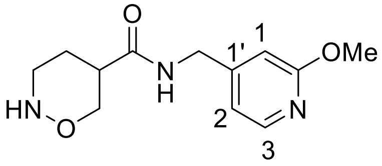
A solution of **62** (0.1 g, 0.4 mmol) in DCM (2 mL) was treated with, _N_-(3-dimethylaminopropyl)-_N_-ethylcarbodiimide hydrochloride (EDC) (0.108 g, 0.56 mmol), and 1-hydroxybenzotriazole hydrate (HOBT) (0.076 g, 0.56 mmol). The suspension was stirred for 5 min at room temperature then cooled to 0°C and treated dropwise with a mixture of 2-methoxypyrdin-4-yl methylamine (0.078 g, 0.56 mmol) and diisopropylethylamine (0.281 g, 2.1 mmol) in DCM (1 mL). The reaction mixture was then allowed to warm at room temperature and stirred for additional 16 hr before it was quenched with water (10 mL) and extracted with DCM (2 x 10 mL). The organic layers were separated, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by silica gel column chromatography eluting with EtOAc:Hexane (35:65) to afford the title amide as an oil (0.105 g, 69%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.08 (d, _J_ = 5.4 Hz, 1H, _H3_), 6.76 (dd, _J_ = 5.3, 1.5 Hz, 1H, _H2_), 6.63 (t, _J_ = 5.9 Hz, 1H, N_H_CO), 6.62 – 6.60 (m, 1H, _H1_), 6.57 (dq, _J_ = 3.4, 1.7 Hz, 1H, N_H_), 4.67 – 4.63 (m, 2H, CH_2-O), 4.43 (d, _J_ = 6.1 Hz, 2H, CH_2-OH), 4.18 (d, _J_ = 3.0 Hz, 2H, CH_2-N), 3.91 (s, 3H, OMe), 1.49 (s, 9H, C(CH_3)$_3$)); ^{13}C NMR (151 MHz, $CDCl_3$) δ 164.9 (NH_C_O), 164.6 (_C_-OMe), 154.7 (_C_O), 149.9 (C3), 147.2 (C1’), 132.0 (_C_=CH), 127.1 (_C_H=C), 115.6 (C2), 108.9(C1), 82.3 (_C_(CH$_3$)$_3$), 67.6 (_C_H$_2$-O), 53.5 (C-O_Me_), 44.4 (_C_H$_2$-N), 42.2 (_C_H$_2$-NHCO), 28.2 (C(_C_H$_3$)$_3$); HRMS (ESI) (_m/z_): Calcd for $C_{17}H_{24}N_3O_5$ $[M+H]^+$ 350.1704; Found: 350.1701.

N-((2-Methoxypyridin-4-yl)methyl) 2-*tert*-butyloxycarbonyl-2H-1,2-oxazinane-2-carboxamide (64).



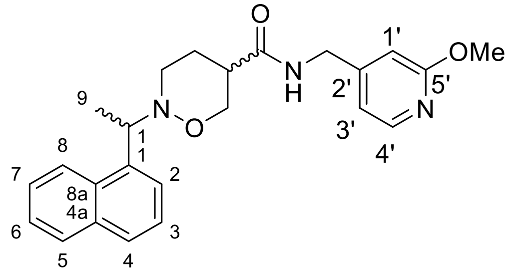
A solution of **63** (0.080 g, 0.22 mmol) in MeOH (2 mL) was treated with 10% Pd(C) (7 mg) and stirred under a hydrogen atmosphere until completion (1.5 hr). The reaction mixture was filtered through a Celite® pad and the filtrate was concentrated under reduced pressure, and the residue purified by silica gel column chromatography eluting with EtOAc:Hexane (1:1) to furnish the title compound as an oil (0.074 g, 92%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.06 (d, *J* = 5.3 Hz, 1H, *H3*), 6.73 (dd, *J* = 5.3, 1.5 Hz, 1H, *H2*), 6.69 (d, *J* = 6.0 Hz, 1H, N*H*CO), 6.62 – 6.55 (m, 1H, *H1*), 4.36 (qd, *J* = 15.8, 6.0 Hz, 2H, CH_2-NHCO), 4.13 (dd, *J* = 11.7, 4.2 Hz, 1H, CH_2-OH), 4.00 (dd, *J* = 11.8, 9.0 Hz, 1H, CH_2-O), 3.94 (dt, *J* = 13.7, 4.5 Hz, 1H, CH_2-N), 3.89 (s, 3H, OMe), 3.37 (ddd, *J* = 13.6, 10.1, 3.5 Hz, 1H, CH_2-N), 2.71 (tt, *J* = 9.1, 4.4 Hz, 1H, C*H*-CO), 2.04 – 1.95 (m, 1H, CH_2-CHCO), 1.94 – 1.86 (m, 1H, CH_2-CHCO), 1.47 (s, 9H, C(CH_3)$_3$)); ^{13}C NMR (151 MHz, $CDCl_3$) δ 171.8 (NH*C*O), 164.6 (*C*-OMe), 154.8 (CO), 150.0 (C3), 147.1 (C1’), 115.6 (C2), 108.8 (C1), 108.8, 82.0 (*C*(CH_3)$_3$), 71.9 (*C*H_2-O), 53.4 (C-O*Me*), 45.0 (*C*H_2-N), 42.1 (*C*H_2-NHCO), 41.5 (*C*-CO), 28.2 (C(*C*H_3)$_3$), 25.8 (*C*H_2-CHCO); HRMS (ESI) (*m*/*z*): Calcd for $C_{17}H_{25}N_3O_5Na$ [M+Na]$^+$ 374.1681; Found: 350.1682.

N-((2-Methoxypyridin-4-yl)methyl)-1,2-oxazinane-5-carboxamide (65).



A stirred solution of **64** (0.040 g, 0.11 mmol) in DCM (3 mL) was treated with trifluoroacetic acid (1 mL) at room temperature and stirred for 2 h, before it was diluted DCM (5 mL), and washed with saturated aqueous $NaHCO_3$ until effervescence stopped. The organic layer was dried on Na_2SO_4 and evaporated to dryness to give a residue that was subjected to flash chromatography over silica gel eluting with MeOH:DCM (2:98) to give **65** (26.7 mg, 93%) as an oil. ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.08 (d, *J* = 5.2 Hz, 1H, *H3*), 6.78 – 6.71 (m, 1H, *H2*), 6.58 (s, 1H, *H1*), 6.48 (d, *J* = 6.1 Hz, 1H, N*H*-CH_2), 4.42-4.35 (m, 2H), 4.12 (dd, *J* = 11.9, 4.0 Hz, 1H, C*H*$_2$-NHCO), 3.95 (dd, *J* = 11.9, 8.0 Hz, 1H, C*H*$_2$-O), 3.99-3.34 (m, 1H, C*H*$_2$-O), 3.90 (s, 3H, MeO), 3.22 (dt, *J* = 13.3, 4.8 Hz, 1H, C*H*$_2$-N), 3.13 (ddd, *J* = 13.1, 8.7, 3.9 Hz, 1H, C*H*$_2$-N), 2.62 (tt, *J* = 8.6, 4.5 Hz, 1H, C*H*CO), 2.12 – 1.90 (m, 2H, C*H*$_2$-CO)); ^{13}C NMR (151 MHz, $CDCl_3$) δ 172.6 (*C*O-NH), 164.6 (C-OMe), 150.0 (C3), 147.2 (C1'), 115.6 (C2), 108.9 (C1), 71.9 (*C*H_2-O), 53.5 (C-O*Me*), 47.1 (*C*H_2-N), 42.1 (*C*H_2-NHCO), 41.8 (*C*H-CO), 26.7 (*C*H_2-CHCO); HRMS (ESI) (*m*/*z*): Calcd for $C_{12}H_{18}N_3O_3$ $[M+H]^+$ 252.1343; Found: 252.1339.

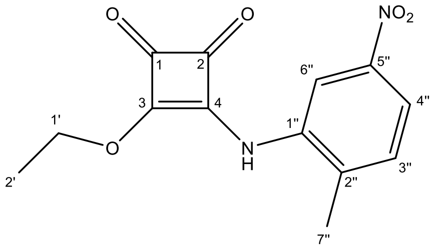
(±)-*N*-((2-Methoxypyridin-4-yl)methyl) 2-(1-(naphthalen-1-yl)ethyl)-1,2-oxazinane-5-carboxamide (66).



A mixture of **65** (0.015 g, 0.059 mmol) and 1-acetonaphthone (0.015 g, 0.089 mmol) was dissolved in MeOH (2 mL) and stirred at room temperature for 4 hr before it was treated sequentially with $NaCNBH_3$ (0.011 g, 0.17 mmol) and glacial acetic acid (0.18 g, 0.29 mmol) and stirred at room temperature for 36 hr. Saturated aqueous $NaHCO_3$ was added and the mixture extracted with DCM (2 x 8 mL). The combined organic layers were dried on Na_2SO_4 and evaporated under reduced pressure to give an oil, which was purified by column chromatography over silica gel eluting with 1:1 EtOAc:Hexane to afford the title compound as a clear oil (8.3 mg, 34%) . ^{1}H NMR (600 MHz, MeOH-d_4) δ 8.43 (d, *J* = 8.5 Hz, 1H, H8), 8.05 (dd, *J* = 5.4, 0.7 Hz, 1H, H4'), 7.98 – 7.84 (m, 1H, H5), 7.79 (d, *J* = 8.1 Hz, 1H, H4), 7.59 (d, *J* = 7.1 Hz, 1H, H2), 7.55 – 7.42 (m, 4H, H3, H6, H7, H3'), 6.85 (dd, *J* = 5.4, 1.4 Hz, 1H, H1'), 6.68 (dt, *J* = 1.7, 0.8 Hz, 1H, NH), 4.43 (d, *J* = 6.7 Hz, 1H, NCH), 4.35 (s, 2H, C*H*$_2$-NH-CO), 4.23 – 4.20 (dd, *J* = 11.5, 4.7 Hz, 1H), 4.16-4.13 (dd, *J* = 11.5, 10.0 Hz, 1H), 3.89 (s, 3H), 2.84 (dt, *J* = 12.0, 4.0 Hz, 1H), 2.64 (m, 1H), 2.03 (m, 1H), 1.85 (bs, 1H), 1.58-1.58 (d, *J* = 6.5 Hz, 3H, C*H*$_3$); ^{13}C NMR (151 MHz, MeOH-d_4) δ 174.2 (*C*O-NH), 164.8 (*C*-OMe), 151.5 (C4'), 146.3 (C4'), 138.7 (C1), 134.2 (C9), , 131.1 (C2'), 128.4, 127.4 (C5), 125.3, 125.1, 125.1(C4), 125.0 (C7), 124.9 (C6), 123.7(C2), 115.4 (C3'), 108.1 (C1'), 108.1, 71.1 (*C*H$_2$-O-N), 63.9 (CH$_3$-*C*H-N), 54.4 (*C*H$_2$-N-O), 52.3 (C-

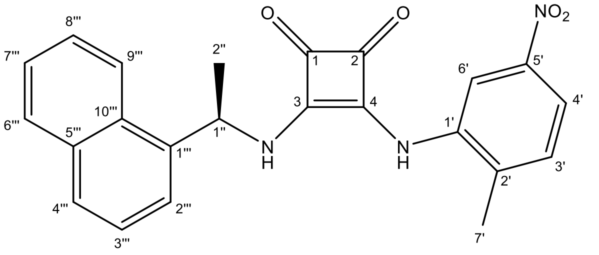
O*Me*)41.3 (*C*H$_2$-NHCO), 29.3 (*C*HCO), 27.1 (*C*H$_2$-CHCO), 19.2 (*C*H$_3$); HRMS (ESI) (*m*/*z*): Calcd for $C_{24}H_{28}N_3O_3$ [M+H]$^+$ 406.2092; Found: 406.2099.

3-Ethoxy-4-((2-methyl-5-nitrophenyl)amino)cyclobut-3-ene-1,2-dione (67).



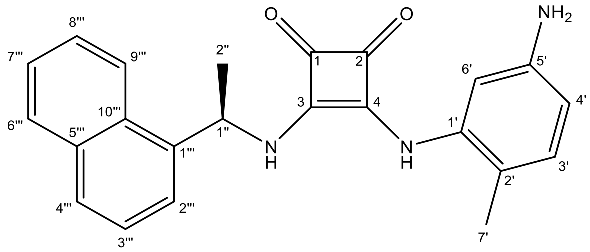
A solution of diethyl squarate (2.05 mL, 13.8 mmol) and 2-methyl-5-nitroaniline (2.00 g, 13.14 mmol) in ethanol was heated under microwave irradiation to 120°C for 4 hr. The reaction mixture was then concentrated under vacuum, suspended in ether, and filtered. The precipitate was washed with MeCN and the filtrate was concentrated to give the title compound (0.647 g, 2.34 mmol, 18%) as a yellow solid. ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.23 – 8.16 (m, 2H, H-6", H-10), 7.95 (dd, *J* = 8.3, 2.3 Hz, 1H, H-4"), 7.37 (d, *J* = 8.3 Hz, 1H, H-3"), 4.89 (q, *J* = 7.0 Hz, 2H, H-1'), 2.49 (s, 3H, H-7"), 1.50 (t, *J* = 7.0 Hz, 3H, H-2'). ^{13}C NMR (126 MHz, CDCl$_3$) δ 188.8 (C-1), 184.8 (C-2), 179.0 (C-3), 168.5 (C-4), 147.0 (C-5"), 136.1 (C-1", C-2"), 131.7 (C-3"), 120.0 (C-4"), 116.1 (C-6"), 71.1 (C-1'), 18.2 (C-7"), 15.8 (C-2'). HRMS (ESI): *m/z*: Calcd for $C_{13}H_{13}N_2O_5$ [M+H]$^+$ 277.0814; Found 277.0811.

N-((2-Methyl-5-nitrophenyl)) N'-®-((1-(naphthalen-1-yl)ethyl))squaramide (68).



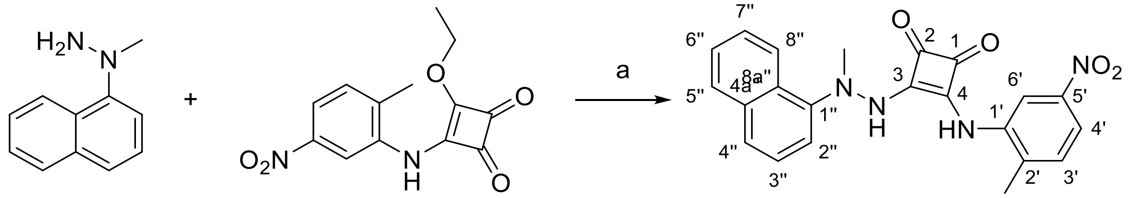
A solution of 3-ethoxy-4-((2-methyl-5-nitrophenyl)amino)cyclobut-3-ene-1,2-dione **67** (0.304 g, 1.10 mmol) and (R)-(+)-1-(1-naphthyl)ethylamine (0.19 mL, 1.2 mmol) in ethanol (3 mL) was heated under microwave irradiation to 50°C for 15 min. The reaction mixture was concentrated and adsorbed on Celite® followed by purification over silica gel eluting with ethyl acetate to give the title compound (0.348 g, 0.87 mmol, 79%) as a yellow solid. $[\alpha]_D^{23}$ = -140.3 (*c* = 1.0, $CHCl_3$), ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.78 (s, 1H, NH), 8.56 (d, *J* = 7.8 Hz, 1H, NH'), 8.24 (d, *J* = 2.3 Hz, 1H, H-6'), 7.97 (d, *J* = 8.5 Hz, 1H, H-9'''), 7.74 (dd, *J* = 8.1, 1.4 Hz, 1H, H-6'''), 7.64 (d, *J* = 8.2 Hz, 1H, H-4'''), 7.57 (dd, *J* = 8.3, 2.3 Hz, 1H, H-4'), 7.48 – 7.41 (m, 2H, H-2''', H-8'''), 7.38 (t, *J* = 7.4 Hz, 1H, H-7'''), 7.28 (t, *J* = 7.7 Hz, 1H, H-3'''), 6.98 (d, *J* = 8.3 Hz, 1H, H-3'), 6.04 (p, *J* = 6.9 Hz, 1H, H-1''), 2.05 (s, 3H, H-7'), 1.70 (d, *J* = 6.7 Hz, 3H, H-2''). ^{13}C NMR (126 MHz, $CDCl_3$) δ 182.6 (C-1), 181.3 (C-2), 170.9 (C-3), 162.4 (C-4), 146.7 (C-5'), 137.5 (C-1'''), 137.0 (C-1'), 135.0 (C-2'), 134.0 (C-5'''), 131.1 (C-3'), 130.2 (C-10'''), 129.1 (C-6'''), 128.8 (C-4'''), 126.6 (C-8'''), 125.9 (C-7'''), 125.5 (C-3'''), 123.0 (C-2'''), 122.7 (C-9'''), 118.4 (C-4'), 115.3 (C-6'), 51.0 (C-1''), 23.1 (C-2''), 18.1 (C-7'). HRMS (ESI) *m/z*: Calcd for $C_{23}H_{19}N_3O_4Na$ $[M+Na]^+$ 424.1268; Found 424.1261.

***N*-((5-Amino-2-methylphenyl)) *N’*-(*R*)-((1-(naphthalen-1-yl)ethyl))squaramide (69).**



Palladium on carbon (10 wt%, 2 mg) was added to a solution of *N*-((2-methyl-5-nitrophenyl)) *N*’-(*R*)-((1-(naphthalen-1-yl)ethyl))squaramide **68** (20 mg, 0.05 mmol) in 1,4-dioxane (1 mL) and the reaction mixture was stirred under H_2 at atmospheric pressure for 2 hr followed by addition of 10% aqueous acetic acid (1.0 mL). After an additional 0.5 hr stirring, palladium on carbon (10 wt%, 6.6 mg) was added and the reaction mixture was stirred for 1 hr before it was filtered and concentrated under vacuum. The crude residue was then dry loaded on Celite® and purified over silica gel eluting with 0 to 10% methanol in CH_2Cl_2 to give the title compound (7.9 mg, 0.02 mmol, 43%) as a brown solid. $[\alpha]_D^{23}$ = -88.29 (c = 0.3, $CHCl_3$), ^{1}H NMR (500 MHz, CD_3OD) δ 8.14 (br s, 1H, H-9'''), 7.89 (dd, *J* = 8.2, 1.4 Hz, 1H, H-6'''), 7.82 (d, *J* = 8.2 Hz, 1H, H-4'''), 7.61 – 7.53 (m, 2H, H-2''', H-8'''), 7.51 – 7.47 (m, 2H, H-3''', H-7'''), 6.85 (d, *J* = 8.1 Hz, 1H, H-3'), 6.74 (br s, 1H, H-6'), 6.41 (dd, *J* = 8.1, 2.3 Hz, 1H, H-4'), 6.24 (br s, 1H, H-1''), 2.09 (s, 3H, H-7'), 1.76 (d, *J* = 6.8 Hz, 3H, H-2''). ^{13}C NMR (126 MHz, CD_3OD) δ 184.1 (C-1), 181.5 (C-4), 168.3 (C-2), 164.8 (C-3), 146.4 (C-5'), 138.1 (C-1'''), 136.5 (C-1'), 134.3 (C-5'''), 130.9 (C-3'), 130.6 (C-10'''), 128.7 (C-6'''), 128.3 (C-4'''), 126.4 (C-8'''), 125.7 (C-3'''), 125.2 (C-7'''), 122.6 (C-9'''), 122.3 (C-2'''), 117.5 (C-2'), 112.3 (C-4'), 108.7 (C-6'), 49.8 (C-1''), 22.0 (C-2''), 15.5 (C-7'). HRMS (ESI) *m/z*: Calcd for $C_{23}H_{22}N_3O_2$ $[M+H]^+$ 372.1707; Found 372.1700.

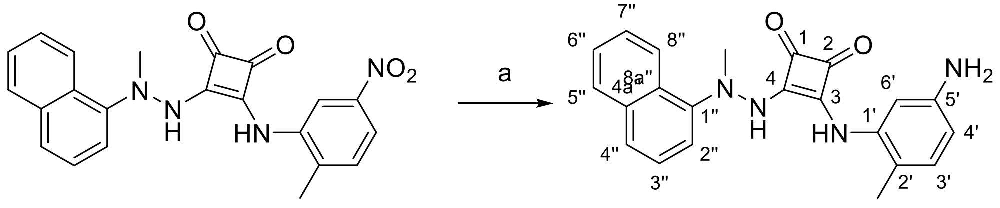
3-[*N*'-(Naphthalen-1-yl)-*N*'-methylhydrazinyl]-4-[(2-methyl-5-nitrophenyl)amino]cyclobut-3-ene-1,2-dione (70).



Reagents and conditions: (a) EtOH, 20°C, 16 hr.

N-Methyl-*N*-(1-naphthyl)hydrazine **18** (90 mg, 0.52 mmol) and 3-ethoxy-4-((2-methyl-5-nitrophenyl)amino)cyclobut-3-ene-1,2-dione **67** (173 mg, 0.63 mmol) were taken up in ethanol (1.7 mL) and stirred at 20°C for 16 hr. The reaction mixture was concentrated and purified by silica gel column chromatography (20 - 60% ethyl acetate/hexane) to yield the title product (63.1 mg, 30%) as a dark brown syrup. ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.64 (s, 1H, H-6′), 8.45 (s, 1H, H-8′), 8.08 (d, *J* = 7.9 Hz, H-5″), 7.89 (d, *J* = 7.9 Hz, 1H, H-4′), 7.73 (m, 2H), 7.52 (m, 4H), 7.14 (d, *J* = 8.3 Hz, 1H), 3.40 (s, 3H, CH_3), 1.90 (s, 3H, -NCH_3). ^{13}C NMR (126 MHz, $CDCl_3$) δ 183.6 (C-1), 179.2 (C-2), 167.0 (C-3), 147.3 (C-2′), 146.5 (C-5), 136.7 (C-1), 134.9, 134.4, 131.1, 129.5, 127.0, 126.6, 126.6, 126.2, 121.6, 119.1, 116.1, 115.9, 48.3 (CH_3), 17.7 (-NCH_3). HRMS (ESI) *m/z*: [M + Na]$^+$ Calcd for $C_{22}H_{18}N_4O_4Na$ 425.1226; Found 425.1230.

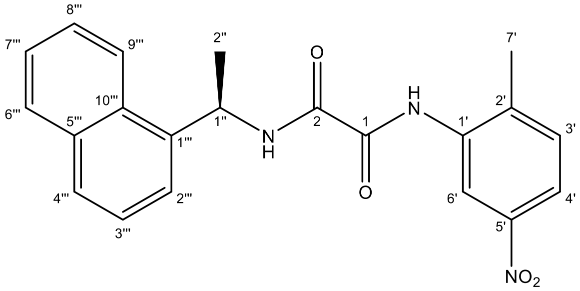
3-[*N*'-(Naphthalen-1-yl)-*N*'-methylhydrazinyl]-4-[(5-amino-2-methylphenyl)amino]cyclobut-3-ene-1,2-dione (71).



Reagents and conditions: (a) Pd/C, H_2, 1,4-dioxane, AcOH, 20°C, 5 hr.

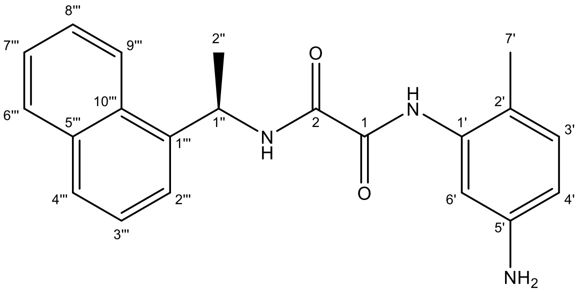
To a stirred solution of 3-[*N*'-(naphthalen-1-yl)-*N*'-methylhydrazinyl]-4-[(2-methyl-5-nitrophenyl)amino]cyclobut-3-ene-1,2-dione **70** (47.7 mg, 0.12 mmol) in a 1:1 mixture of 1,4-dioxane (1.2 mL) and 10% aqueous acetic acid (0.6 mL) was added Pd/C (6.31 mg, 0.01 mmol). The resulting mixture was stirred at 20°C under hydrogen (1 atm) for 5 h, then was filtered and the filtrate concentrated to give a residue, which was purified by silica gel column chromatography (20 - 60% ethyl acetate/hexane) to furnish the title product (38.4 mg, 87%) as an off-white solid. ^{1}H NMR (500 MHz, $CDCl_3$ + CD_3OD) δ 8.00 (br s, 1H, H-8″), 7.83 (d, *J* = 9.5 Hz, 1H, H-5″), 7.66 (d, *J* = 8.0 Hz, 1H, H-4″), 7.47 (m, 2H, H-6″, H-3″), 7.43 (d, *J* = 7.9 Hz, 1H, H-2″), 7.39 (t, *J* = 7.7 Hz, 1H, H-7″), 7.02 (br s, 1H, H-6′), 6.77 (d, *J* = 8.3 Hz, 1H, H-3′), 6.33 (dd, *J* = 8.3, 2.3 Hz, 1H, H-4′), 3.27 (s, 3H, -N*CH₃*), 1.78 (s, 3H, $-CH_3$). ^{13}C NMR (126 MHz, $CDCl_3$ + CD_3OD) δ 182.4 (C-1), 166.7 (C-2), 146.2 (C-5′), 136.3 (C-1′), 134.9 (C-4a″), 131.3 (C-3′), 129.1 (C-8a″), 126.9 (C-5″), 126.5 (C-4″), 126.5 (C-3″), 125.6 (C-2″), 122.1 (C-8″), 115.0 (C-7″), 112.2 (C-4′), 107.9 (C-6′), 47.7 (-N*CH₃*), 16.4 (CH_3). HRMS (ESI) *m*/*z*: $[M + H]^+$ Calcd for $C_{22}H_{20}N_4O_2$ 373.1659; Found 373.1680.

N^1-(2-Methyl-5-nitrophenyl) N^2-(*R*)-(1-(naphthalen-1-yl)ethyl)oxalamide (72).



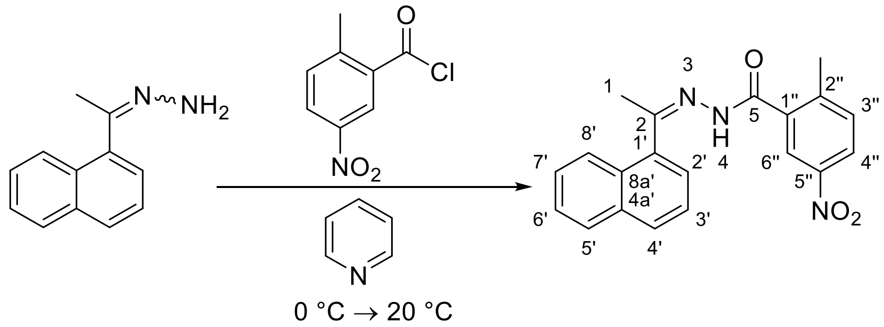
A solution of 2-methyl-5-nitroaniline (101 mg, 0.66 mmol) in CH_2Cl_2 (3 mL) was added to a stirred solution of oxalyl chloride (0.17 mL, 1.98 mmol) and K_2CO_3 (0.273 g, 1.98 mmol) in CH_2Cl_2 (3 mL) at 0°C. After 10 min the reaction mixture was concentrated under vacuum and dissolved in CH_2Cl_2 (12 mL) followed by addition of (*R*)-(+)-1-(1-naphthyl)ethylamine (0.11 mL, 0.69 mmol) at 0°C. The reaction mixture was stirred for 15 min then diluted with EtOAc and washed with 1N HCl and brine. The organic layer was dried with Na_2SO_4, filtered, and concentrated. The crude residue was adsorbed on Celite® and purified over silica gel eluting with 70% to 100% CH_2Cl_2 in hexanes to give the title compound (0.151 g, 0.40 mmol, 60%) as a white solid. $[\alpha]_D^{23}$ = -59.71 (*c* = 0.07, $CHCl_3$), ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.42 (s, 1H, NH), 8.99 (d, *J* = 2.4 Hz, 1H, H-6'), 8.07 (d, *J* = 8.5 Hz, 1H, H-9'''), 7.95 (dd, *J* = 8.4, 2.4 Hz, 1H, H-4'), 7.90 – 7.85 (m, 2H, NH', H-4'''), 7.82 (d, *J* = 8.2 Hz, 1H, H-6'''), 7.59 – 7.54 (m, 2H, H-2''', H-3'''), 7.53 – 7.44 (m, 2H, H-7''', H-8'''), 7.34 (d, *J* = 8.4 Hz, 1H, H-3'), 5.93 (qd, *J* = 6.9, 6.9, 1.6 Hz, 1H, H-1’’), 2.42 (s, 3H, H-7'), 1.77 (d, *J* = 6.9 Hz, 3H, H-2''). ^{13}C NMR (126 MHz, $CDCl_3$) δ 158.5 (C-2), 157.7 (C-1), 147.1 (C-5'), 137.1 (C-1'''), 135.4 (C-1'), 135.2 (C-2'), 134.1 (C-5'''), 131.2 (C-3'), 130.8 (C-10'''), 129.2 (C-4'''), 128.9 (C-6'''), 126.8 (C-3'''), 126.1 (C-8'''), 125.4 (C-7'''), 122.8 (C-2'''), 122.8 (C-9'''), 120.2 (C-4'), 115.9 (C-6'), 46.0 (C-1''), 21.1 (C-2''), 17.8 (C-7'). HRMS (ESI) *m/z*: Calcd for $C_{21}H_{19}N_3O_4Na$ $[M+Na]^+$ 400.1269; Found 400.1263.

N^1-(5-Amino-2-methylphenyl) N^2-(*R*)-(1-(naphthalen-1-yl)ethyl)oxalamide (73).



Palladium on carbon (10 wt%, 5.2 mg) was added to a solution of N^1-(2-methyl-5-nitrophenyl) N^2-(*R*)-(1-(naphthalen-1-yl)ethyl)oxalamide **72** (33.7 mg, 0.09 mmol) in 1,4-dioxane (2 mL) and the reaction mixture was stirred under H_2 for 1 hr followed by addition of further palladium on carbon (11 mg). After an additional 1 hr of stirring glacial AcOH (0.25 mL) was added, and the reaction mixture was stirred for 9 hr before filtration and concentration under vacuum. The residue was adsorbed on Celite® and purified over silica gel eluting with 25% to 100% ethyl acetate in hexanes to give the title compound (16.3 mg, 0.05 mmol, 53%) as a yellow film. $[\alpha]_D^{23}$ = -40.34 (*c* = 0.7, $CHCl_3$), ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.27 (s, 1H, NH), 8.09 (d, *J* = 8.4 Hz, 1H, H-9'''), 7.92 (d, *J* = 8.5 Hz, 1H, NH'), 7.87 (dd, *J* = 8.1, 1.4 Hz, 1H, H-6'''), 7.81 (d, *J* = 8.2 Hz, 1H, H-4'''), 7.59 – 7.53 (m, 2H, H-2''', H-8'''), 7.52 – 7.48 (m, 2H, H-6', H-7'''), 7.46 (dd, *J* = 8.2, 7.2 Hz, 1H, H-3'''), 6.95 (d, *J* = 8.1 Hz, 1H, H-3'), 6.42 (dd, *J* = 8.1, 2.4 Hz, 1H, H-4'), 5.92 (dq, *J* = 8.5, 6.9 Hz, 1H, H-1''), 2.22 (s, 3H, H-7'), 1.74 (d, *J* = 6.9 Hz, 3H, H-2''). ^{13}C NMR (126 MHz, $CDCl_3$) δ 159.3 (C-2), 157.2 (C-1), 145.3 (C-5'), 137.4 (C-1'''), 135.2 (C-1'), 134.1 (C-5'''), 131.2 (C-3'), 130.9 (C-10'''), 129.1 (C-6'''), 128.7 (C-4'''), 126.8 (C-8'''), 126.0 (C-7'''), 125.4 (C-3'''), 122.9 (C-9'''), 122.8 (C-2'''), 117.7 (C-2'), 112.4 (C-4'), 107.8 (C-6'), 45.8 (C-1''), 21.1 (C-2''), 16.5 (C-7'). HRMS (ESI) *m/z*: Calcd for $C_{21}H_{21}N_3O_2Na$ $[M+Na]^+$ 370.1526; Found 370.1518.

(*E/Z*)-2-Methyl-*N′*-(1-(naphthalen-1-yl)ethylidene)-5-nitrobenzohydrazide (75).



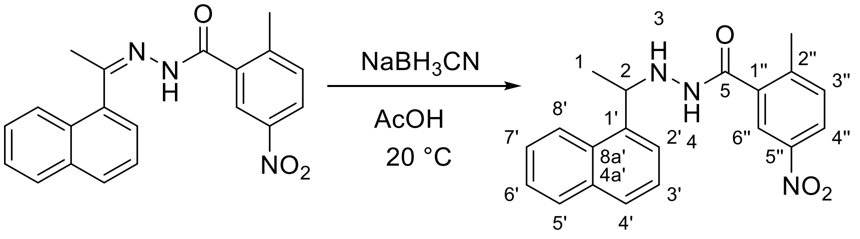
(*E/Z*)-(1-(Naphthalen-1-yl)ethylidene)hydrazine **74**[13] (4 g, 21.71 mmol) was dissolved in pyridine (3 mL) and anhydrous CH_2Cl_2 (2 mL) and cooled down to 0°C before 2-methyl-5-nitrobenzoyl chloride (4.77 g, 23.88 mmol) in anhydrous CH_2Cl_2 (4 mL) was added dropwise. The reaction mixture then was allowed to warm up to 20°C and was stirred for 15 min before it was filtered and the white precipitate was washed with cold CH_2Cl_2 (3×10 mL) and hexanes (3×10 mL) to give a mixture of (*E/Z*)-isomers of the title compound as a white solid (6.8 g, 90 %). ^{1}H NMR (500 MHz, DMSO-*d6*) δ 11.38 (s, 1H, NH), 11.13 (s, 1H, NH(a)[2]), 8.37 (d, *J* = 2.5 Hz, 1H, H-6’’), 8.30 – 8.25(m, 2H, H-4’’, H-5’), 8.21 (d, *J* = 2.5 Hz, 1H, H-6(a)’’), 8.08 (dd, *J* = 8.5, 2.5 Hz, 1H, H-4(a)’’), 8.03 – 7.96 (m, 2H, H-5(a)’, H-4’), 7.92 – 7.81 (m, 3H, H-4(a)’, H-8’, H-8(a)’), 7.67 – 7.53 (m, 5H, H-3’’, H-3(a)’’, H-7’, H-7(a)’, H-3’), 7.51 – 7.39 (m, 4H, H-3(a)’, H-2’, H-2(a)’, H-6’), 7.23 (td, *J* = 6.8, 3.3 Hz, 1H, H-6(a)’), 2.55 (s, 3H, 3×H-1), 2.46 (s, 6H, CH_3, CH_3(a)), 2.38 (s, 3H, 3×H-1a). ^{13}C NMR (126 MHz, DMSO-*d6*) δ 170.1 (C-5), 163.7 (C-5(a)), 157.7 (C-2), 152.0 (C-2(a)), 145.3 (C-5’’), 145.2 (C-5(a)’’), 144.5 (C-2’’), 142.5 (C-2(a)’’), 138.2 (C-4a’), 137.5 (C-4a(a)’), 136.7 (C-8(a)’, C-8a’), 133.4 (C-4’), 133.3 (C-4(a)’), 131.9 (C-1’’), 131.3 (C-1(a)’’), 130.0 (C-1’), 129.6 (C-1(a)’) , 129.0 (C-5’, C-5(a)’), 128.4 (C-7), 128.3 (C-7(a)’),

---

[2] a – corresponds to the second isomer

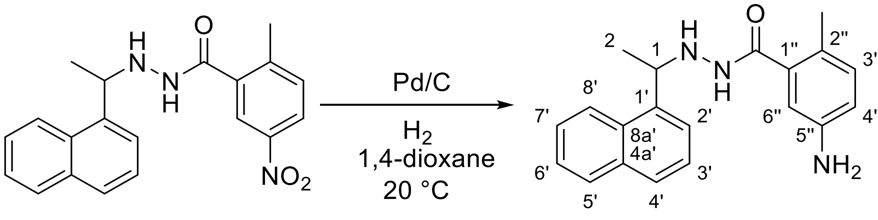
126.6 (C-3’’, C-3(a)’’), 126.1 (C-6’), 126.0 (C-6(a)’), 125.93 (C-8’), 125.88 (C-8(a)’), 125.6 (C-3’), 125.3 (C-3(a)’), 125.2 (C-2’), 125.1 (C-2(a)’), 124.3 (C-4’’), 123.30 (C-4(a)’’), 122.8 (C-6’’), 121.8 (C-6(a)’’), 19.7 (CH_3), 19.67 (CH_3(a)), 19.2 (C-1), 18.8 (C-1(a)). HRMS (ESI) *m/z*: Calcd for $C_{20}H_{18}N_3O_3$ $[M+H]^+$ 348.1343; Found 348.1337.

2-Methyl-*N′*-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzohydrazide (76).



(*E/Z*)-2-Methyl-N'-(1-(naphthalen-1-yl)ethylidene)-5-nitrobenzohydrazide **75** (2.3 g, 6.62 mmol) was suspended in MeOH—THF (1 : 1, (v/v)) (20 mL) and $NaBH_3CN$ (428 mg, 6.81 mmol, 1.03 eq.) was added followed by drop-wise addition of glacial AcOH (570 µL, 9.97 mmol) until pH 4—5. The reaction mixture was stirred 16 hr at 20°C (completion was detected by LCMS) and then quenched with 1M aq. NaOH (60 mL) and extracted with Et_2O (3×80 mL). The combined organic layers were dried over $MgSO_4$ and concentrated to dryness. The crude product was purified by flash column chromatography over silica gel (eluent: hexanes : EtOAc (EtOAc, 0 → 30 %)) to give the title compound as a yellowish solid (1.55 g, 67%).R_f 0.49 (hexanes : EtOAc 7 : 3 (v/v); UV). ^1H NMR (500 MHz, DMSO-*d*6) δ 9.95 (s, 1H, N<u>H</u>CO), 8.32 (t, *J* = 4.2 Hz, 1H, H-6’’), 8.13 (dd, *J* = 8.4, 2.5 Hz, 1H, H-4’’), 7.96 – 7.90 (m, 2H, H-8’,H-5’), 7.85 – 7.78 (m, 2H, H-4’, H-2’), 7.58 – 7.48 (m, 3H, H-6’, H-3’, H-7’), 7.46 (d, *J* = 8.4 Hz, 1H, H-3’’), 5.58 (br s, 1H, N<u>H</u>-NHCO), 5.13 (q, *J* = 6.6 Hz, 1H, H-2), 2.25 (s, 3H, CH_3), 1.49 (d, *J* = 6.6 Hz, 3H, H-1). ^{13}C NMR (126 MHz, DMSO-*d*6) δ 166.8 (C-5), 145.7 (C-5’’), 144.7 (C-1’’), 139.8 (C-1’), 137.0 (C-2’’), 134.0 (C-4a’), 132.4 (C-3’’), 131.6 (C-8a’), 129.2 (C-4’), 127.8 (C-5’), 126.5 (C-7’), 126.1 (C-6’), 126.0 (C-3’), 124.6 (C-8’), 124.3 (C-1’), 123.7 (C-6’’), 122.6 (C-4’’), 54.8 (C-2), 21.5 (CH_3), 19.7 (C-1). HRMS-ESI: m/z calcd. for $C_{20}H_{19}N_3NaO_3^+$ $[M+Na]^+$ 372.1324, found 372.1308.

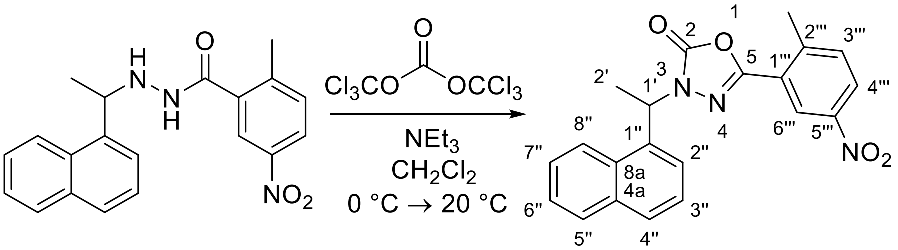
5-Amino-2-methyl-N'-(1-(naphthalen-1-yl)ethyl)benzohydrazide (77).



2-Methyl-*N'*-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzohydrazide **76** (150 mg, 0.429 mmol) was dissolved in 1,4-dioxane (4 mL) and Pd/C (10% wt) (27.4 mg, 0.025) was added. The reaction mixture then was stirred for 4 hr under H_2 (balloon) at 20°C. After completion, detected by LCMS, the reaction mixture was filtered through Celite and the filter cake was additionally washed with Et_2O (3×15 mL). The filtrate was concentrated to dryness and the crude product was purified by flash column chromatography over silica gel (eluent: hexanes : EtOAc (EtOAc, 0 → 30 %)) to give the title compound as a yellowish resin (69 mg, 51 %). R_f 0.15 (hexanes : EtOAc 7 : 3 (v/v); UV). ^1H NMR (500 MHz, $CDCl_3$) δ 8.31 (d, *J* = 8.3 Hz, 1H, H-8’), 7.86 (dd, *J* = 8.0, 1.6 Hz, 1H, H-4’), 7.79 – 7.69 (m, 2H, H-5’, H-2’), 7.52 – 7.41 (m, 3H, H-2’, H-6’, H-3’), 7.16 (s, 1H, H-6’), 6.84 (d, *J* = 8.1 Hz, 1H, H-3’’), 6.47 (dd, *J* = 8.1, 2.5 Hz, 1H, H-4’’), 6.28 (d, *J* = 2.5 Hz, 1H, H-6’’), 5.15 (m, 1H, H-1), 3.39 (br, 1H, NH), 2.11 (s, 3H, CH_3), 1.56 (d, *J* = 6.6 Hz, 3H, H-2), 1.24 (br, 2H, NH_2). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.7 (CO), 143.9 (C-5’’), 138.9 (C-2’’), 134.7 (C-4a’), 134.0 (C-8a’, C-1’), 131.7 (C-1’), 131.6 (C-4’), 128.9 (C-1’’), 127.9 (C-5’), 126.1 (C-

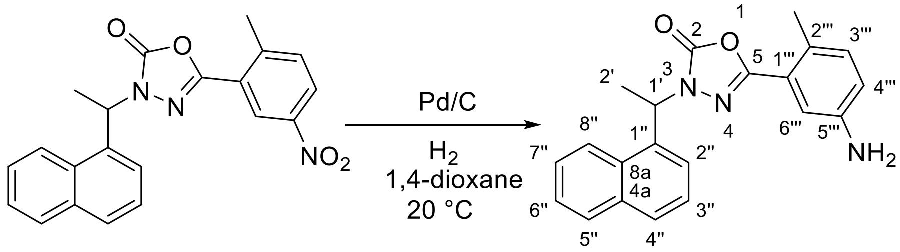
7’), 125.7 (C-3’’), 125.6 (C-6’), 123.8 (C-3’), 123.3 (C-8’), 117.0 (C-6’’), 113.7 (C-4’’), 55.8 (C-1), 20.9 (CH_3), 18.4 (C-2). HRMS (ESI): *m/z* Calcd for $C_{20}H_{21}N_3NaO$ $[M+Na]^+$ 342.1577; Found 342.1558.

5-(2-Methyl-5-nitrophenyl)-3-(1-(naphthalen-1-yl)ethyl)-1,3,4-oxadiazol-2(3*H*)-one (78).



To a stirred solution of triphosgene (135.9 mg, 0.457 mmol) in anhydrous CH_2Cl_2 (6 mL) at 0°C was added 2-methyl-*N'*-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzohydrazide **76** (400 mg, 1.14 mmol) in anhydrous CH_2Cl_2 (6 mL) and then, dropwise, NEt_3 (160 µL, 1.14 mmol). After warming to 20°C the reaction mixture was stirred for 4 hr under an Ar atmosphere. The reaction then was quenched with H_2O (20 mL), and the reaction mixture was extracted with CH_2Cl_2 (3×30 mL). The combined organic layers were dried over $MgSO_4$ and concentrated to dryness. The crude product was purified by flash column chromatography over silica gel (eluent: hexanes : EtOAc (EtOAc, 5 → 25 %)) to give the title compound as a yellow foam (310 mg, 72%). R_f 0.20 (hexanes : EtOAc 9 : 1 (v/v); UV). ^1H NMR (500 MHz, $CDCl_3$) δ 8.61 (d, $J = 2.5$ Hz, 1H, H-6’’’), 8.29 (d, $J = 8.6$ Hz, 1H, H-3’’’), 8.16 (dd, $J = 8.6, 2.5$ Hz, 1H, H-4’’’), 7.91 – 7.83 (m, 2H, H-7’’, H-6’’), 7.79 – 7.74 (m, 1H, H-8’’), 7.65 – 7.58 (m, 1H, H-4’’), 7.56 – 7.49 (m, 2H, H-2’’, H-5’’), 7.42 (d, $J = 8.4$ Hz, 1H, H-3’’), 6.28 (q, $J = 7.0$ Hz, 1H, H-1), 2.66 (s, 3H, CH_3), 2.02 (d, $J = 7.0$ Hz, 3H, CH_3(H-2)). ^{13}C NMR (126 MHz, $CDCl_3$) δ 152.0 (C-5), 151.7 (C-2), 146.3 (C-5’’’), 144.8 (C-1’’’), 134.7 (C-2’’’), 133.8 (C-4a’’), 132.8 (C-8a’’), 130.7 (C-1’’), 129.1 (C-4’’), 129.0 (C-5’’), 127.0 (C-7’’), 125.9 (C-3’’’), 125.1 (C-6’’), 125.0 (C-3’’), 124.3 (C-8’’), 123.8 (C-2’’), 122.8 (C-6’’’), 122.5 (C-4’’’), 50.8 (C-1’), 22.2 (CH_3), 19.2 (C-2’). HRMS (ESI) *m/z*: Calcd for $C_{21}H_{17}N_3NaO_4$ $[M+Na]^+$ 398.1113; Found 398.1097.

5-(5-Amino-2-methylphenyl)-3-(1-(naphthalen-1-yl)ethyl)-1,3,4-oxadiazol-2(3*H*)-one (79).



5-(2-Methyl-5-nitrophenyl)-3-(1-(naphthalen-1-yl)ethyl)-1,3,4-oxadiazol-2(3*H*)-one **78** (132 mg, 0.35 mmol) was dissolved in 1,4-dioxane (4 mL) and Pd/C (10% wt) (22.4 mg, 0.021 mmol) was added. The reaction mixture was stirred for 4 hr under H_2 (balloon) at 20°C. After completion, detected by LCMS, the reaction mixture was filtered through Celite and the filter cake was additionally washed with Et_2O (3×15 mL). The filtrate was concentrated to dryness and crude product was purified by flash column chromatography over silica gel (eluent: hexanes : EtOAc (EtOAc, 0 → 30 %)) to give the title compound as a yellow syrup (70 mg, 57 %). R_f 0.43 (hexanes : EtOAc 7 : 3 (v/v); UV). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.27 (d, *J* = 8.6 Hz, 1H, H-8’’), 7.89 – 7.79 (m, 2H, H-4’’, H-5’’), 7.76 – 7.70 (m, 1H, H-2’’), 7.59 (ddd, *J* = 8.6, 6.8, 1.4 Hz, 1H, H-7’’), 7.55 – 7.44 (m, 2H, H-6’’,H-3’’), 7.06 (d, *J* = 2.6 Hz, 1H, H-6’’’), 7.01 (d, *J* = 8.2 Hz, 1H, H-3’’’), 6.66 (dd, *J* = 8.1, 2.5 Hz, 1H, H-4’’’), 6.21 (q, *J* = 7.0 Hz, 1H, H-1’), 3.21 (s, 2H, NH_2), 2.41 (s, 3H, CH_3), 1.96 (d, *J* = 7.0 Hz, 3H, H-2’). ^{13}C NMR (126 MHz, $CDCl_3$) δ 153.8 (C-5) , 152.8

(C-1), 144.3 (C-5’’’), 135.37 (C-2’’’), 133.8 (C-4a’’), 132.6 (C-8a’’), 130.8 (C-1’’), 129.0 (C-4’’), 128.9 (C-5’’), 127.5 (C-7’’), 126.9 (C-3’’’), 125.9 (C-6’’), 125.2 (C-3’’), 124.3 (C-8’’), 123.1 (C-1’’’), 122.8 (C-1’’), 117.9 (C-6’’’), 114.0 (C-5’’’), 50.6 (C-2’), 20.9 (CH_3), 19.4 (C-2’). HRMS (ESI) *m/z* Calcd for $C_{21}H_{19}N_3NaO_2$ $[M+H]^+$ 346.1550; Found 346.1532.

References

1. Ghosh, A. K.; Takayama, J.; Aubin, Y.; Ratia, K.; Chaudhuri, R.; Baez, Y.; Sleeman, K.; Coughlin, M.; Nichols, D. B.; Mulhearn, D. C.; Prabhakar, B. S.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D., Structure-Based Design, Synthesis, and Biologial Evaluation of a Series of Novel and Reversible Inhibitors for the Severe Acute Respiratory Syndrome-Coronavirus Papain-Like Protease. *J. Med. Chem.* **2009,** *52*, 5228-5240.

2. Portoghese, P. S.; Sultana, M.; Takemori, A. E., Design of peptidomimetic delta-opioid receptor antagonists using the message address concept. *J. Med. Chem.* **1990,** *33*, 1714-1720.

3. Korth, H.-G.; Sustmann, R.; Lommes, P.; Paul, T.; Ernst, A.; de Groot, H.; Hughes, L.; Ingold, K. U., Nitric Oxide Cheletropic Traps (NOCTs) with Improved Thermal Stability and Water Solubility. *J. Am. Chem. Soc.* **1994,** *116*, 2767-277.

4. Li, G.; Huan, Y.; Yuan, B.; Wang, J.; Jiang, Q.; Lin, Z.; Shen, Z.; Huang, H., Discovery of novel xanthine compounds targeting DPP-IV and GPR119 as anti-diabetic agents. *Eur. J. Med. Chem.* **2016,** *124*, 103-116.

5. Báez-Santos, Y. M.; Barraza, S. J.; Wilson, M. W.; Agius, M. P.; Mielech, A. M.; Davis, N. M.; Baker, S. C.; Larsen, S. D.; Mesecar, A. D., X-ray Structural and Biological Evaluation of a Series of Potent and Highly Selective Inhibitors of Human Coronavirus Papain-like Proteases *J. Med. Chem.* **2014,** *57*, 2393-2412.

6. Ghosh, A. K.; Takayama, J.; Rao, K. V.; Ratia, K.; Chaudhuri, R.; Mulhearn, D. C.; Lee, H.; Nichols, D. B.; Baliji, S.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D., Severe Acute Respiratory Syndrome-Coronavirus Papain-Like Protease Inhibitors: Design, Synthesis, Protein-Ligand X-ray Structure and Biologial Evaluation *J. Med. Chem.* **2010,** *53*, 4968-4979.

7. Radchenko, D. S.; Grygorenko, O. O.; Komarov, I. V., Synthesis of 2-azaspiro[3.3]heptane-derived amino acids: ornitine and GABA analogues. *Amino Acids* **2010,** *39*, 515-521.

8. Palmer, A. M.; Chiesa, V.; Schmid, A.; Munch, G.; Grobbel, B.; Zimmermann, P. J.; Brehm, C.; Buhr, W.; Simon, W.-A.; Kromer, W.; Postius, S.; Volz, J.; Hess, D., Tetrahydrochromenoimidazoles as Potassium-Competitive Acid Blockers (P-CABs): Structure-Activity Relationship of Their Antisecretory Properties and Their Affinity toward the hERG Channel. *J. Med. Chem.* **2010,** *53*, 3645-3674.

9. Lopez, J. C.; Blanco, S.; Lesarri, A.; Alonso, J. L., Internal Dynamics in Azetidine: A Microwave and ab initio study. *J. Phys. Chem.* **2001,** *114*, 2237.

10. Belostotskii, A. M.; Markevich, E., Conformational Dynamics in Nitrogen-Fused Azabicycles. *J. Org. Chem.* **2003,** *68*, 3055-3063.

11. El Ashry, E. S. K., Y.; Hamid, H. A.; El-Zemity, S. R.; Boghdady, S., Synthesis of functionalised derivatives of pentaerythritol. *J. Chem. Res., Miniprint* **2003**, 111-128.

12. Marion, F.; Calvet, S.; Marié, J.-C.; Courillon, C.; Malacria, M., Silylated Vinyloxiranes – Recent Advances and Synthetic Applications. *Eur. J. Org. Chem.* **2006**, 453-462.

13. Takacs, A.; Farkas, R.; Petz, A.; Kollar, L., Synthesis of 2-naphthylacrylamides and 2-naphthylacrylates via homogeneous catalytic carbonylation of 1-iodo-1-naphthylethene derivatives. *Tetrahedron* **2009,** *65*, 4795-4800.

APPENDIX C

SUPPLEMENTARY MATERIAL FOR CHAPTER 5

Table S5.1 Data collection and refinement statistics

	Chicken OASL Ubl
Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	28.9, 100.9, 55.8
α, β, γ (°)	90.0, 105.0, 90.0
Resolution (Å)	28.53-2.23 (2.29-2.23)
$R_{pim}(\%)$	6.0 (81.7)
$R_{\mathrm{merge}}(\%)$	8.3 (103.6)
$cc_{1/2}$	0.993 (0.775)
$I/\sigma I$	8.30 (1.30)
Completeness (%)	93.8 (80.0)
Redundancy	2.9 (2.3)
Refinement	
Resolution (Å)	27.9-2.23 (2.31-2.23)
No. reflections	13956
$R_{\mathrm{work}}(\%)/ R_{\mathrm{free}}(\%)$	22.82/25.84
No. atoms	
Protein	2556
Ligand/ion	0
Water	60
B-factors	
Protein	48.8
Ligand/ion	
Water	39.2
R.m.s deviations	
Bond lengths (Å)	0.007
Bond angles (º)	0.85

*Highest resolution shell is shown in parenthesis.