INVESTIGATING THE SUBSTRATE SPECIFICITY OF THE EQUIVALENT PAPAIN-LIKE PROTEASE 2 DOMAIN OF NSP3 ACROSS ALPHA- AND BETA-CORONAVIRUSES

by

Jozlyn Clasman

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Biological Sciences

West Lafayette, Indiana

May 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Andrew Mesecar, Chair
Department of Biological Sciences
Dr. Cynthia Stauffachern
Department of Biological Sciences
Dr. Richard J. Kuhn
Department of Biological Sciences
Dr. Angeline Lyon
Department of Chemistry

Approved by:
Dr. Janice Evans
Head of the Graduate Program

For my superhero, fiancé Dan Cholger & two kitties, Independence and St. Paddy

ACKNOWLEDGMENTS

Although graduate school has been a roller coaster ride, I was not the only passenger along this journey. I am incredibly thankful for all of my amazing friends, family, colleagues, and mentors for all the support and encouragement that has shaped me into the scientist and person I am today. First of course, I'd like to thank my advisor and mentor Dr. Andrew Mesecar who has kept me motivated every step of the way and taught me that science is fun, pushing the boundaries of discovery. Thanks for always saying "keep doing what you're doing" even when I was in doubt. I cannot thank you enough for your patience and guidance in writing and thoughtful conversation about research and life. You are a great person and boss who I have always felt I can be open and honest with.

Next, I'd like to thank my committee members, Dr. Cynthia Stauffacher, Dr. Angeline Lyon, and Dr. Richard Kuhn who have challenged me with insightful discussions and kept me learning, engaged and on track in my research. They have taught me the importance of clearly communicating scientific ideas and have boosted my confidence as a presenter, writer, and scientific researcher, and with that I cannot thank them enough for their time, knowledge, and inspiration.

I'd also like to acknowledge our amazing collaborators Dr. Susan Baker and Dr. Scott Pegan and members in their labs. Thanks for everything Susan; you are so kind and truly an inspirational role model. Thanks for giving me the chance to present my research to the experts in this field and being patient with me throughout the years. Your lab produces such beautiful work that continues to fuel my passion for this project. Special thanks to Dr. Robert Mettelman and Amornrat O'Brien for directly contributing to Chapter 2. Rob, your paper edits were so helpful. Also, thanks Dr. Xufang Deng, Dr. Anna Mielech, Matthew Hackbart, and Aaron Volk as well as

Rob for sharing your exciting work in our combined lab meetings to stimulate discussions and new ideas. Also, thanks Scott for working with me, being patient with me, on edits for the paper. Thanks Courtney Daczkowski and John Dzimianski as well as Octavia Goodwin for allowing me to contribute to your exciting work, which is condensed in Chapter 3.

I would like to thank all the current and past members of the Mesecar lab, Dr. Yahira Báez-Santos, Dr. Nicole David-Hjortland, Dr. Yafang Chen, Dr. Sarah St. John, Dr. Yu-Chen Yen, Julia Luciano-Chadee, Dr. Qing Zhou, Aya Ryuzoji, Dr. Renata Everett, Brandon Anson, Corey Moore, Dr. Kristina Kesely, Dr. Courtney Daczkowski, Emma Lendy, Samadhi Kulathunga, Mackenzie Chapman, Adam Hamdani, Noah Danielson, and last but not least in the slightest, the well-organized lab manager and one of the nicest people I've had the pleasure to work with, Nicole Klinkner. I am so grateful to have had the chance to work with all of you. Every one of you have challenged me, offered constructive criticism, and have directly contributed to my growth in graduate school. In addition, I'd like to thank Dr. Katy and Sam Jenson for your willingness to meet with me in Boston, and Dr. Kiira Ratia for your awesome science starting this field with SARS-CoV PLpro.

I want to extend a special thanks to Yahira for kick starting the project in my rotation, meeting with me after you left the lab to go over kinetics and helping me develop my skills in writing and beautifying figures. Your edits directly contributed to Chapter 2. I was so lucky that you were my mentor and also thanks for those Catan game sessions. Thanks to Aya for helping me learn the ins and outs of FPLC systems, and Yafang for teaching me ubiquitin chain-cleavage assays. Also, special thanks to Renata for making human ISG15-PA in Chapter 4 and teaming up with me to model the AYE group in the PLpro-ISG15 crystal structure. Thanks for being a great friend and helpful co-worker. In addition, I'd like to specifically acknowledge Nicole Davis and Corey Moore for contributing their precious human USP proteins for Chapter 2. Thanks to Corey

and Brandon for your help with so many things I don't have enough time to name them all. We are the three musketeers! Thanks, Brandon, for always taking time out of your day to make me laugh and always being so helpful, whether it is filling the liquid nitrogen, making competent cells, talking about science, offering your buffers and aliquots, or just supportive talks; you made my graduate experience so much better. Also, thanks to Dr. Yu-Chen Yen and Chun-liang Chen; you both have offered so much help and expertise for my development. Yu-Chen thanks for helping me with specific activity and Chun thanks for being a great mentor when I was rotating in the Stauffacher lab and teaching me how to collect data on the home source and the tricks to refining structures in Phenix.

Thank you to all my undergraduates and new Ph.D. rotation students for all of their help with the projects in this dissertation. My students have helped me develop as a mentor and in the process, I've learned so much about myself. Thank you, Zachary Beck and Bethany Manning, for your patience with me as I was just starting out in the program. Thanks, Sarah Zalar, for all of your positivity and passion that you brought to the lab, and for all of your help troubleshooting the expression and purification of FIPV Ubl2-PLP2 in Chapter 5. You brought out my newfound passion for teaching. It was so rewarding seeing you grow as a scientist and you equally allowed me to grow as a mentor. Thank you to Karthik Srinivasan for your astounding work-ethic and keeping up with me and the project like a boss. You truly pushed the project along in Chapter 4 when you help with the mutagenesis of MERS-CoV PLpro mutants. Thank you, Julia Weeder, for your willingness to help and also bringing life into the lab with your talented drawings and smiling face. Furthermore, I'd like to thank Daniel Wesenberg for your ambition and going above and beyond as an undergraduate student. I was so impressed with all of your work over the summer 2018 and your persistence to keep the project moving. You helped me tremendously not only with mutagenesis work, but also purifying and doing preliminary characterization of FIPV and PEDV

PLP2 mutants in Chapter 5. Thanks for your patience with me, for fixing my watch battery, and for all of the laughs over the summer. I'd also like to thank Sarah McGovern for your excitement and passion about crystallography and research. I had a lot of fun with you crystallizing PEDV PLP2-Ub bound complex in the appendix section.

Furthermore, I'd like to thank Dr. Carol Post for the Hockmeyer socials, which brings the structure group together for stimulating talks. Thank you, Dr. Chitta Das, Dr. Angeline Lyon, Dr. Richard Kuhn, and Dr. Christine Hrycyna for serving on my preliminary examination committee. Your patience and guidance throughout the process helped me tremendously.

I'd also like to acknowledge the professors, Dr. Dave Leonard, Dr. Rachel Powers, and Dr. Brad Waller in the Chemistry Department at Grand Valley State University, and specifically my mentor Dave in my undergraduate school. All of your training and teaching allowed me to have the upper hand in graduate school. I was so lucky to have the opportunity to work with such a passionate team that is dedicated to developing new scientists. I cannot thank you enough for taking a chance on me.

Finally, I'd like to acknowledge my friends and family for all the emotional support. My loving and supportive parents, Charlie and Kristen Clasman, who both say to keep "living the dream". My dad for bringing the fun and joy to my life every single day, and my mom for pushing me and believing in me no matter what. I cannot thank them enough for the sacrifices they have made for my education. My mom turned down a position in science and chose me first, so my tenacity and drive to succeed is because of her. I'd also like to thank my loving, goofy, and caring fiancé, Dan Cholger who tells me every day that "I am beautiful and motivated". I am so grateful my best friend was with me along this crazy ride to pick me up when I am down and celebrate with me through all of the milestones. Also, for his amazing cooking and brain food, which has now reached Gordan Ramsey status. I'd also like to thank my cuddly, cute, and fuzzy kitties,

Independence and St. Paddy for the breows, purring, and unconditional love. Thanks to Posty and Jeremy Soule for the jams that keep me going.

I am incredibly thankful for the life-long friends I have met along this journey. Thanks to my friends on the Hockmeyer social committee for their awesomeness, including Matt Therkelsen, Dr. Lenna Peterson, Frank Vago, Elisabeth Garland-Kuntz, Brandon, and Corey for all the fun coming up with songs and themes as well as picking out the best beverages. Special thanks to Matt for being such a kind and helpful friend who gave me tons of writing edits and much needed pep talks. Thanks, Lenna and Kevin Peterson as well as Dr. Nathan Gardner for contributing to discussions while playing magic the gathering. Thanks to Chuck and Lizzy for talking science, playing games, and being one of the nicest people I've met here. Thanks to Adam and May Hamdani for the games, laughs, talks, and ice cream. Thanks to Chemistry Yoga with Cory for helping me stay strong mentally. For Keelan Trull, Veronica Heintz, Dan, May, and Matt for joining in. Lastly, I'd like to thank my best friends and sisters at heart, Korinn and Sabrina Detter, for the endless laughs and talks I needed to recharge my brain for science.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACE2	angiotensin-converting enzyme 2
AMC	7-amino-4-methylcoumarin
APN	aminopeptidase N
AYE	allylamine
BL	Blocking loop
BME	β-mercaptoethanol
BSA	bovine serum albumin
CD	circular dichroism
CDC	Centers for Disease Control & Prevention
CoV	coronavirus
CV	column volume
deISG	deISGylating
DMO	double membrane organelle
DMV	double membrane vesicles
DPP4	dipeptidyl peptidase 4
DPUP	domain preceding Ubl2 and PLpro/PLP2
DTT	dithiothreitol
DUB	deubiquitinating
E	envelope
ERGIC	endoplasmic reticulum-Golgi intermediate compartment
FIPV	feline infectious peritonitis virus
HCoV	human CoV
IBV	infectious bronchitis virus
Interferon	IFN
IRF3	interferon regulatory factor 3
ISG15	interferon-stimulated gene 15
M	membrane
Mac	macrodomain
MAVS	mitochondrial antiviral signaling protein
MDM2	mouse double minute 2

MERS-CoV	Middle East respiratory syndrome coronavirus
MESNa	sodium 2-mercaptoethanesulfonate
MHV	mouse hepatitis virus
MW	molecular weight
N	nucleocapsid
NFκB	nuclear factor κB
nsp	nonstructural protein
ORFs	Open reading frames
OTU	ovarian tumor domain
PA	propargylamine
PDB	Protein Data Bank
PEDV	porcine epidemic diarrhea virus
PLP/PLpro	papain-like protease
pp	polyprotein
RIG-I	retinoic acid-inducible gene I
RMSD	root-mean-square-deviation
S	spike
SADS-CoV	swine acute diarrhea syndrome coronavirus
SARS-CoV	severe acute respiratory syndrome coronavirus
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
STING	stimulator of IFN genes
SUb	ubiquitin recognition subsite
TBK1	TANK-binding kinase 1
TEVp	tobacco etch viral protease
TGEV	transmissible gastroenteritis virus
T_m	thermal melting temperature
TM	transmembrane domain
Ub	ubiquitin
Ubl	ubiquitin-like
USP	ubiquitin-specific protease
WHO	World Health Organization
WT	wild-type

ABSTRACT

Author: Clasman, Jozlyn, R. Ph.D.
Institution: Purdue University
Degree Received: May 2019
Title: Investigating the Substrate Specificity of the Equivalent Papain-like Protease 2 Domain of nsp3 across Alpha- and Beta-Coronaviruses
Committee Chair: Andrew Mesecar

The papain-like protease (PLP) domain of nonstructural protein 3 (nsp3) of the coronavirus (CoV) genome promotes viral replication by processing the CoV polyprotein (protease) and also antagonize innate immune responses by deubiquitinating (DUB) and deISGylating (deISG) host substrates. Selectively removing the DUB/deISG activities of PLP while keeping the protease activity intact is a potential strategy for designing a live attenuated virus. However, it is unclear in the literature the precise mechanism by which PLPs support CoV evasion of the innate immune system. Deciphering the substrate specificity of PLPs for host ubiquitin (Ub) and interferon stimulated gene 15 (ISG15) can therefore help in the design of PLP mutants that selectively lack one activity for evaluating the DUB and deISG mechanism in CoV pathogenesis and replication.

In this dissertation, we investigate the structure and function of the single PLP (PLpro) from beta-CoVs, severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), which are dangerous viral pathogens that emerged from a zoonotic source to cause infectious disease in the human population. Additionally, we translate the knowledge gained to the equivalent PLP2 from alpha-CoV porcine epidemic diarrhea virus (PEDV) and feline infectious peritonitis virus (FIPV), which cause fatal disease in suckling piglets on industrial pork farms and household cats, respectively. The primary objective of this work is to rationally design PLP mutants across beta- and alpha-CoVs to help attenuate CoV infection, as no antiviral or vaccine exist for human CoVs and the efficacy of PEDV vaccines are an ongoing research topic.

In Chapter 1, different human, animal, and the bat origin CoV strains are introduced. The CoV life-cycle and virion structure are outlined, along with the replicase complex for viral replication. The multidomain nsp3 from alpha- and beta-CoV genomes are also described with a focus on the PLP domain and its proposed cleavage sites of the viral polyprotein. The discovery of the first viral protease DUB and the multiple activities of PLPs are defined, which includes a proposed model of how DUB versus deISG activities may act in the innate immune response. This leads into the therapeutic potential of PLP for an antiviral or live attenuated vaccine, which is followed by the introduction of live attenuated vaccines and the reverse genetics system. Next, proof of concept studies on PLP2 mutants are described and the introduction is concluded by stating the ultimate goal for the design of PLP mutants.

In Chapter 2, we hypothesize that the flanking ubiquitin-like (Ubl2) domain of MERS-CoV PLpro is not required for its enzymatic function. We characterize the specific activity, kinetics, substrate specificity, and inhibition of the PLpro enzyme with and without the Ubl2 domain and reveal that the Ubl2 domain does not significantly alter PLpro function. We determine the structure of the core PLpro, smallest catalytic unit to 1.9 Å resolution and observed no structural changes compared to the wild-type. Additionally, we demonstrate that a purported MERS-CoV PLpro inhibitor is nonselective in non-reducing conditions and should not be pursed for therapeutic use. We show that the core PLpro enzyme i.e. without the Ubl2 domain is a stable and robust construct for crystallization and is also thermally stable based on thermal melting studies with utility for structure-based drug design.

In Chapter 3, we shed light on the specificity of SARS-CoV PLpro towards Ub versus ISG15 by characterizing the specific activity and kinetic parameters of SARS-CoV PLpro mutants. In addition, the structure of SARS-CoV PLpro in complex with the C-terminal domain of ISG15 is determined and compared with the Ub-bound structure. Based on the structure and kinetic results,

the altered specificities of SARS-CoV PLpro mutants Arg167Glu, Met209Ala, and Gln233Glu are compared with the wild-type. Arg167Glu mutant exhibits DUB hyperactivity and is expected to adopt a more favorable interaction with the Arg42 of Ub. At the same time, ARG167GLU contains a shorter side-chain that hinders interaction with the unique Trp123 of ISG15 for deISG activity compared to the wild-type. These results aid in the development of SARS-CoV PLpro mutants that have directed shifts in substrate specificity for Ub versus ISG15.

In Chapter 4, the process and antiviral activity of ISGylation is reviewed and how viruses can modulate host-derived versus virus-derived machineries to counteract ISGylation for viral infection. MERS-CoV PLpro is cross-reactive for Ub, but less is known about its specificity towards ISG15. In this study, we determine the structure of MERS-CoV PLpro bound with ISG15 to 2.3 Å resolution and reveal a small hydrophobic pocket of ISG15 that consists of P130 and W123, which differs from Ub hydrophobic patch. We design and determine the kinetic parameters for 13 PLpro mutants and reveal that MERS-CoV PLpro only has a single ubiquitin recognition (SUb1) site. Kinetic studies show that removing the charge of the R1649 greatly enhances DUB/protease activity while mutating in an Arg near R42 of Ub or ISG15 hydrophobic region is detrimental to both DUB/deISG activities. Kinetic experiments and probe-reactivity assays showed that Val1691Arg, Val1691Lys, and His1652Arg mutants are drastically reduced DUB/deISG activities compared to the wild-type. Overall, MERS-CoV PLpro mutants with alter kinetic profiles will be useful for discovery tools and DUB/deISG deficient mutants are great candidates for removing host cell antagonism activity by PLpro for live attenuated vaccines.

In Chapter 5, the goal is to translate the knowledge gained in Chapters 2-4 on beta-CoVs PLpro and evaluate the substrate specificity of alpha-CoVs FIPV and PEDV PLP2 for mutagenesis experiments. First, we design and purify the core PLP2 enzymes for kinetics. PLP2s are efficient DUBs that prefer Ub to ISG15 *in vitro*, and this preference is conserved in beta-CoV MHV PLP2

as well as alpha-CoV NL63 PLP2. We determine the structure of alpha-CoV PEDV PLP2 to 1.95 Å resolution and reveal the unique Zn-finger coordinating Cys_3-His arrangement of the alpha-CoV genus that differs from past beta-CoV PLP crystal structures. To determine residues of the SUb1 site, we generate a homology model of FIPV PLP2 and overlay our PLP2 structures with MERS-CoV PLpro bound with Ub. In addition, we create electrostatic surface maps across coronaviral PLP subfamilies to evaluate the charge distribution of the SUb1 for the rational design of several FIPV and PEDV PLP2 mutants. We evaluate the turnover of PLP mutants for FRET-based substrates and reveal that $His101Arg^{FIPV}$ and $Asn101Arg^{PEDV}$ are drastically reduced in Ub-AMC activity while their peptide activities are within 2-fold of the wild-type. These mutants show delayed reactivity for Ub probes and no longer cleave Ub-chains displaying isopeptide bonds compared to the wild-type. Results from this study reveal a hot spot in both alpha- and beta-CoVs that can be used to selectively remove DUB activity of PLPs for generating a DUB deficient PLP enzyme.

In this dissertation, we investigate the substrate specificity of PLPs across alpha- and beta-CoVs and develop a fingerprint for Ub and also shed light on ISG15 recognition. Specifically, hot spots were identified in the SUb1 site of different PLPs, which recognize R42 and hydrophobic Ile44 of Ub. Position 97-98 of PLPs can be used to remove DUB activity by substituting an Arg, but usually effect protease function. Substituting an Arg at position 101 and 136 of coronaviral PLPs serve as the best strategy to remove DUB function while not hindering active site functionality. The DUB/deISG deficient mutants described will be useful for inhibiting the ability of PLPs to function in the innate immune response. Ultimately, this work provides a guide for identifying attenuating mutants in existing CoVs for live attenuated vaccines and also a blueprint for engineering PLPs from new emerging CoVs.

CHAPTER 1. INTRODUCTION

1.1 Overview of Coronaviruses

1.1.1 Outbreak of SARS and MERS coronavirus

Coronaviruses (CoVs) are positive-sense single-stranded RNA viruses that can cause serious respiratory and gastrointestinal diseases in both humans and animals. In late 2002, severe acute respiratory syndrome coronavirus (SARS-CoV) emerged in humans and led to devastating outbreak by a human CoV to date. The capacity of SARS-CoV for airborne human transmission created a world-wide panic as necessary public efforts were taken to quarantine infected patients and limit air travel out of China. Although travel advisories were initiated by the World Health Organization (WHO) and the Centers for Disease Control & Prevention (CDC) during this time [1], SARS-CoV infection eventually spread to 26 countries. By the next year, the SARS-CoV had infected 8,000 individuals with a ~10% fatality rate; roughly 800 people succumbed to SARS.

The first reported case of SARS occurred in late 2002 in Guangdong Province of South China, and soon after, the virus was found to circulate in palm civets and raccoon dogs from wild-animal markets in this region. The first animals reported to have been infected by SARS-CoV were the Chinese ferret badger and the raccoon dog [2, 3]. The SARS-CoV lineage was later traced back to a similar virus in horseshoe bats suggesting it originated in bats. In general, bats are great natural reservoirs hosts for viruses because viral infection is asymptomatic [3, 4]. Nipah and Hendra, and likely Marburg and Ebola are other viruses that circulate in bats [5]. Since SARS-CoV was thought to have originated in bats, and jumped to multiple intermediate hosts before emerging in humans, researchers thought that SARS was a suitable pathogen to emerge or re-emerge in the human population [3].

Nearly ten years later another novel virus was identified to cause a SARS-like disease. The SARS-like virus was first isolated from the lungs of a 60-year-old elderly patient from Saudi

Arabia [6]. The virus was originally named human CoV-Erasmus Medical Center (HCoV-EMC) where the viral genome was first sequenced, until it was renamed as Middle East respiratory syndrome coronavirus (MERS-CoV), which is the region where the most cases of MERS were reported. Over 2000 total cases have been reported globally with most of these cases centralized in the Arabian Peninsula. Specifically, >80% of MERS-CoV cases have reported in Saudi Arabia [7]. In 2015, a brief 3-month outbreak also occurred in South Korea, which resulted in 39 deaths among 186 cases [7].

The symptoms of MERS manifest as a SARS-like illness starting with headache, malaise, fever, chills, cough, leading to severe acute respiratory pneumonia of the lungs. Other symptoms often include diarrhea and in serious cases, renal failure, which may explain the higher fatality rate of MERS of ~35% compared to SARS. Older men with comorbidities, based on cases analyzed in Saudi Arabia and South Korea, are at a higher risk of dying from MERS-CoV infection [7]. Like SARS-CoV, MERS-CoV can easily spread through international air travel. In 2014, the CDC reported two cases of MERS-CoV in Indiana and Florida, which were spread by travelers from Saudi Arabia [8]. To date, MERS-CoV has now spread to 27 countries.

Similar to SARS-CoV, MERS-CoV has also been thought to have originated from viruses found in bats, including CoV HKU4 from *Tylonycteris pachypus*, CoV HKU5 from *Pipistrellus*, and CoV from a *Neoromicia* South African bat [9]. Neutralizing antibodies and infectious MERS-CoV were found in several dromedary camel populations in the Middle East and Africa, specifically in Oman, Spain, and Egypt, indicating that camels were primary intermediate hosts [10, 11]. In cell culture, MERS-CoV has shown to be able to replicate in kidney and lung goat cell-lines as well as camel kidney cells [11, 12].

The existing MERS-CoV threatens public health and new emerging or re-emerging CoVs are likely to occur in humans. As of now, there are no antiviral or vaccines available to treat human

CoV infection. This emphasizes the critical need to understand key virus-host interactions that facilitate virus propagation and disease pathogenesis to effectively treat infected patients.

1.1.2 Origin, phylogeny, and taxonomy of Coronaviruses

Coronaviruses, including SARS-CoV and MERS-CoV, belong to the *Coronaviridae* family of the *Nidovirales* order (Figure 1.1), together with viruses from the *Arteriviridae*, *Mesoniviridae*, and *Roniviridae* families. Nidoviruses at large are non-segmented, enveloped, single-stranded positive sense RNA viruses that infect a variety of animal species. The *Mesoniviridae* family, closely related to the *Roniviridae* family, is more divergent because the virus only infects insects (i.e. mosquitos) [13]. Nidoviruses are recognized to produce 3' nested sets of subgenomic mRNAs during transcription [14]. The Latin word, *nido* or "nest" is where the *Nidovirales* name derived from.

Before the SARS outbreak in late 2002, only two human CoVs (HCoVs), 229E and OC43 were known, which were identified in the 1960s [15]. Following the SARS outbreak in 2002, many studies were done on patients presenting acute respiratory symptoms, and not surprisingly, two other HCoVs, NL63 and HKU1, were soon discovered thereafter in 2004 and 2005, respectively. These viruses cause mild upper respiratory tract infections and are responsible for many common colds. Approximately 3-11% of patients with respiratory tract illness are caused by a HCoV strain and most cases occur in infants and immunocompromised adults [16, 17]. Although HCoV-NL63 was discovered in the 2000s, the virus may have diverged from bat CoV strains approximately 500-800 years ago [15].

While there are only six human CoVs, over 40 different animal CoVs have been discovered over the last decade and new viruses continue to arise, including the recent swine acute diarrhea syndrome coronavirus (SADS-CoV) in October 2016. A study in April 2018 revealed that SADS-CoV was related to a bat coronavirus HKU2 [18]. As new CoVs are discovered, new classifications

have followed. Currently, animal, human, and bat CoVs are grouped into four genus types: alpha-, beta-, gamma-, and delta-CoV (Figure 1.1 and 1.2). Beta-CoVs are organized into four groups: A, B, C, and D. Mouse hepatitis virus (MHV), a model system of the coronaviral infection, is found in group A with bovine coronavirus (BCoV), and two human CoVs, OC43 and HKU1. SARS-CoV belongs to group B, and MERS-CoV is a group C beta-CoV together with similar bat CoVs, HKU4 and HKU5. HCoV-NL63, feline infectious peritonitis (FIPV) virus, porcine epidemic diarrhea virus (PEDV), together with transmissible gastroenteritis virus (TGEV) are within the alpha-CoV clade. Additionally, bat coronavirus HKU2 and SADS-CoV are also alpha-CoVs.

PEDV and FIPV cause gastrointestinal and neurological disease in suckling pigs and domestic cats, respectively. The outbreak of alpha-CoV PEDV in May 2013 devastated pork farms across the U.S. and placed a tremendous financial burden on the pork industry [19]. The PEDV strain in the U.S. was genetically related to the Chinese PEDV strain reported in 2011-2012 [20], while PEDV was first isolated in Belgium in 1978 [21]. It has been suggested that the genetic variation of PEDV strains may have arose from sporadic or unpredictable recombination events [22]. FIPV arose from a mutation in the spike protein of a common enteric strain of feline CoV, which enabled the virus to effectively replicate in macrophages and infect cats (see Chapter 5). While feline CoVs cause mild diarrhea in cats and are often asymptomatic, FIPV causes incurable and fatal disease. Approximately 90% of cats in multi-cat households are infected with a feline CoV and 5% of feline CoV cases will mutate to FIPV [23, 24]. PEDV and FIPV are limited to their natural host, but both are similar to alpha-CoV strains to those found in bats.

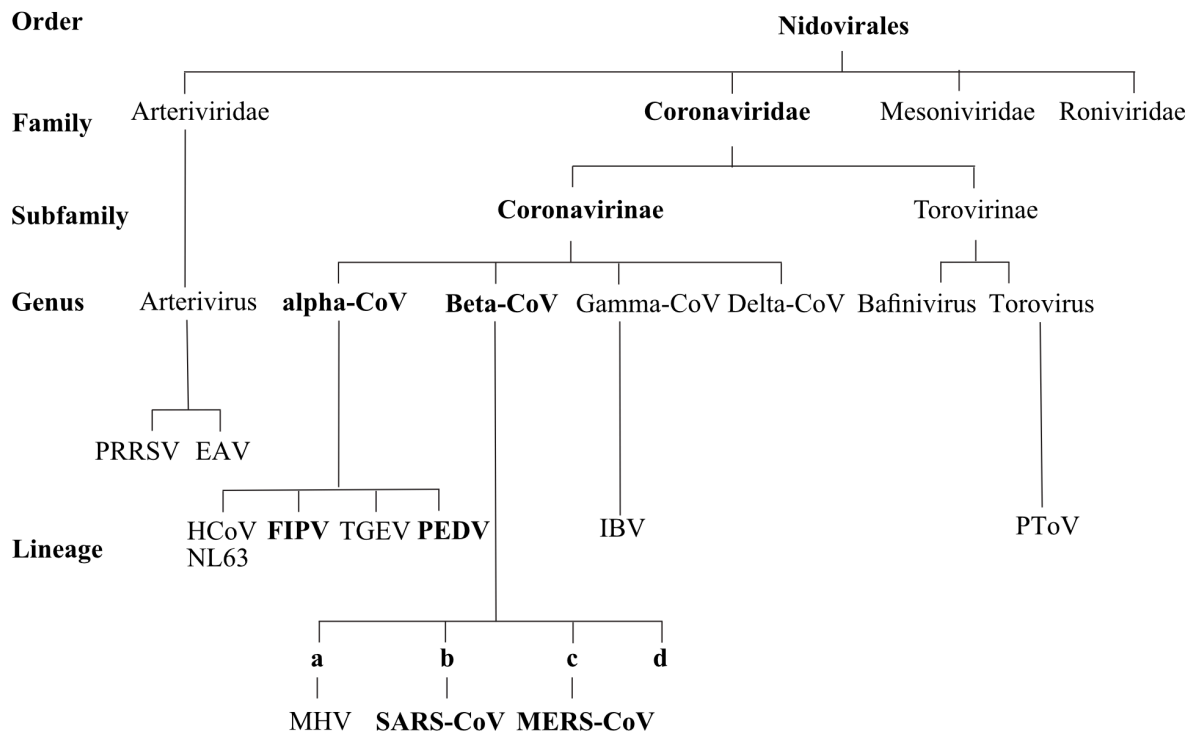


Figure 1.1 Taxonomy of CoVs of the nidovirus order

Different coronaviruses, arteriviruses, and toroviruses classifications are shown within their order, family, subfamily, and genus types. FIPV, PEDV, SARS-CoV, and MERS-CoV are bolded. *Mesoniviridae* and *Roniviridae* family of viruses are not shown for clarity. Figure was adapted from reference [25].

Beyond the alpha- and beta-CoVs are the gamma-CoVs, which infect chickens, turkeys, and also marine life, including the beluga whale and bottlenose dolphin [10, 26, 27]. Probably the most well-known gamma-CoV is the avian infectious bronchitis virus (IBV). IBV causes acute respiratory disease in chickens and outbreaks each year have led to significant economic losses within the poultry industry. Over 50 different serotypes of IBV have been reported in the U.S. since its first isolation in the 1930s, and it appears to have a more divergent genomic organization to other CoVs [28]

Finally, the least studied CoVs reside in the delta-CoV branch. Delta-CoVs infect mostly wild birds, such as the night heron, magpie robin, sparrow, but a pig CoV HKU15 was recently identified in 2014 as a delta-CoV [10, 29-31].

Bat CoVs have been found to be related to both human and animal pathogens, and therefore show a wide range of genetic diversity [32]. Bat caves in Yunnan Province, China have shown to harbor SARS-related CoV strains [33]. Similarly, a strain of MERS-CoV isolated in *Pipistellus hesperidus* in Uganda suggests that a recombination event is likely how MERS-CoV emerged into humans [34]. Genetic variation of these similar bat CoV strains was found in the viral spike protein while other regions of the viral genome showed less divergence. For example, two mutations of bat HKU4 were found within the MERS-CoV spike protein, which enabled the virus to replicate in human cells [35]. This illustrates that slight alterations and mutations of the spike protein may enable bat CoVs to infect humans or other intermediate hosts.

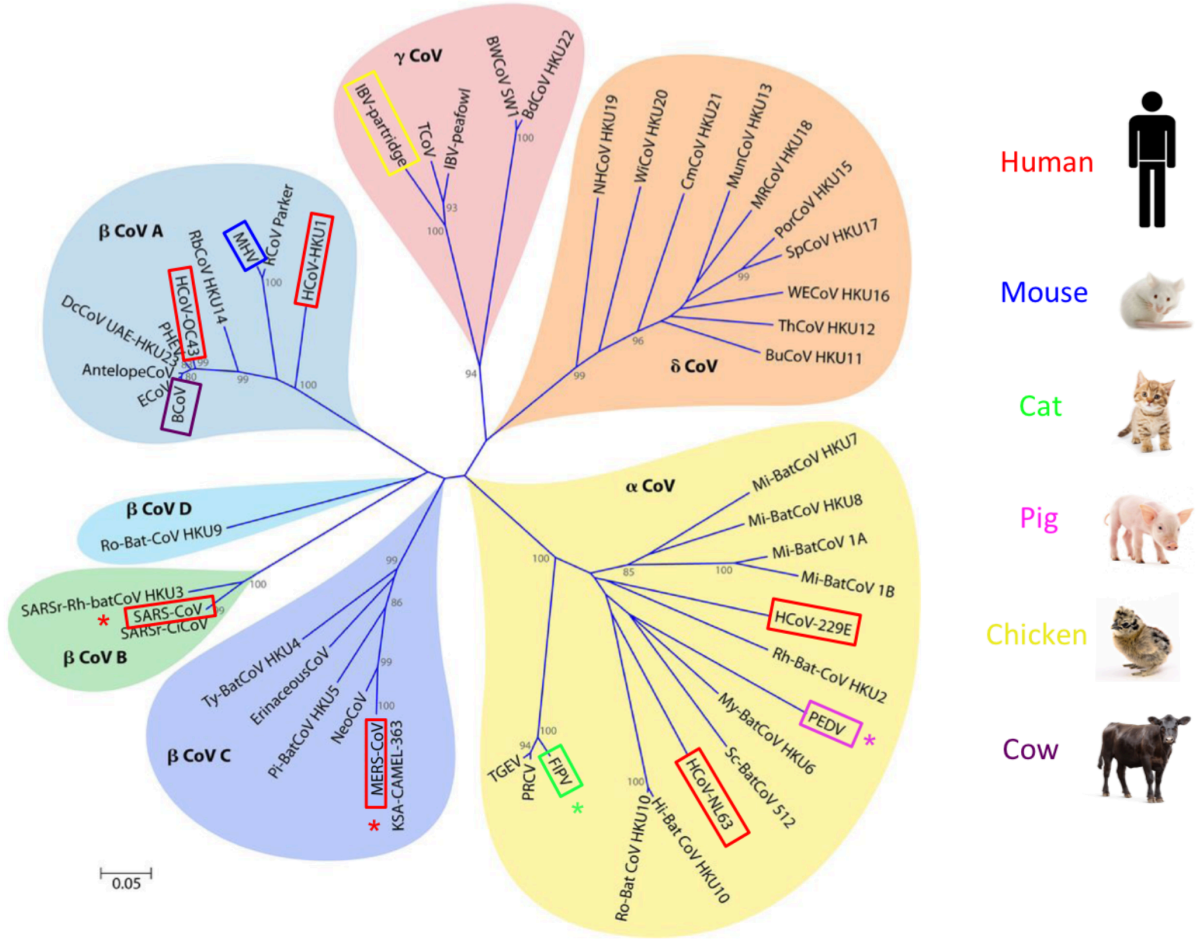


Figure 1.2 Coronavirus phylogeny and broad host range

Phylogenic tree of 50 CoVs. All human CoVs and selected animal CoVs are boxed in the phylogenic tree and colored according to the host the CoV is known to infect. The CoVs focused on in this dissertation, SARS-CoV, MERS-CoV, FIPV, and PEDV are starred. Color coding is as follows: Red, HCoVs; blue, MHV; green, FIPV; pink, PEDV; yellow, IBV; purple, BCoV. Other pig CoVs, such as TGEV and PRCV are not colored for clarity. These sequences only partially contained the RNA-dependent RNA polymerase. The scale bar represents the number of substitutions per 20 nucleotides. Figure was adapted from reference [10].

1.1.3 Coronavirus genome and virion structure

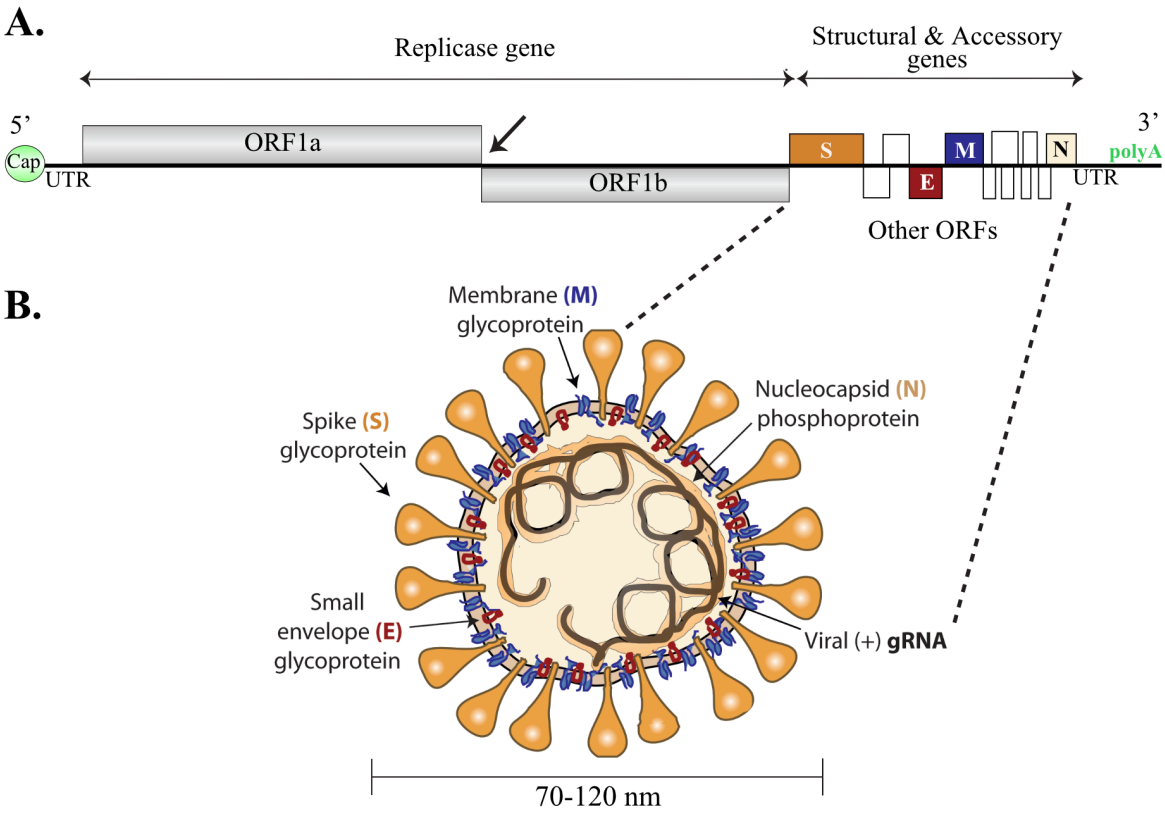
CoVs have the largest viral genomes among RNA viruses, ranging from 26-32 kb in size. The genome mimics host mRNA with a 5'-cap and 3'-poly-A tail to enable translation by host ribosomes (Figure 1.3.A). Similarly, the genome includes untranslated regions (UTR) at the 5'- and 3'-ends. UTRs adopts a series of different RNA structures (i.e. multiple stem-loops and a pseudoknot) that are important in transcription and viral replication [13]. The largest part of the viral genome is the replicase gene (~20 kb), which spans two-thirds of the viral genome between two large open reading frames (ORFs), ORF1a and ORF1b. The replicase gene encodes for the 16 nonstructural proteins (nsps) involved in virus replication. The CoV genome is polycistronic. It contains a "slippery" RNA sequence and a pseudoknot between ORF1a and ORF1b, which causes a ribosomal -1 frame shift [13]. This ribosomal frame shift enables two polyproteins, pp1a and pp1ab, to be expressed from the replicase gene. Finally, structural genes (~10 kb) are at the 3'-end of the RNA genome and contain additional ORFs that encode for proteins of virion structure. These ORFs are transcribed into subgenomic RNAs by the replicase complex. Accessory genes are also in this structural region, which may be important in viral pathogenesis [13].

CoV virions are enveloped, usually pleomorphic or spherical, with a diameter spanning from ~70 to 120 nm in length (Figure 1.3.B). The structural glycoproteins of the virion, encoded at the C-terminal end of viral genome, include the spike (S), envelope (E), and membrane (M). The spike protein (~150 kDa) mediates host cell attachment, which is essential for membrane fusion and entry of the virion into the host cell [13, 15]. M protein (~25-30 kDa) likely gives the virion its shape. It is approximately 6-8 nm thick and contains three transmembrane domains [13]. The E protein is the smallest protein (~8-12 kDa) of the virion and the least abundant. Although it is small, the E protein appears to have many functions, including viral assembly and release [13]. Finally, the nucleocapid (N) phosphoprotein found within the particle packages the coronaviral

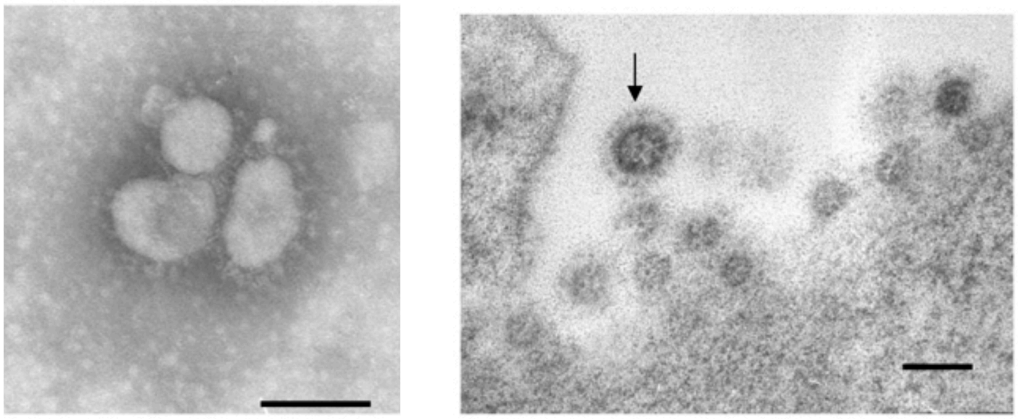
positive strand genomic RNA. Figure 1.3.C shows a typical SARS-CoV virion imaged by transmission electron microscopy [36]. Club-like glycoprotein spikes projecting from the enveloped membrane of the virion adopt a “crown” or halo-like appearance. The Latin word, *corona*, “crown” is where the name coronavirus was derived from. Due to their unique morphology, SARS virions are easily detected via chest radiography or in other bodily secretions (i.e. feces, urine, tears) of patients [3].

Figure 1.3 CoV genome organization and SARS-CoV virion morphology

(A) The general RNA genome organization of CoVs is shown as a schematic diagram. The large replicase component of the genome is shown at the 5'-end with the flanking structural and accessory region at the 3'-end. The arrow between ORF1a and ORF1b of the replicase represents the region where ribosomal frame shift occurs. (B) The structural components of the virion are shown as a cartoon representation. S protein, orange; M protein, blue; E protein; red; The N protein (tan) contains the genomic RNA (brown). Panel A-B was adapted from reference [37]. (C) Electron micrograph of SARS-CoV under high (left) and low (right) magnifications was adapted from reference [36]. Scale bar represents 100 nm. A typical SARS-CoV virion is indicated by the arrow, which shows the virus outside of an infected cell in close proximity to the cell membrane.


A.
Replicase gene
Structural & Accessory genes
5'
3'
Cap
UTR
ORF1a
ORF1b
S
E
M
N
polyA
UTR
Other ORFs
B.
Membrane (M) glycoprotein
Nucleocapsid (N) phosphoprotein
Spike (S) glycoprotein
Small envelope (E) glycoprotein
Viral (+) gRNA
70-120 nm

C.



1.1.4 Coronavirus life-cycle and formation of the replicase complex

The CoV life-cycle starts with attachment of virion S protein of the virion to the cellular receptor for the virus to gain entry into the host cell (Figure 1.4). The S1 region of the S protein generally contains a receptor binding domain that facilitates interaction with the receptor [13]. The S protein and the receptor are then separated by proteases at two distinct cleavage sites, which exposes a fusion peptide. The fusion peptide initiates membrane fusion at the S2 region of the S protein usually within acidic endosomes or in the case of MHV, at the plasma membrane [13, 38].

The spike glycoproteins are variable among CoVs and dictate tissue tropism and host infectivity. The cellular receptors utilized by CoVs are often peptidases, but their interactions rather than enzymatic activity facilitate viral entry [13]. SARS-CoV and HCoV-NL63 share the same host receptor, angiotensin-converting enzyme 2 (ACE2) while many alpha-CoVs, including HCoV-229E, TGEV, PEDV, FIPV, and canine CoV use aminopeptidase N (APN). MHV utilizes mCEACAM, and MERS-CoV enters into human cells via the dipeptidyl-peptidase 4 (DPP4) [13]. In the case of FIPV, the virus effectively utilizes APN receptor to replicate in macrophages. FIPV-infected macrophages have been shown to produce the cytokine THF-alpha, which increases FIPV replication by elevating the expression of the APN receptor [39].

Cells that do not present the appropriate cellular receptor for CoV entry can be problematic for researchers studying basic CoV replication as well as analyzing potential antivirals and vaccines in animal models. For example, MERS-CoV is unable to interact with mouse DPP4 to enter mice cells. To address this problem, a transgenic mouse model was developed using the adenoviral vector system to express the human DPP4 receptor gene in the lungs of mice [13, 40, 41]. This is a promising system for researchers to help develop therapeutics for human CoVs in animal models with the appropriate cellular receptors.

Following membrane fusion, the positive sense genomic RNA of the virion is distributed into the cytoplasm, where it serves as the blueprint for host ribosomes to produce viral pp1a and pp1ab. Due to the ribosomal frame shift, pp1a and pp1ab expression is variable. Studies have suggested that the incidence of the ribosome frameshifting is likely high ~25%, but it is still unclear why the viruses utilize the frameshift to control genome expression [13].

Once the large overlapping polyprotein is generated, the pp1a/pp1ab is processed into 16 nsps by coronaviral cysteine proteases. Either one or two papain-like proteases (referred to either as PLPs or PLpro), cleave at three sites of the viral polyprotein [42, 43]. Alpha-CoVs, including HCoV-NL63, HCoV-229E, PEDV, FIPV, and TGEV, all have two PLPs while only group A beta-CoVs have two PLPs (i.e. HCoV-HKU1, HCoV-OC43, and MHV). SARS and MERS-CoV of group B and C encode for a single PLpro, along with gamma-CoV IBV, beluga SW1 and delta-CoV munia HKU13. While PLPs cleave the amino-terminal portion, the main protease or chymotrypsin-like protease (3CLpro) cleaves at the remaining 11 sites from 3 to 16 of pp1ab. Following maturation of pp1ab and release of nsps is the formation of the large membrane-bound replicase-transcription complex, which facilities RNA replication as well as transcription of the nested set of subgenomic RNAs necessary for the virion structure (Figure 1.4.B).

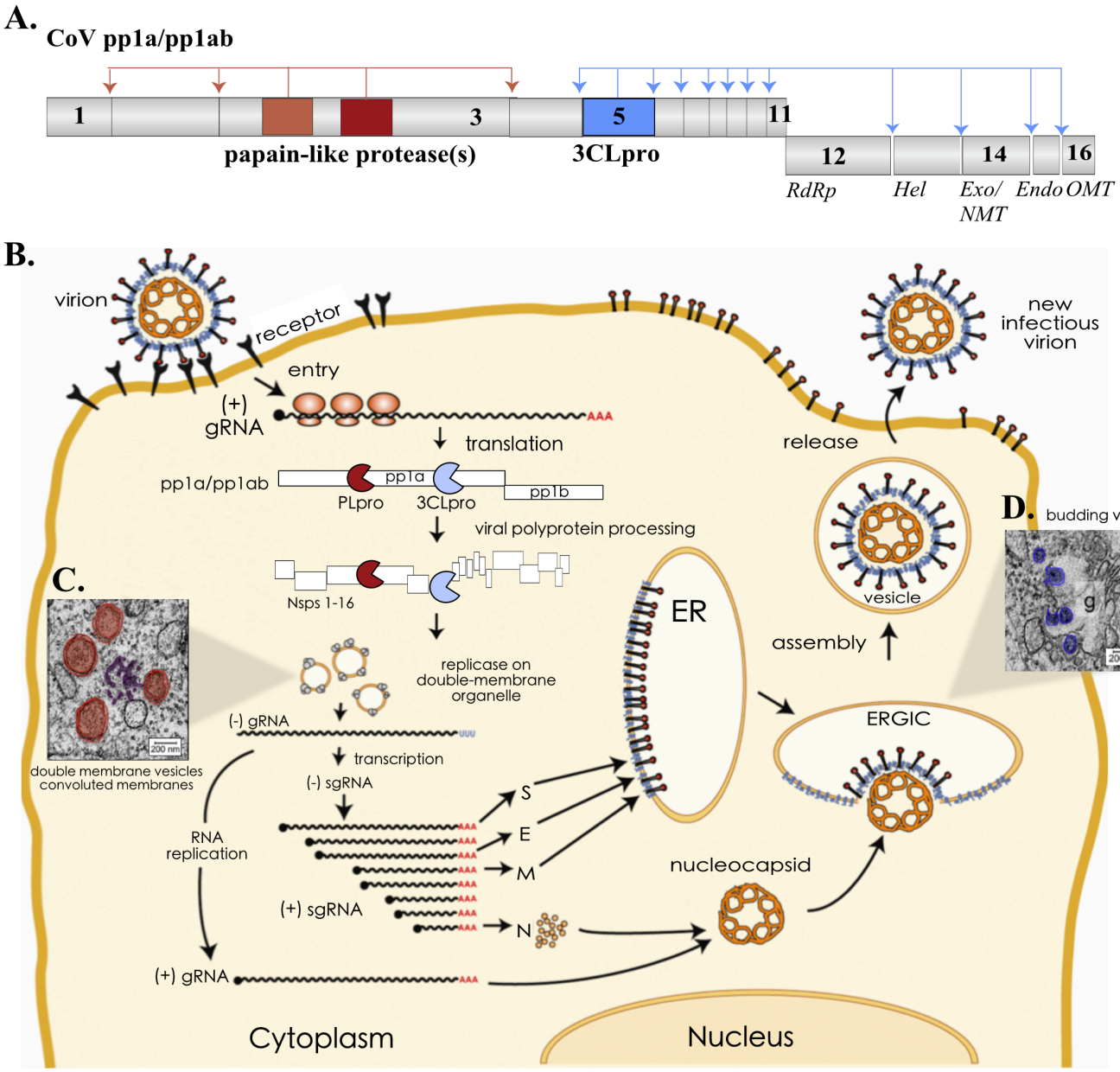
The replicase complex form in double-membrane replicative organelles (DMOs) at the endoplasmic reticulum (ER), which contains two lipid bilayers about 20 nm thick. Recent electron microscopy and tomography studies have revealed that DMOs consist of membrane paired convoluted membranes (CM) and spherules, as well as double membrane vesicles (DMVs) [43]. An example electron micrograph depicting defined DMVs and folded CMs of the ER during MHV replication are shown in Figure 1.4.C [43]. In addition, studies with IBV provided evidence that spherules form from ‘zippered’ cisternae of ER membranes [44]. Nsp3, nsp4, and nsp6, are released by nsp5 and are required for the formation of the DMO, likely through nsp-nsp and nsp-

host interactions on both sides of the membrane lumen [43]. While many studies have revealed structures of the DMO, the precise function of these structures is still unclear, and their organization may vary between viruses. For example, studies with a pathogenic IBV strain showed that a low amount of spherules structures is not necessarily indicative of low RNA production [45]. Nevertheless, evidence points to DMOs as the primary site of RNA replication.

While nsp3-nsp6 are important in forming the DMO, other nsps contain enzymes that function to promote RNA replication (Figure 1.4.A). For example, nsp12 harbors the important RNA-dependent RNA polymerase (RdRp) domain for synthesizing RNA, and nsp13 contains the RNA helicase domain for unwinding double RNA strands. Furthermore, nsp14 encodes the exoribonuclease for proofreading the large RNA genome and maintaining fidelity during replication [15]. In addition, two methyltransferases of nsp14 and nsp16, respectively are important in maintaining the 5'-cap to protect against immune system detection [43]. Finally, an enigmatic endoribonuclease (EndoU) is encoded in nsp15. EndoU was first identified as part of the replicase complex but recent studies in Susan Baker's lab showed that EndoU is a key virulence factor that plays an important role in evading host innate immune sensors and reduces apoptosis in macrophages [46, 47]. Specifically, EndoU is suggested to shield dsRNA intermediates from host dsRNA sensors in DMVs. However, the precise mechanism of action of EndoU is still under investigation.

Figure 1.4 CoV polyprotein organization and general scheme of the CoV life-cycle

(A) The CoV polyprotein 1ab altogether contains 16 nsps. The papain-like protease (red) of nsp3 and 3CLpro of nsp5 (blue) are colored along with their respective cleavage sites across pp1a/pp1ab. Depending on the virus, either one or two papain-like protease(s) are/is located in nsp3 to cleave sites 1-3. Important enzymes for viral replication are also labeled in nsps 12-16. RdRp, RNA-dependent RNA polyperase; Hel, RNA helicase; Exo, exoribonuclease (ExoN); NMT, N7-methyltransferase; OMT, 2'-O-methyltransferase. A newly discovered virulence factor is also shown in nsp15. Endo, endoribonuclease (EndoU). (B-D) First, the spike protein of the CoV virion attaches to the appropriate cellular receptor to initiate membrane fusion and entry into the host cell. The positive sense genome RNA, (+) gRNA, of the CoV is then released into the cytoplasm and directly translated by host ribosomes to generate two polyproteins, pp1a and pp1ab. These pp1a and pp1ab are matured into 16 nsps by virus-encoded cysteine proteases (PLpro, red; 3CLpro, blue). Cleaved nsps are released to form the membrane-bound replicase complex, which occurs in double-membrane organelle structures, including double membrane vesicles (DMVs, red) and convoluted membranes (CM, purple). (C) Electron micrograph depicting DMVs and CMs in MHV-infected cells is from reference [43]. Next, the replicase complex generates (-) gRNA for RNA replication to produce more (+) gRNA. In addition, the replicase complex generates a nested set of subgenomic RNAs (sgRNA), which encode for the structural proteins of the virion, including the S, E, and M proteins. The structural proteins are inserted into the host ER and move to the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) through the secretory pathway. The N protein is also translated for packaging of the newly transcribed (+) gRNA. The N protein and (+) gRNA form the nucleocapsid, which combines with the structural proteins in the ERGIC to assemble the virion. The mature virion finally exits the cell via exocytosis. (D) Electron micrograph of budding and new virions in MHV-infected cells during assembly and release. Cartoon figure in panel B was adapted from [37, 48].


A.
CoV pp1a/pp1ab
1
3
5
11
12
14
16
papain-like protease(s)
3CLpro
RdRp
Hel
Exo/ NMT
Endo
OMT
B.
virion
receptor
entry
(+) gRNA
translation
pp1a/pp1ab
pp1a
PLpro
3CLpro
pp1b
viral polyprotein processing
Nsps 1-16
replicase on double-membrane organelle
(-) gRNA
transcription
(-) sgRNA
S
E
M
N
(+) sgRNA
RNA replication
(+) gRNA
nucleocapsid
ER
ERGIC
assembly
vesicle
release
new infectious virion
C.
200 nm
double membrane vesicles convoluted membranes
D.
budding virions
g
200 nm
Cytoplasm
Nucleus

After the structural proteins of the virion are produced by the replicase complex, they assemble in the ER and travel along the secretory pathway to the endoplasmic reticulum-Golgi intermediate compartment (ERGIC). The structural proteins embedded in the ERGIC bud with the nucleocapsid (N protein packaged with new gRNA to form the mature virion. The M and E proteins appear to be critical for forming new virions, N protein enhances their formation while the S protein is not required [13]. Finally, the mature virion is transported in a smooth-walled vesicle to the cell surface where it is released to the extracellular space via exocytosis to continue CoV infection. Figure 1.4.D shows an electron micrograph of immature and mature MHV virions during assembly and release.

1.2 Multidomain structure of Nsp3

The PLP domain resides in the largest nsp of the pp1ab, nsp3. Nsp3 is a large multidomain protein that is roughly 200 kDa in size and contains 1500-2000 amino acids (Figure 1.5.A). The structure of nsp3 has been limited to single domains structures. The domain organization of nsp3 is shown in Figure 1.5.A for beta-CoVs and alpha-CoVs, using SARS-CoV and HCoV-NL63 nsp3 as examples. As the functions of each domain of nsp3 have been characterized and better understood, the nomenclature of each domain has changed [49].

The first domain of nsp3 is the N-terminal ubiquitin-like 1 domain (Ubl1), which is well-conserved in CoVs. The Ubl1 has been implicated in binding single RNA strands and nucleocapsid N protein and has been suggested to act as a switch between viral transcription and replication. Studies with SARS-CoV suggest it may also bind AUA repeats in the 5'-UTR [49]. After Ubl1 is the Glu-rich acidic region i.e. the hypervariable region (HVR), which is present in all CoV nsp3. The function of the HVR is currently unknown, but its acidic character could potentially facilitate protein-protein or metal-ion interactions.

In alpha-CoVs and group A beta-CoV MHV, the next domain is the papain-like protease 1 (PLP1). PLP1 is absent in beta-CoVs, including MERS- and SAR-CoV, and gamma-CoVs. PLP1 domain cleaves one or two of three sites of the N-terminal portion of the viral polyprotein together with the PLP2 domain [42]. In the case of TGEV, PLP1 cleaves the nsp2/nsp3 site and recognizes a KMGG cleavage sequence [50]. TGEV PLP1 has deubiquitinating (DUB) activity towards K48- and K63-linked chains, similar to ortholog PLP2s [51], and has recently been shown to antagonize IFN-β [52].

The next domain following either the PLP1 or HVR is the macrodomain 1 (Mac1) i.e. X domain, ADP-ribose-1'-phosphatase (ADRP), which is present in all CoVs. Recent studies suggest that ADRP activity is probably not the physiological function of Mac1. Mac1 has been shown to act as a hydrolase and to reverse mono-ADP-ribosylation (MARylation) or poly-ADP-ribosylation (PARylation) by removing single or multiple ADP-ribose(s) from proteins. ADP-ribose is a post-translational modification that has been implicated in many cellular responses, such as the host innate immune defenses i.e. NFκB signaling. While Mac1 is not involved in viral replication, its de-MARylation and de-PARylating could be an additional mechanism CoVs use to counteract innate immune responses. A Beaudette strain of IBV contains a glycine to serine mutation in a triple glycine motif of Mac1, which is a region required for binding ADP-ribose. The Beaudette strain of IBV is attenuated compared to a M41 strain, which does not have a mutation in Mac1. It has been hypothesized that this mutation in Mac1 could have produced the attenuating phenotype exhibited by the Beaudette strain [49]. A recent study, however, suggests that this mutation alone is not sufficient to attenuate IBV [53]. Nevertheless, the enzymatic functions of Mac1 as well as its potential for protein-protein interactions still remain to be investigated.

The next three domains of nsp3 are macrodomain 2 (Mac2), macrodomain 3 (Mac3), and the domain preceding Ubl2 and PLpro/PLP2 (DPUP). Before the discovery of DPUP in MHV

[54], these were named “SARS unique” (SUD) domains at the N-terminal, middle, and C-terminal regions (SARS-N, SARS-M, and SARS-C). Mac2, Mac3, DPUP only seem to exist in beta-CoV genus while this region in alpha-CoVs has yet to be discovered. Mac2/3 do not bind ADP-ribose like Mac1, but their lysine patches have been shown to bind oligo(G) stretches; DPUP may fine tune these oligo(G) interactions. In the case of SARS, Mac3 is essential for viral replication [49]. These domains may mediate interaction with other nsps as well as the C-terminal regions of nsp3 itself.

The next two domains of nsp3 are the ubiquitin-like domain 2 (Ubl2) followed by either PLpro or PLP2. The function of the Ubl2 domain is enigmatic, but it may modulate the function of PLpro. PLpro has multiple enzymatic functions in processing the viral polyprotein and removing ubiquitin (Ub) and interferon-stimulated gene 15 (ISG15), from targeting proteins, i.e. deubiquitinating (DUB) and deISGylating (deISG) activities, respectively, for the suppression of the innate immune responses in CoVs. The precise mechanism by which PLpro antagonizes the innate immunity, such as interferon (IFN) and NFκB signaling, through its DUB and/or deISG activity is still unclear.

PLPs are conserved among CoVs and arteriviruses, including equine arteritis virus (EAV) and porcine reproductive and respiratory syndrome virus (PRRSV). N-terminal of PLP2 in the viral polyprotein, PRRSV contain additional papain-like protease domains, termed α/β, in its viral genome [42]. EAV only contains Pβ and PLP2. Toroviruses and bafiniviruses of the family *Coronaviridae* are also predicted to have papain-like protease domain in their viral genome [43].

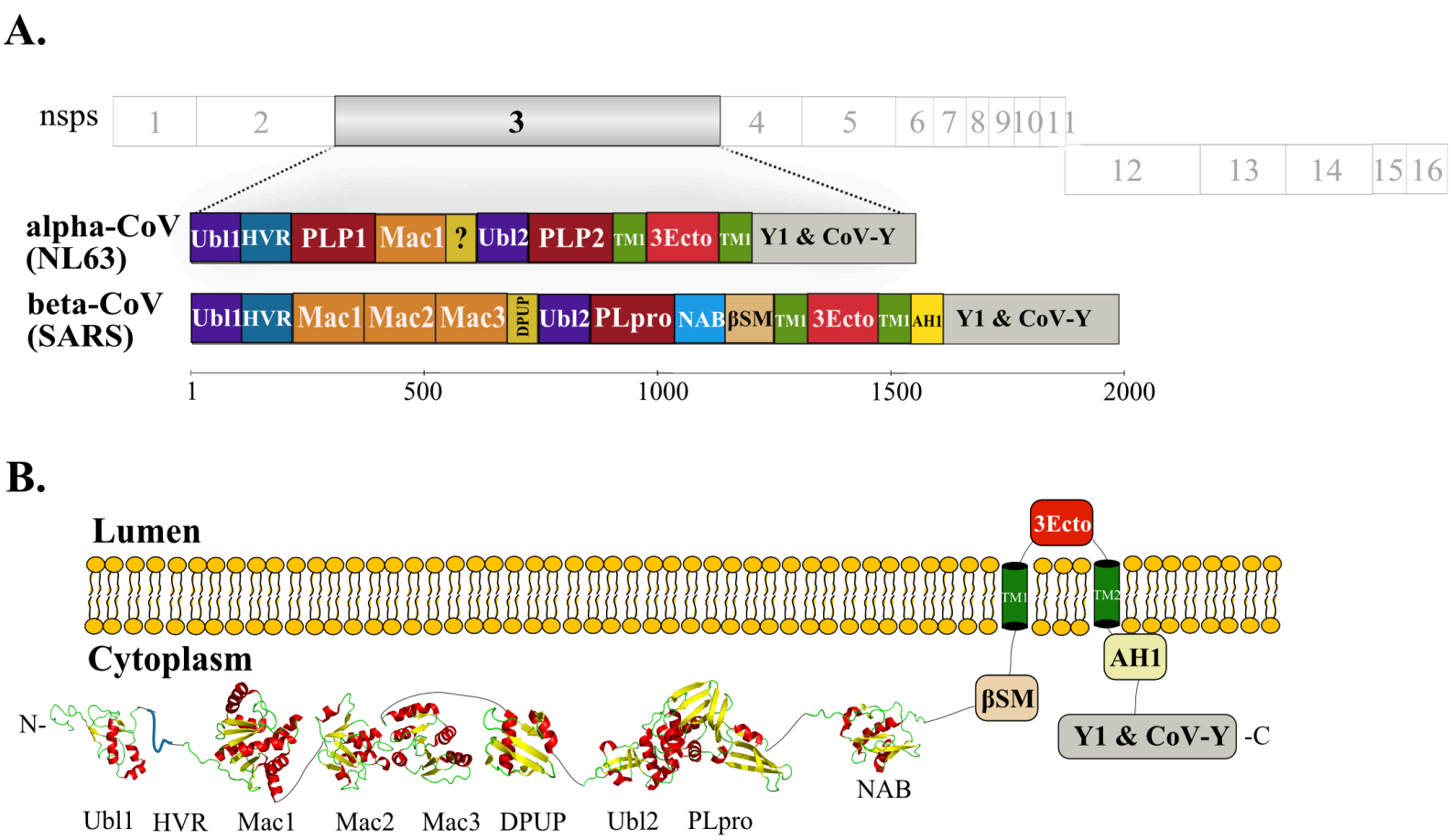


Figure 1.5 Multidomain organization of the nsp3 in alpha- and beta-CoVs genomes

(A) CoV pp1ab with nsps labeled 1-16 is shown highlighting the nsp3 domain. Each domain of nsp3 in alpha- and beta-CoV genus are labeled accordingly using HCoV-NL63 and SARS-CoV as representative virus from each clade. Nomenclature of each domain follows the recent review of nsp3 [49]. (B) Structural model of SARS-CoV nsp3 embedded in the membrane of DMVs from the ER. Two transmembrane domains TM1 and TM2 pass through the membrane. 3Ecto domain is shown on the luminal side while other domains are shown in the cytosol. AH1 is likely attached to the membrane. PDB X-ray/NMR structures determined to date are shown as a ribbon while structures not determined yet are shown as a box, colored according to panel a. Panel b was adapted from reference [54].

A gene for porcine torovirus papain-like protease gene (ToV-PLP) that had DUB and deISG activity as well as the ability to suppress IFN-β expression was recently discovered in a novel enterovirus species G strain isolated from neonatal pig presenting diarrhea [55]. These results stress the dangers of cross-order recombination between viruses of the order *Nidovirales* and order *Picornavirales* in the pig population and possible dangers of this occurring in humans for new emerging viruses.

In the case of beta-CoV SARS and MERS, a single PLpro cleaves the pp1a/pp1ab at all three nsp1/nsp2, nsp2/nsp3, and nsp3/nsp4 sites by recognizing the LXGG motif (residue positions P4-P1; Figure 1.6), which is similar to the RLRGG sequence at the C-terminus of Ub and ISG15. Gamma-CoV IBV and delta-CoV bulbul (song bird) contain a single PLpro that cleaves at two sites nsp2/nsp3 and nsp3/nsp4, due to an absence of the nsp1 domain [42].

For viruses that have two PLPs, the cleavage patterns tend to differ and sometimes the PLP1 and PLP2 can cleave at the same site (Figure 1.6). In the case of MHV, PLP2 was shown to only cleave at the last nsp3/nsp4 site while PLP1 cleaves the first two sites [42, 56]. For HCoV-NL63 the opposite is true; PLP2 has been shown to cleave at the last two sites while PLP1 alone cleaves at the first site [57]. In the case of HCoV-229E, both PLPs cleave at the first two sites. PLP1 is more efficient at cleaving the first nsp1/nsp2 site while PLP2 prefers the second nsp2/3 site [58]. From these results, it is possible that PLP1 and PLP2 may work together or have a preference for particular sites of the CoV polyprotein during processing.

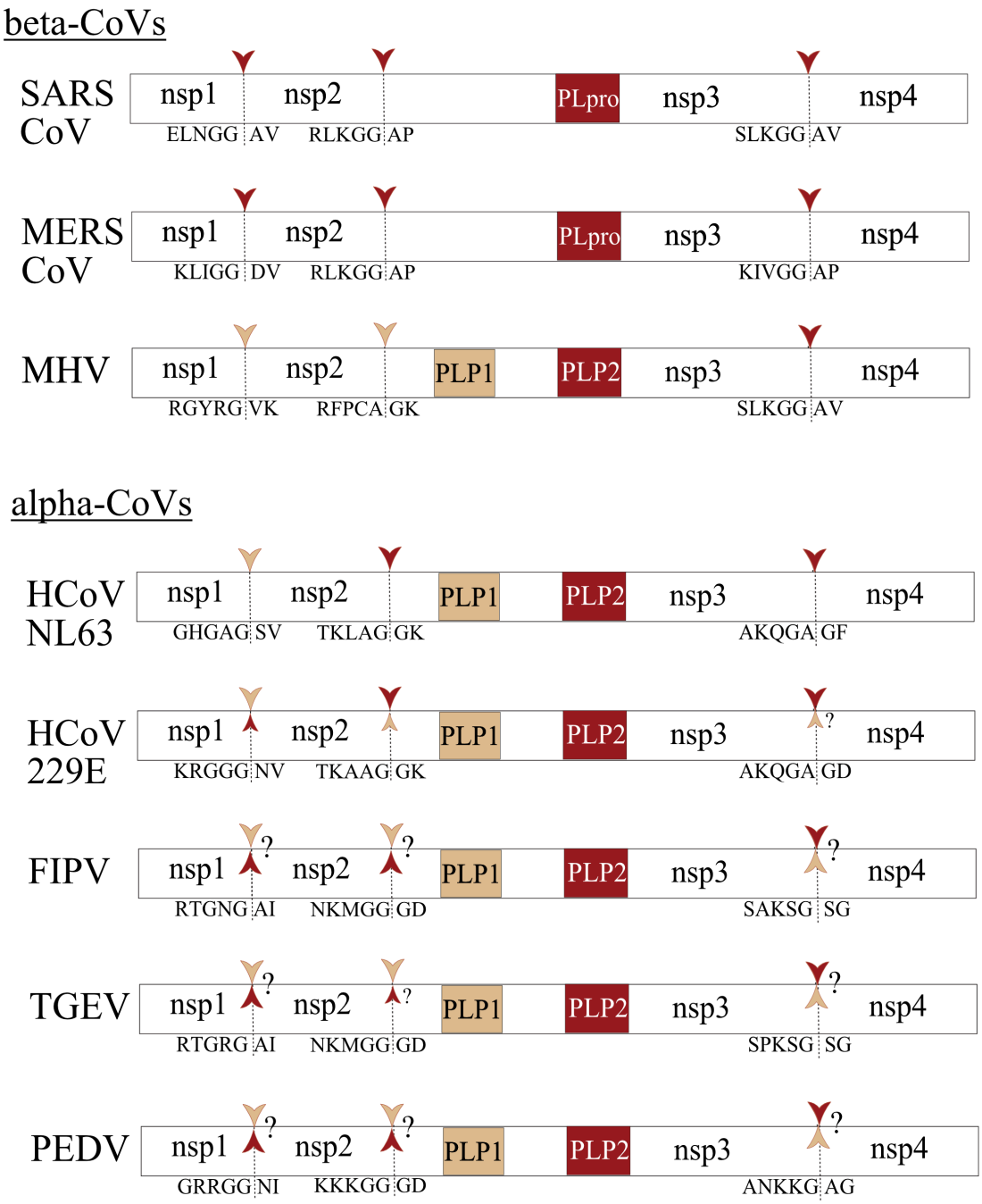


Figure 1.6 Predicted and confirmed cleavage sites of PLPs from different alpha- and beta-CoVs

The nsp1-nsp4 region of the viral polyprotein is shown with cleavage sites of either PLpro (red) or PLP1 (tan) and PLP2 (red). Question marks are indicated where cleavage site specificity is unknown. Below the viral polyprotein is the P5-P4-P3-P2-P1-P1'-P2' residues of the cleavage site. These amino acid sequences were determined either by bioinformatics and mass spectroscopy.

Beyond the PLpro domain is the nucleic acid-binding (NAB) domain and the beta-CoV specific marker (βSM) i.e. nsp3e. Gamma-CoVs also contain a γSM, but alpha-CoVs and delta-CoV do not have a specific marker [43]. NAB has been shown to bind ssRNA and unwind dsDNA and may have a similar function as Mac3; although there is not a clear connection between these domains. βSM is an intrinsically disordered region with unknown function [49].

The C-terminus of nsp3 weaves through the membrane and is conserved in all CoVs. These domains include the first transmembrane domain 1 (TM1) followed by nsp3 ectodomain (3Ecto) i.e. zinc-finger domain, luminal loop of nsp3, another transmembrane domain 2 (TM2), an amphipathic helix (AH1), and the C-terminal Y1 domain of nidoviruses and CoV-Y of coronaviruses. The function of the Y1 and CoV-Y domains are unknown making their nomenclature not well-defined. SARS-CoV has been the most studied compared to other CoVs, and also has the most structures available in the PDB; however, no structural information is available yet for the C-terminus of nsp3. Using the individual domain structures of SARS-CoV, a model of nsp3 embedded in the membrane was assembled (Figure 1.5.B). The 3Ecto domain is the only domain in the ER luminal space while all the other domains of nsp3 reside in the cytosol. It has been suggested that nsp3 can arrange itself in the luminal space rather and in that case, the 3Ecto would be in the cytosol [43].

The membrane-associated region of nsp3 appears to be important in forming CMs and DMVs with nsp4 [59, 60]. 3Ecto domain was previously called the Zn-finger domain with the potential to bind metal ions; however, this is probably unlikely since this domain only contains two cysteine residues. 3Ecto has been shown to have N-glycosylation sites. It is possible that these N-glycans may facilitate interactions with nsp4 for DMO formation. Additionally, in the case of SARS-CoV and MHV, the TM1-3Ecto-TM2 region is required for PLpro or PLP2 to process the

nsp3/nsp4 site of the polyprotein. Therefore, the TMs of nsp3 may also be critical for bringing PLPs to their cleavage sites to release nsp3 and nsp4 to facilitate their interaction.

Depending on the construct of nsp3 used in experiments, researchers have identified different binding partners for nsp3. More functional and structural information is therefore required to understand how nsp3 may arrange itself in the membrane to make critical nsp-nsp, nsp-host, and nsp-RNA interactions.

1.3 Discovery of a viral protease DUB with a USP-fold

Before 2006, PLP was thought to only have viral polyprotein processing activity and be important for the formation of the replicase complex. The first structure of PLpro revealed that the viral protease has a similar thumb-palm-finger architecture as the human ubiquitin-specific protease (USP) family of DUBs [61-63]. PLpros' striking similarity to the USP family was the first clue that the viral USP (vUSP) might be a DUB. Enzymes with DUB activity are encoded by all different types of viruses, including herpesviruses, nairoviruses, enteroviruses, and tymoviruses [64]. These DUBs adopt either a USP or an ovarian-tumor domain (OTU) fold. Figure 1.7 shows the structure of SARS-CoV PLpro in comparison with human USP7 and USP14 (RMSD of ~3 Å). Human USPs contain two blocking loops (BL1 and BL2) for binding Ub and ubiquitin-like proteins but only the second BL2 loop β-turn is conserved in viral USPs. The BL2 loop accommodates the RLRGG tail of Ub and ISG15 at the interface between the palm and thumb. SARS-CoVs also contain a structural zinc-finger with four cysteine (Cys_4) ligands coordinating a single zinc atom (Figure 1.7.A). The Cys_4 configuration adopted by SARS-CoV and other beta-CoVs, including MERS-CoV and MHV, may not be conserved in some members of the alpha-CoV genus of PLPs.

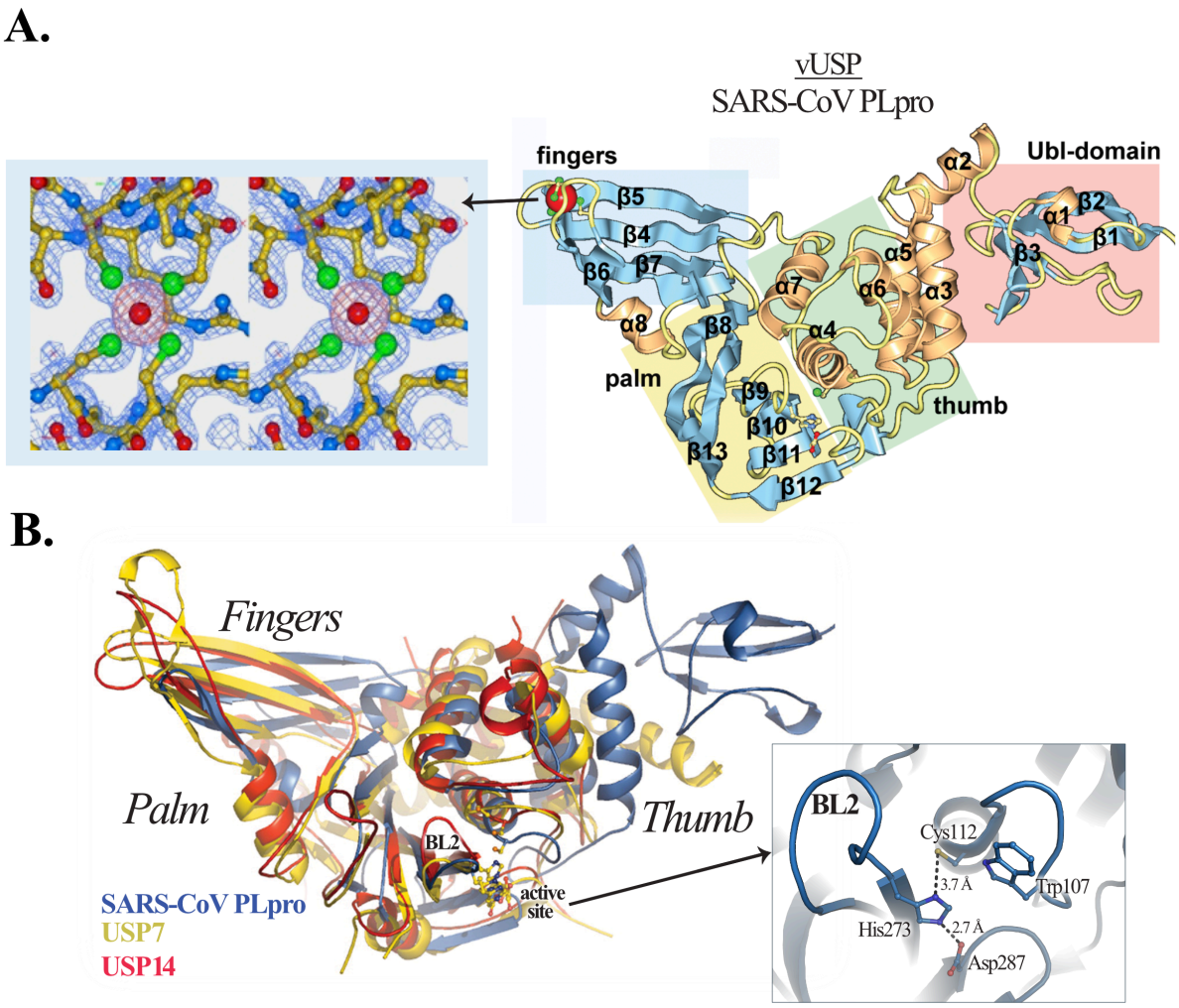


Figure 1.7 Structure of SARS-CoV Ubl2-PLpro revealed a viral USP-fold

(A) Left image –Stereoview of the four cysteine ligands of the fingers domain coordinating the zinc atom in a tetrahedral configuration. Electron density $2F_o$-F_c and F_o-F_c omit maps are contoured at 1.8σ (blue, residues) and 8σ (red, zinc atom), respectively. Right image – Structure of SARS-CoV Ubl2-PLpro with thumb, palm, finger labeled as well the N-terminal Ubl2 domain. (B) Comparison of SARS-CoV Ubl2-PLpro (blue, PDB code 2FE8) with USP7 (yellow, PDB code 1NB8) and USP14 (red, PDB code 2AYN) catalytic domain. Active site and BL2 loop are indicated and magnified on the right. Active site residues are shown as a ball and stick model. Figure was modified from reference [63].

1.3.1 PLP is a multifunctional enzyme

The discovery of the viral protease DUB revealed that PLPs have multiple enzymatic functions. As was previously described, PLPs help mature the viral polyprotein by releasing the first four nsps for the formation of the replicase complex and RNA replication (Figure 1.8.1). Additionally, PLPs are able to reverse the processes of ubiquitination and ISGylation in innate immune response pathways, by acting as a DUB and deISG enzyme to counteract the antiviral state for the host (Figure 1.8).

Ubiquitination and ISGylation are both three-step enzymatic processes that involve the enzymatic action of the E1 activation, E2 conjugation and the E3 ligation to attach Ub or ISG15 to target proteins by an isopeptide bond (Figure 1.8). Ub and ISG15 contribute their C-terminal glycine residues for the isopeptide bond while the lysine residue comes from the target protein. Ub is conserved in eukaryotes and can form polymeric chains while ISG15 is only induced upon viral infection, varies between species, and does not form chains. To form poly-Ub chains, another Ub molecule is covalently linked by its Gly residue with one of the seven Lys (or its N-terminal Met residue) of the former Ub. This is accomplished by the E3 ligase enzyme. There are many different types of Ub-chains that can form and signal for different cellular events [65]. The classic signal for proteasomal degradation is K48 tetraubiquitin while K48 and K63-linked chains are abundant in innate immune response pathways [66, 67]. SARS-CoV, MERS-CoV, and IBV PLpro, HCoV-NL63 and PEDV PLP2, and TGEV PLP1 have been shown to cleave mono-Ub as well as multiple Ub moieties forming chains [51, 57, 68-70]. The DUB activity of PLP can modulate the expression of proteins as well as protein signals in the innate immune responses.

Recent work shows that PLPs may stabilize E3 ligases, which are important for p53 degradation and promoting viral replication. SARS-CoV, MERS-CoV, and HCoV-NL63 can interact with E3 ubiquitin ligase ring-finger and CHY zinc-finger domain-containing 1 (RCHY1),

and HCoV-NL63 PLP2 has been shown to deubiquitinate mouse double minute 2 (MDM2) [71, 72]. RCHY1 and MDM2 both target p53 for degradation, and p53 has been shown to inhibit viral replication; therefore, PLPs could be one mechanism CoVs use to inhibit p53 activity. Although SARS-CoV PLpro interacts with RCHY1, it does not deubiquitinate the E3 ligase. The precise mechanism of how PLPs stabilize E3 ligases in the p53 pathway is still under investigation.

ISG15 structurally resembles a K48-linked diubiquitin (Ub_2) chain with C-terminal and N-terminal ubiquitin-like (Ubl) lobes. The C-terminal Ubl domain aligns well with mono-Ub with an RMSD of 0.967 Å across 418 C_α atoms (Figure 1.8). These Ubl domains in ISG15 are connected by an interdomain linker or hinge rather than an isopeptide linkage as in Ub_2. Each Ubl lobe adopts a putative β-grasp fold, which is present in a variety of Ubl proteins, including SUMO, NEDD8, ATG8, ATG12, URM1, UFM1, FAT10 [73]. FAT10 is another Ubl protein that contains two Ubl domains like ISG15. In the host, USP18 is a selective deISGylase responsible for removing solely ISG15 in cells [74]. Viral PLPs, in contrast, are thought to be cross-reactive in their ability to recognize both Ub and ISG15.

While studies with the human CoVs have determined that MERS-CoV and SARS-CoV are robust DUBs and deISGylases [75-77], beta-CoV MHV PLP2 prefers Ub to ISG15 [54, 78]. In the case of alpha-CoV NL63, PLP2 removes both ubiquitinated and ISGylated proteins in cells [79] and is an efficient DUB and deISGylase *in vitro* [48]. Although HCoV-NL63 PLP2 can recognize both Ub and ISG15, it is the only human CoV to prefer Ub over ISG15. Alpha-CoV FIPV and PEDV PLP2 prefer Ub over ISG15 *in vitro* and may not be physiological deISG enzymes. Therefore, PLP2s may not all be deISGylases.

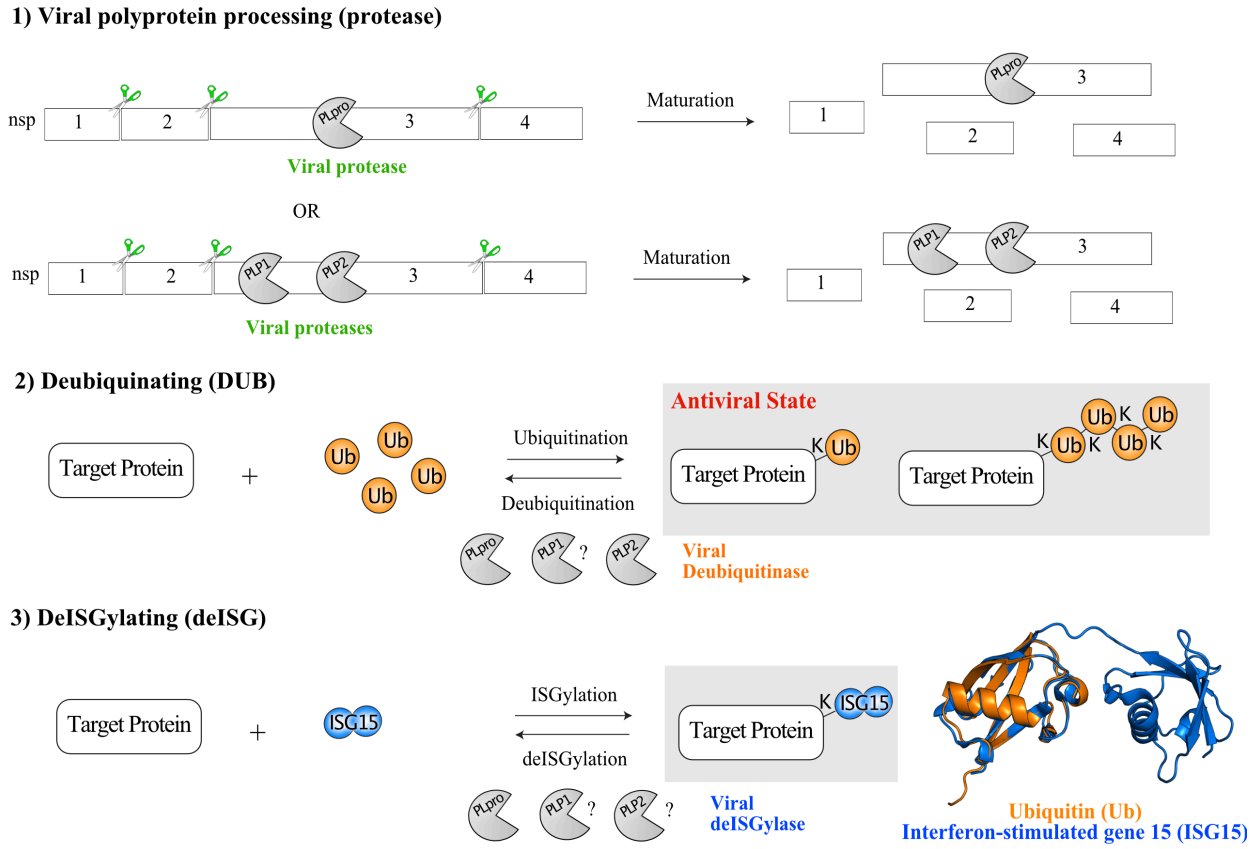


Figure 1.8 Multifunctional activities of CoV PLPs

(1) Viral polyprotein processing involves the cleavage by PLpro or PLP1/PLP2 at three sites leading to the release of nsps 1-4 for formation of the replicase complex. (2) PLPs deubiquitinating (DUB) activity reverses the process of ubiquitination by cleaving the isopeptide bond between mono- and extended polymeric Ub chains of target proteins. PLpro and PLP2s have been shown to have DUB activity as well as PLP1 from TGEV, but whether PLP1 acts as a DUB in other CoVs still remains in question. (3) DeISGylating (deISG) activity of PLP reverses the process of ISGylation by cleaving at the isopeptide bond of ISG15 conjugated proteins. PLpro from SARS-CoV and MERS-CoV and NL63 PLP2 have been shown to have robust deISG activity while PLP1 and other PLP2s deISG activity are weak, absent, or overall not well characterized. DUB/deISG activities of PLPs counteract the antiviral state of the host. Comparison of the Ub (orange, PDB, 1UBQ) and human ISG15 (blue, PDB code 1Z2M) crystal structures.

While PLpro and most PLP2s have shown to have DUB and deISG activities, PLP1 has yet to be tested for deISG activity, and most PLP1s have not been tested for DUB activity. Based on bioinformatics of the substrate-binding loop, HCoV-229E PLP1 may have DUB activity while HCoV-HKU1, MHV, HCoV-OC43, and BCoV PLP1 are most likely not DUBs [61]. HCoV-NL63 PLP1 did not show DUB activity compared to PLP2 in cells [79], and as mentioned previously, alpha-CoV TGEV PLP1 does have DUB activity. Since many cleavage sites of PLP1 share similar residues of the RLRGG sequence at the C-terminus of Ub and ISG15, it would not be surprising if PLP1 have evolved DUB and deISG activity like the PLP2 ortholog in certain CoVs.

1.3.2 Catalytic mechanism

PLPs are cysteine proteases that contain the canonical catalytic triad with the Cys-His-Asp arranged in the active site (Figure 1.7.B). The active site is adjacent to the BL2 loop of the protease and appears to be in a productive conformation for catalysis in crystal structures. The proposed chemical mechanism for SARS-CoV PLpro catalysis is shown in Figure 1.9. The catalytic triad is hydrogen bonded and properly aligned for catalysis. In the free enzyme "E" form, the general base His273 abstracts a proton from the Cys112 thiol group. It is important to note that there is no data confirming the protonation state of the Cys residue, and this mechanism uses the classic thiolate form of Cys112 [68]. The peptide substrate comes into the active site forming the "ES" state (Figure 1.9.a) and the deprotonated Cys112 thiolate nucleophile attacks the carbonyl carbon of the peptide substrate to form the first negatively charged tetrahedral-transition state (T1-1 or FP; Figure 1.9.b). The oxyanion is stabilized by the Trp107 amino group of the oxyanion hole in SARS-CoV PLpro [63, 80].

From mutagenesis experiments, the Trp107 residue is critical for SARS-CoV PLpro catalysis, as Trp107Ala mutant destroyed the ability of SARS-CoV PLpro to cleave in a trans-cleavage assay [63]. A Trp107Leu/Ala108Ser double mutant of SARS-CoV PLpro also

completely lost its catalytic activity against fluorogenic substrates [81], which was most likely due to destroying the oxyanion hole [49]. However, Trp107 is not conserved in CoV PLPs. Rather an Asn or Gln residue is commonly found in this region of cysteine proteases. For example, USP7 and USP14 contain an Asn, and TGEV PLP1 contains Gln27 at the equivalent Trp107 position in SARS-CoV PLpro [51]. Similarly, MHV PLP2 has a Gln [54]. Interestingly, MERS-CoV PLpro contains a Leu106 in the place of Trp107. When the Leu was mutated to the equivalent Trp in SARS-CoV, MERS-CoV PLpro exhibited 40-fold enhancement in activity towards the RLRGG-AMC substrate [82]. Results support that MERS-CoV PLpro is less efficient at stabilizing the oxyanion hole compared to SARS-CoV PLpro. It has been proposed that the oxyanion hole of MERS-CoV PLpro is likely stabilized by backbone amides of Asn residues and the catalytic Cys (Asn1590, Asn1591, Cys1592 in PDB code 4RF1 from reference [83]). These Asn residues are conserved in CoVs PLPs and may therefore contribute to stabilizing the oxyanion hole in the case where the equivalent Trp107 position is not enough for stabilization or not involved.

Once the TI-1 is formed, it readily collapses. In this step, the scissile bond is cleaved, and the C-terminal amine of the peptide is released, which forms the thioester intermediate “F” (Figure 1.9.c). This acyl-enzyme intermediate is broken by the addition of the nucleophilic water molecule, which forms the second tetrahedral transition state (TI-2) or “FQ”. The oxyanion is again stabilized by the oxyanion hole residue of PLPs (Figure 1.9.d). Next, the Cys112 abstracts the proton from the general base to collapse the TI-1 state, which cleaves the next scissile bond to release the N-terminal carboxylic acid or the product “Q” (Figure 1.9.e). The product may be transiently stabilized in the active site before it is released (Figure 1.9.f), and the “E” state is regenerated for another round of catalysis.

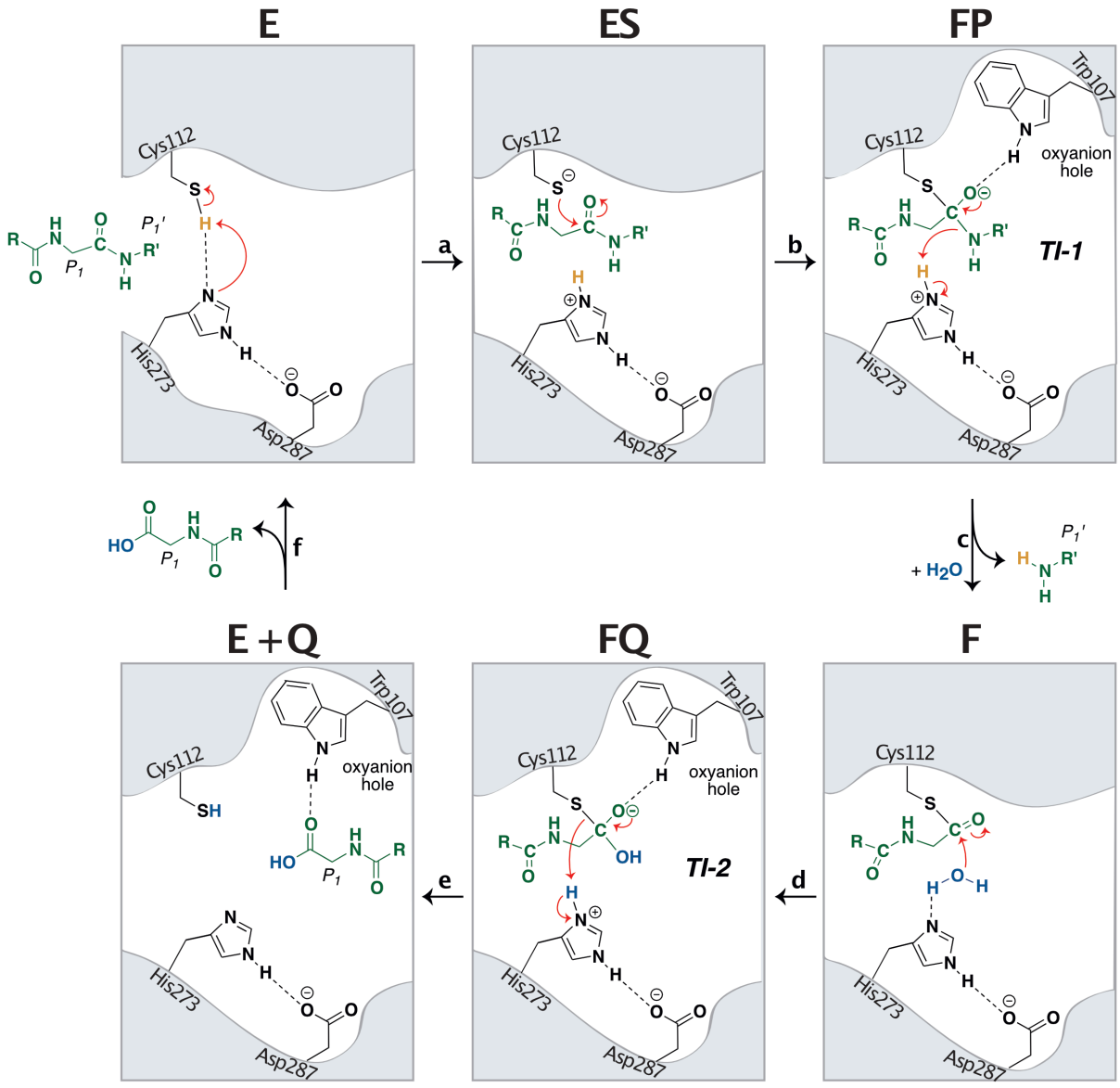


Figure 1.9 Proposed mechanism of SARS-CoV catalysis with a peptide substrate

Residues of the catalytic triad of SARS-CoV are shown as well as Trp107, which coordinates the oxyanion hole in the two tetrahedral transition (TI-1 and TI-2) states. The peptide and the catalytic water are shown in green and blue, respectively. Red arrows represent electron pushing in the chemical reaction while dotted-lines illustrate hydrogen bonds. Each intermediate state is labeled accordingly. Figure was adapted from reference [68].

1.3.3 PLP are IFN antagonists

Pathogen-associated molecular patterns (PAMPs) are recognized by pattern recognition receptors (PRRs) to activate innate immune response pathways. PRRs, such as toll-like receptors or retinoid acid inducible gene (RIG)-I-like helicase family, which includes cytosolic RIG-I and melanoma differentiation-associated protein (MDA5), are able to detect viral RNA and other replicative intermediates to activate the host antiviral response. This includes the recruitment of kinases and adapter proteins to translocate transcription factors, such as IFN regulatory factor-3 (IRF3) and nuclear factor κB (NFκB) into the nucleus (Figure 1.10). Once in the nucleus, IRF3 and NFκB synthesize type I interferons, such as IFN-β, as well as pro-inflammatory cytokines, which transduce downstream pathways, such as the JAK/STAT pathway. STAT transcription factors induce the expression of IFN stimulated genes (ISGs), which are involved in establishing the antiviral state of the host for viral clearance [66, 84].

The DUB/deISG function of PLPs, with an emphasis on DUB activity, has been implicated in antagonizing host innate immune responses to promote viral replication [77, 83, 85]. However, the precise mechanism of how PLPs use DUB and deISG activity to modulate host innate immunity is still unclear. A model of how SARS-CoV PLpro may act on different branches to antagonize the innate immune response is shown in Figure 1.10. In general, PLPs from arteriviruses are known to block IFN responses [42], and DUB activity is a critical factor in this process [85]. So far, beta-CoVs SARS-CoV [76] and MERS-CoV PLpro [84], MHV PLP2 [78, 86], and alpha-CoV TGEV PLP1 [52], and HCoV-NL63 [79] and PEDV PLP2 [70] have all demonstrated IFN antagonism activity. Both MERS-CoV and SARS-CoV PLpro reduce the expression of proinflammatory cytokines and chemokines, including CCL5, IFN-β, CXC10 in stimulated cells [87].

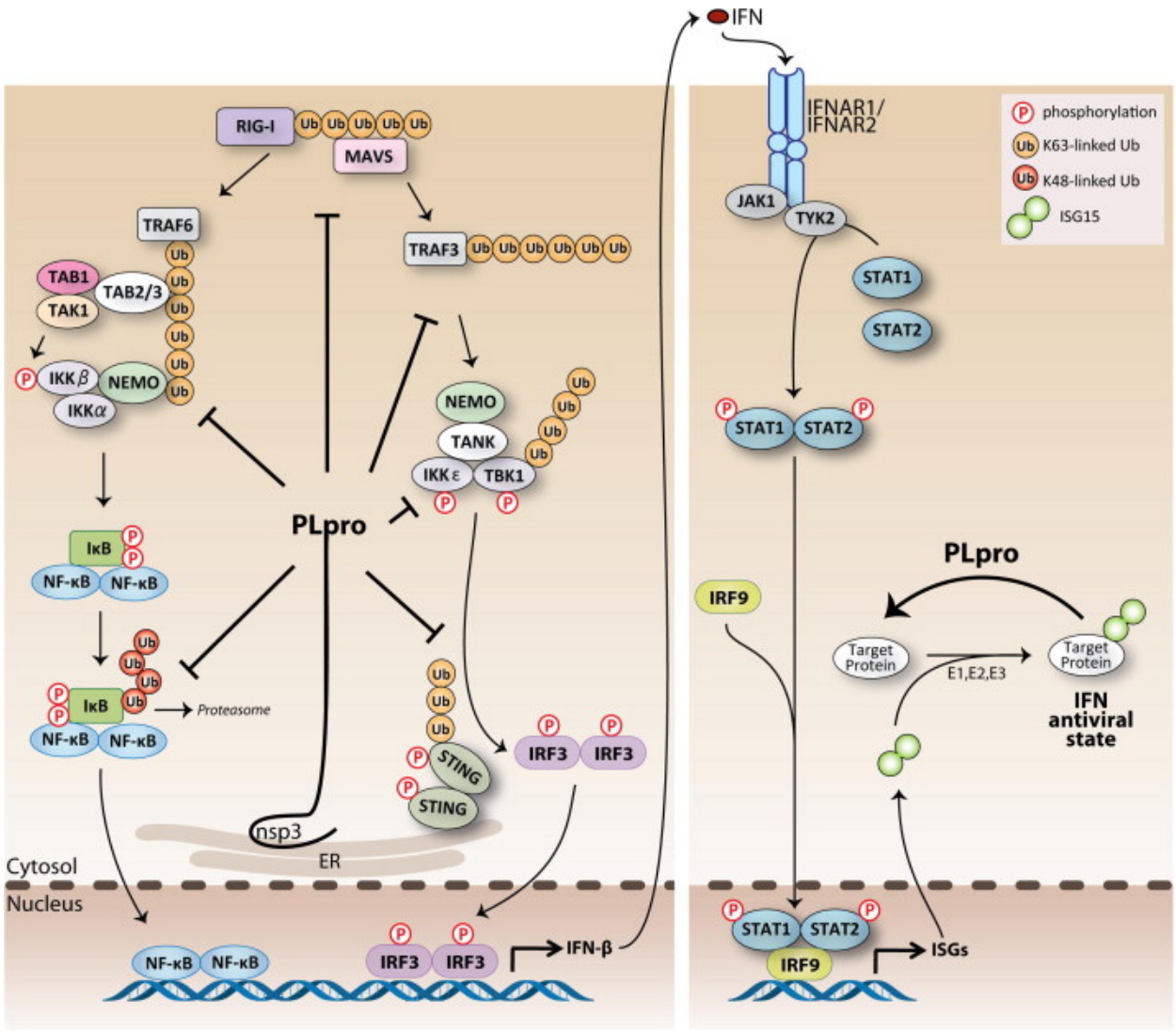


Figure 1.10 Proposed mechanism for PLP-mediated antagonism of IFN response

Upon viral infection, pattern recognitions receptors RIG-I and MDA5 sense viral RNA, which recruits mitochondrial antiviral signaling protein (MAVS) to the mitochondria to stimulate kinase complexes, such as IκB kinase and TANK-binding kinase 1 (TBK1). These kinases target transcription factors, IRF3 and NFκB into the nucleus. Stimulator of IFN genes (STING) resides in the ER. IκB must be degraded by the proteasome for NFκB activation, and the phosphorylated IRF3 dimer is required for translocation into the nucleus. IRF3 and NFκB then activate type I interferons to target the STAT transcription factor into the nucleus for expression of ISGs. ISG15, the most abundant ISG, is conjugated onto target proteins via the E1, E2, E3 enzymes and establishes the antiviral state. SARS-CoV PLpro of nsp3, membrane-bound in the ER, may act on different branches of the host innate immune response by directly interacting with proteins and/or deubiquitinating, or deISGylating proteins upstream and downstream in this pathway. PLpro activities counteract the host antiviral state for SARS-CoV evasion of the host immune system. Figure is reproduced from [68].

Devaraj et al. showed that SARS-CoV PLpro is a potent IFN antagonist that directly interacts with IRF3 to inhibit its phosphorylation and dimerization, and ultimately its translocation into the nucleus [75]. Another study supported that SARS-CoV PLpro acts in the IRF3 pathway, specifically through its interaction with the STING-TRAF3-TBK1 complex [88]. SARS membrane-anchored PLpro reduced the levels of ubiquitination of RIG-I, STING, TRAF3, TBK1, and IRF3 to suppress the IFN response. MHV PLP2 also was shown to deubiquitinate TBK1 and reduce IFN-β response through the IRF3 pathway [78]. Furthermore, NL63-CoV PLP2 also inhibits the STING-mediated IFN pathway [84], and PEDV PLP2 was shown to deubiquitinate RIG-I and STING to inhibit IFN-β [70]. Bailey-Elkin et al. and colleagues reported that MERS-CoV PLpro inhibits IFN-β induced by either RIG-I, MAVS, or IRF3 [83]. They also demonstrated that a single V1691R mutation selectively targeting the DUB activity of MERS-CoV PLpro removed the ability of PLpro to act as an IFN antagonist, supporting that DUB activity is involved in the IFN antagonism activity of PLpro. In a recent report, TGEV PLP1 also inhibited RIG-I-, MAVS-, STING- and TBK-1-mediated IFN-β expression [52]. This was the first example of a CoV PLP1 acting as an IFN antagonist.

In addition to the IRF3 pathway, Frieman et al. reported that SARS-CoV PLpro can also block NFκB pathway by stabilizing IκBα, a protein that is normally degraded to activate NFκB and its proinflammatory cytokines [76]. Another study by Clementz et al. similarly showed that SARS-CoV PLpro could inhibit TNFα-induced NFκB activation, but also demonstrated that a selective PLpro inhibitor could be used to remove the antagonizing effect mediated by PLpro [79]. Deng et al. also described a chimeric Sindbis virus-mouse model system for evaluating small molecule PLpro inhibitors using the deISG activity of PLpro as a reporter in mice. IFN receptor knockout mice infected with chimeric viruses expressing both the ISG15 machinery and PLpro

blocked the ability of ISG15 to protect the mice while a selective inhibitor targeting the deISG activity of PLpro allowed for ISG15-mediated protection of mice to clear Sindbis virus [89].

Although there has been extensive work on the IFN antagonism activities of PLPs, there have been controversial results in the literature. For example, one report noted the MHV PLP2 does not act as an IFN antagonist [76]. In addition, binding partners in the innate immune response have varied between reports. SARS-CoV PLpro was shown to interact with IRF3 only when the TM region was attached [76, 88].

Additionally, there has been discrepancies on whether the enzymatic activity of PLPs are necessary for IFN antagonism. For example, reports have described that catalytic mutants remove PLP-mediated IFN antagonism for MERS-CoV PLpro [87], PEDV PLP2 [70], and for SARS-CoV PLpro (only in the case of NF-κB activity [79]) while other reports have demonstrated that catalytic mutants have varying degrees of IFN antagonism that is reduced but not completely abolished (observed in SARS-CoV PLpro [76] and MHV PLP2 [86]). In addition, different constructs have produced varying results. For example, Yang et al. demonstrated that membrane anchored MERS-CoV PLpro-TM did not require its catalytic function for RIG-I induced IFN-β promoter activity [90] while the report by Bailey-Elkin et al. demonstrated that the Ubl2-PLpro construct required catalytic function and DUB activity for IFN antagonism [83]. Similarly, Devaraj et al. showed that SARS-CoV PLpro-TM inhibits IFN expression independent of catalytic activity [75]. In the case of HCoV-NL63, however, at higher concentrations both the PLP2 and PLP2-TM form catalytic mutants were able to reduce IFN response [79]. Overall, these results supported the possibility that additional IFN antagonists may be present in nsp3 and catalytic mutants of PLP may not be sufficient to remove IFN antagonism activity. Also, it is possible that catalytically inactive PLP could still facilitate IFN antagonism by protein-protein interactions.

Recently, Niemeyer et al. and others showed that SARS-CoV related strains derived from bat *Rhinolophus* species contain catalytically dependent PLPs with IFN antagonism function while infectious SARS-CoV strains have protease-independent anti-IFN function [91]. The protease-independent anti-IFN activity exhibited by SARS-CoV strains is related to PLPs' ability to bind Ub, as a M209R mutant targeting the Ub binding site of SARS PLpro showed less anti-IFN activity compared to the wild-type. This study supported that PLP from epidemic SARS-CoV have enhanced IFN antagonism activity compared to bat-derived strains.

There have also been conflicting reports whether the Ubl2 domain adjacent to PLPs are important in PLP-mediated IFN antagonism. In the case of SARS-CoV PLpro, two reports have shown conflicting results. Frieman et al. and colleagues show that the Ubl2 domain is required [76] while Clementz et al. reports the opposite that PLP-mediate IFN antagonism is independent of the Ubl2 domain [79].

1.4 PLP for antiviral or live-attenuated vaccine

Since PLPs function in both viral replication and suppression of innate immune responses, the multifunctional enzyme is an attractive candidate for both antiviral drug discovery and protein engineering studies for live-attenuated vaccines. Extensive work has gone into designing potent and selective PLpro inhibitors for anti-SARS drugs [68, 92, 93]. These SARS-CoV PLpro inhibitors also target HCoV-NL63 PLP2, which has a conserved aromatic residue in the substrate-binding loop where the inhibitor binds [94]. However, small molecules that target other CoV PLPs, such as MERS-CoV, FIPV and PEDV, have yet to be developed, and therefore high-throughput inhibitor screens are necessary to identify unique scaffolds for these proteases.

Based on its ability to modulate innate immunity, PLPs could be engineered to create a live-attenuated CoV vaccine. By selectively targeting the ubiquitin binding site of PLP through structure-guided mutagenesis, the ability of PLP to suppress innate immune responses

(DUB/deISG activities) can be removed, at the same time maintaining PLPs viral polyprotein processing ability (protease) for viral replication. The net effect of designing PLP mutants that are selectively deficient in DUB/deISG activities in the virus may activate the host innate immunity during viral infection, and hence attenuate CoV pathogenesis.

1.5 Overview of live attenuated vaccines

Historically, live attenuated vaccines, among other strategies i.e. inactive or subunit vaccines, have proven to be successful at providing intervention against viral pathogens causing infectious disease. For example, live attenuated vaccines against smallpox, poliovirus, yellow fever, measles, mumps, and rotavirus have controlled the spread of disease in the human population [95]. These live-attenuated vaccines trigger robust immune responses and provide long-lasting protection for establishing herd immunity. Although live attenuated vaccines are the most cost effective and successful vaccines, on-going studies continue to monitor their efficacy and safety in practice, as the virus can possibly revert to the virulent form. In the case of the polio, the type 3 vaccine strain showed reversion, recombination, and genetic drift in under 50 days post vaccination in two children case studies [95]. The more virulent type 1 strain in animal models has shown the least reported vaccine-associated diseases, which suggests that many weak attenuating mutations may put less selective pressure on the virus to mutate. Ultimately, after identifying an attenuated virus strain, it is important to understand what is causing the phenotype, the reversion rate in virus passages, and the possibility of recombination events to prevent the vaccine from causing lethal disease.

Live-attenuated vaccines have originated from passage rather than rational design. The yellow fever vaccine is a classic example. An attenuating strain was developed after the virus was serially passed in chick embryos in the late 1930s [95]. Similarly, one of the most effective live attenuated viruses against a CoV was developed in the 1960s in the Netherlands. After 120 serial

passages of the IBV strain in embryonated eggs, virus H120 produced strongly attenuated IBV disease in young chickens and protected against challenge with wild-type IBV [96]. Similarly, genomes mutated using multiple passages in Vero cells led to the development of swine enteric CoV vaccines [97].

In the case of TGEV and PEDV, there have been many different types of vaccines developed other the last decade, including live attenuated vaccines, inactivated bivalent TGEV/PEDV vaccine, and even a trivalent TGEV/PEDV/rotavirus vaccine [98]. Since there has been a demand for more efficacious vaccines in China, many vaccines were developed without documentation of efficacy and safety. In the U.S., there are two conditionally licensed PEDV vaccines developed by Harrisvaccines™ and Zoetis. The first was designed using an alphavirus replicon system and incorporated an insertion in the spike PEDV protein while the second consisted of an inactivated vaccine [97]. Although there has been extensive work on PEDV vaccines, published data on its efficacy in protecting suckling piglets is lacking [98].

While attenuated virus strains arising from passage have led to the design of effective live-attenuated vaccines, the process is costly, unpredictable, and time-consuming. In the 1990s, reverse genetics plasmid-based technology was developed, which revolutionized the study of viral infection as well as vaccine development. The reverse-genetics system has served as a way for researchers to evaluate vaccine candidates in appropriate animal models to rationally develop live-attenuated or inactivated vaccines. Researchers have utilized this technology for influenza vaccines to rapidly compensate for antigenic drift of surface glycoproteins HA and NA of influenza strains.

The reverse genetics system has allowed for improvement of previous attenuated vaccine strains developed through passage. For example, a recent study showed that deletions of IBV accessory genes 3ab and/or 5ab produced an attenuated phenotype and immune protection in

chickens for a next generation IBV vaccine, which may be effective against multiple IBV serotypes [99]. Furthermore, reverse genetic experiments can help in identifying attenuating mutations for an overall more stable and safe vaccine. Jimenez-Guardeño et al. and colleagues used reverse genetics to evaluate attenuating mutations of SARS-CoV in mice [100]. They showed that recombinant viruses with attenuating mutations in the E protein and the nsp1 gene of SARS-CoV were able to maintain fitness in the attenuating phenotype and protect mice challenged with wild-type virus.

Our collaborators in Susan Baker's lab used reverse genetics system in mice to study the role of PLPs in CoV replication and pathogenesis. One particular attenuating mutation (Val787S) was identified in the Ubl2 domain of MHV. This Val to Ser mutation reduced the thermal stability and DUB activity of the adjacent PLP2 domain at increased temperatures. When this mutation was reinserted into the virus, mice infected with mutant viruses not only survived but were protected when challenged with wild-type virus [101].

1.1 Justification for the study of PLP mutants

Recent mutagenesis studies by previous lab member Dr. Yafang Chen demonstrated that PLP is a possible target for attenuating CoV infection. Based on her MHV PLP2-Ub bound structure, she generated a D250A mutant that exhibited 56% reduced Ub-AMC compared to wild-type activity with retained RLRGG-AMC activity. The mutant also exhibited slightly enhanced ISG15-AMC activity [102]. Preliminary results from Susan Baker's lab have revealed that the D250A mutant virus is slightly attenuated in mice and enhances immune responses (personal communication). These results support that reducing DUB activity of PLP is a potential strategy for attenuating CoV infection in an animal model.

The mutation R281A was the first example of a PLP mutant with selectively reduced deISG activity (4% remaining activity) while protease activity (88% remaining activity) and DUB activity

(69% remaining activity) were less affected. When this mutant was reinserted back into the MHV genome by reverse genetics, the mutant virus was strongly attenuated in mice, and the mice were protected against challenge with wild-type virus. After the mutant virus was deep sequenced, however, another mutation in the EndoU domain of nsp15 was discovered, which was determined to be responsible for virus attenuation. In addition, a virus with the single Thr98Met mutant in nsp15 has never been recovered to date [46], suggesting that the PLP mutant in nsp3 stabilizes the T98M mutation.

1.5.1 Structure-guided design of PLP mutants

The primary objective of this dissertation is to selectively remove the ability of PLP to suppress innate immune responses (DUB/deISG activities) while maintaining PLPs ability to function in viral replication (protease activity) for the design of a live-attenuated CoV. One major challenge in designing a selectively DUB/deISG deficient PLP is that PLPs rely on the same functional active site. Therefore, targeting the active site will remove all three activities, yielding a nonviable virus. To overcome this challenge, we will use a structure-guided approach to target the ubiquitin-binding site of the protease located outside of the active site.

Another goal of this research is to tease out the DUB and deISG activities of PLPs to delineate how each activity functions in antagonizing the innate immune response. In the literature, there has been an emphasis on understanding DUB activity and a lack of structural and function studies that fully characterize the recognition of PLPs towards ISG15. In addition, designing a DUB or deISG deficient PLP is not trivial because mono-Ub and the C-terminal domain of ISG15 share the same binding pocket. Therefore, a majority of the PLP mutations that target this binding pocket will alter both of these activities in some way.

Figure 1.11 shows the activities of the wild-type PLPs in comparison to PLP mutants we aim to generate from our mutagenesis studies. There are three different categories of PLP mutants:

DUB/deISG deficient (Figure 1.11.B), selectively deISG deficient (Figure 1.11.C), and selectively DUB deficient (Figure 1.11.D). Again, protease function is retained for all three mutants. The first set of PLP mutants remove DUB and deISG activities. The net effect would likely remove host cell antagonism function. The second and third set of mutants would test for the DUB and deISG mechanism in host cell antagonism, respectively. The second set of PLP mutants yield only a DUB enzyme while its deISG function is completely removed. One would expect that PLP with intact DUB activity would still function in IFN antagonism, but a mutant that is selectively deISG deficient would look further the DUB mechanism of PLPs. Lastly, the third set of PLP mutants would yield only a deISG enzyme with removed DUB function.

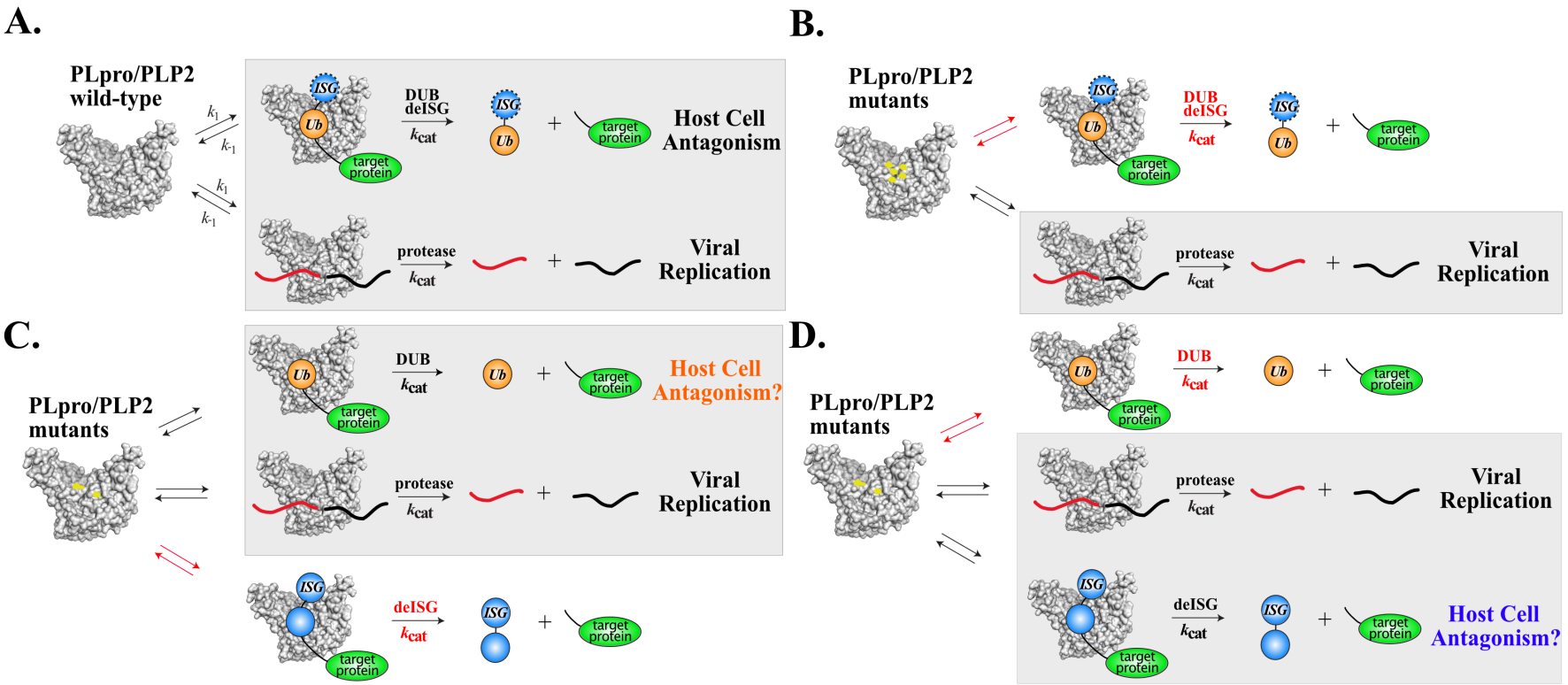


Figure 1.11 Schematic of desired PLpro mutants with altered activities

(A) PLP wild-type binds and hydrolyzes both Ub/ISG15 substrates (top reaction, DUB/deISG) for host cell antagonism and viral polyprotein substrates (bottom reaction, protease) for viral replication. Wild-type activities are shown with a gray box. (B-D) All desired PLP mutants retain the ability to hydrolyze the viral polyprotein substrates for viral replication. First set of mutants (B) are unable to bind and hydrolyze both Ub/ISG15 substrates (top reaction, DUB/deISG deficient) removing their ability to antagonize host cell innate immunity. The second and third set of mutants either retain their ability to hydrolyze Ub (C, top reaction) or ISG15 (D, bottom reaction) for IFN antagonism but are unable to bind the other, either ISG15 (C, deISG deficient) or Ub (D, DUB deficient). It is questionable whether DUB or deISG activity of PLP alone are sufficient for evasion of the innate immune system.

1.6 Central hypothesis and statement of purpose

The central hypothesis is that selectively removing PLP DUB/deISG activities will activate the innate immune response and reduce viral pathogenesis. The goal of this research project is to design PLPs with separated DUB and/or deISG activity for investigating DUB and deISG mechanism of PLPs in coronavirus replication and pathogenesis. The primary objective is to design a live-attenuated coronavirus by engineering PLPs to no longer suppress the innate immune system, i.e. by eliminating its DUB and/or deISG activities, while maintaining its viral polyprotein processing ability for viral replication.

In this dissertation, we investigated the structure and function of MERS-CoV PLpro (Chapter 2 and Chapter 4), SARS-CoV PLpro (Chapter 3), and translated the knowledge gained from beta-CoV PLPs to alpha-CoV FIPV and PEDV PLP2 (Chapter 5) for structure-guided mutagenesis. Results from kinetic, biochemical, and crystallographic studies revealed a common hot spot for altering Ub specificity of PLPs across the alpha- and beta-CoV genera. The PLP mutants described here will help researchers characterize the IFN antagonism activity of PLP in cellular studies. In addition, this work will help geneticists identify mutations for live-attenuated CoV vaccines.

CHAPTER 2. X-RAY STRUCTURE AND ENZYMATIC ACTIVITY PROFILE OF A CORE PAPAIN-LIKE PROTEASE OF MERS CORONAVIRUS WITH UTILITY FOR STRUCTURE-BASED DRUG DESIGN

A version of this chapter has been published in the following journal article [103]*.Cellular data in this chapter was done by our collaborators in Dr. Susan Baker's lab at Loyola University.*

Ubiquitin-like domain 2 (Ubl2) is immediately adjacent to the N-terminus of the papain-like protease (PLpro) domain in coronavirus polyproteins, and it may play a critical role in protease regulation and stability as well as in viral infection. However, our recent cellular studies reveal that removing the Ubl2 domain from MERS PLpro has no effect on its ability to process the viral polyprotein or act as an interferon antagonist, which involves deubiquitinating and deISGylating cellular proteins. Here, we test the hypothesis that the Ubl2 domain is not required for the catalytic function of MERS PLpro *in vitro*. The X-ray structure of MERS PLpro-ΔUbl2 was determined to 1.9 Å and compared to PLpro containing the N-terminal Ubl2 domain. While the structures were nearly identical, the PLpro-ΔUbl2 enzyme revealed the intact structure of the substrate-binding loop. Moreover, PLpro-ΔUbl2 catalysis against different substrates and a purported inhibitor revealed no differences in catalytic efficiency, substrate specificity, and inhibition. Further, no changes in thermal stability were observed between enzymes. We conclude that the catalytic core of MERS PLpro, i.e. without the Ubl2 domain, is sufficient for catalysis and stability *in vitro* with utility to evaluate potential inhibitors as a platform for structure-based drug design.

2.1 Introduction

Coronaviruses (CoVs) are enveloped, positive sense, single-stranded RNA viruses that cause mild to severe upper respiratory tract infections in humans. Approximately 10 years after

emergence of the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002/2003, Middle East respiratory syndrome coronavirus (MERS-CoV) emerged and has been identified so far in 26 countries with a case-fatality rate over 30% [10, 104]. Although these CoVs are well-recognized global pathogens, there are no antiviral interventions available. Thus, a better understanding of the molecular mechanisms that facilitate viral pathogenesis and replication may permit the design of targeted therapeutics against CoVs.

MERS-CoV is classified in the sub-lineage C genus *Betacoronavirus* with a conserved genomic size of ~30 kb among other CoVs [10]. The first 22 kilobases located at the 5'-end of the RNA genome is encoded in two open reading frames (ORF1a/ORF1b) that are translated by host ribosomes to generate two respective viral polyproteins (pp1a & pp1ab). Pp1a and pp1ab are processed by two virus-encoded cysteine proteases, termed the papain-like protease (PLpro) and the 3C-like protease (3CLpro). Together, these two proteases cleave the polyproteins to produce 16 nonstructural proteins (nsps), which are essential for the formation of the replicase complex and hence RNA replication. This study focuses on the multifunctional and putative drug target, PLpro, located in nonstructural protein 3 (nsp3; Figure 2.1.A).

In addition to its function of cleaving the viral polyprotein into the requisite nsps, SARS-CoV PLpro is also a viral ubiquitin-specific protease (vUSP), having a structural fold almost identical to the human USP family [61-63]. SARS PLpro is a highly efficient deubiquitinating (DUB) enzyme having the ability to rapidly hydrolyze isopeptide bonds of proteins that are post-translationally modified by cellular ubiquitin-like (Ubl) molecules, such as ubiquitin (Ub) and interferon-stimulating gene 15 (ISG15), which are two key regulators of the innate immune response [75-77]. We and others have also shown that SARS-CoV PLpro and other CoV PLpros display substantially different substrate specificities for ISG15 and certain poly-Ub chains [54, 94, 105]. More importantly, the DUB/deISGylating activities have been shown to play an important

role in antagonizing host innate immune responses to promote viral replication [77, 83, 85] although the precise roles for each activity in this antagonism have yet to be determined.

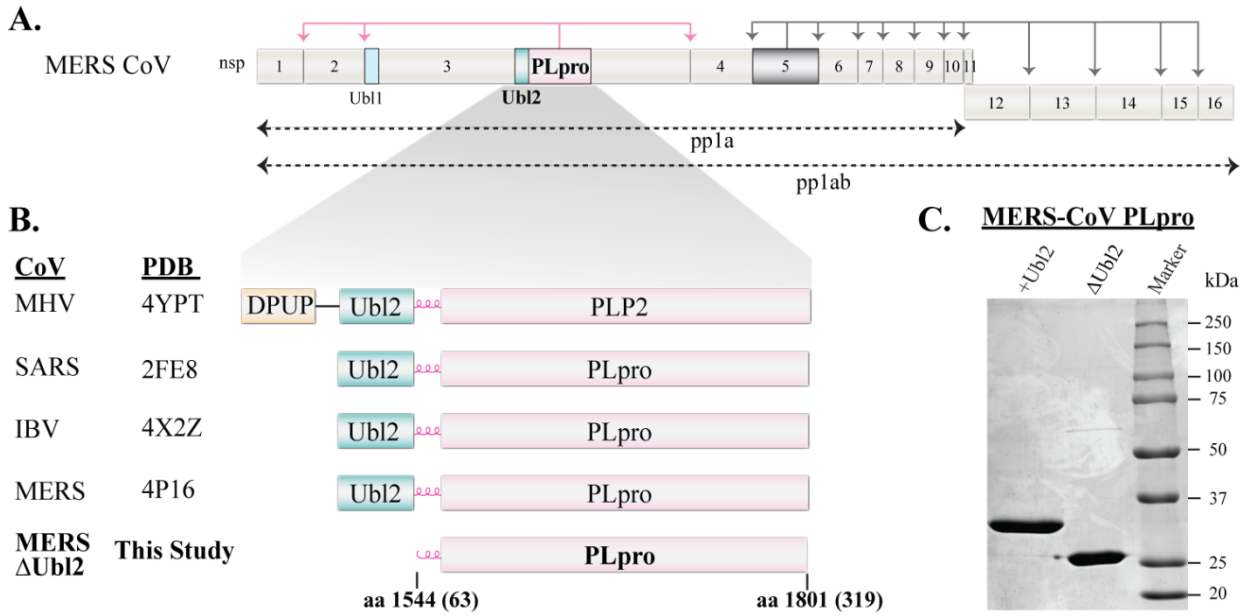


Figure 2.1 MERS-CoV polyprotein organization and design rationale for the MERS-CoV PLpro-ΔUbl2 construct of nsp3

(A) Non-structural proteins (nsps) are numbered 1-16 within the MERS-CoV viral polyprotein 1a and 1ab. MERS PLpro is colored in pink in nsp3 and 3CLpro, which is in nsp5, is colored in gray and their respective cleavage sites are colored accordingly and indicated by arrows. The Ubl1 (light blue) and Ubl2 (green) domains of nsp3 are indicated. (B) Summary of the current PLP X-ray structures and the smallest catalytic unit determined in this study. The PDB codes of the X-ray structures that were determined first for MHV PLP2, SARS-CoV, IBV and MERS-CoV PLpro, all containing the Ubl2 domain, are indicated. (C) SDS-PAGE (12.5%) analysis of purified MERS-CoV PLpro-Ubl2 (36 kDa) and –ΔUbl2 (29 kDa) used for activity assays. The proteins are estimated to be >95% purity by densitometry.

Interestingly, the CoV RNA genome encodes for two Ubl domains within nsp3 that are denoted as Ubl1 [106] and Ubl2 [63] according to their location in the nsp3 multi-domain protein. The Ubl2 domain of SARS-CoV, previously named the Ubl, was first identified by our lab through X-ray structural studies where it was found to reside directly adjacent to the N-terminus of the PLpro catalytic domain [63]. Since our original structure, the Ubl2 domain has been found to be conserved among CoVs to date, including MERS-CoV [82], murine hepatitis virus (MHV) [54], and infectious bronchitis virus (IBV) [69]. However, the functional roles for Ubl2 in viral

pathogenesis and RNA replication remain enigmatic. So far, the majority of studies aimed at understanding the roles of viral Ubls in CoV replication have focused on the Ubl2 domain due to its location in the RNA genome and potential to modulate the enzymatic activity of PLpro. For example, we investigated the function of Ubl2 in SARS-CoV [76] and MHV [101] and found that the Ubl2 fold is crucial for maintaining PLpro structural integrity *in vitro*. Interestingly, in cell-based assays, SARS-CoV PLpro without its Ubl2 domain was no longer able to antagonize the pathways involved in the host innate immune response and thereby act as an IFN antagonist [76]. However, the mechanism that leads to this loss of function is not well understood because the enzyme retained its protease and DUB activities. In the case of MHV, a single point mutation in the Ubl2 domain was found to reduce the thermal stability of the papain-like protease 2 (PLP2) domain rendering the enzyme DUB deficient in cells [101]. When this mutation was inserted back into the virus, the mutant virus was attenuated in infected mice with the ability to replicate and induce a protective immune response against wild-type virus. This study revealed that the Ubl2 domain could be used as a strategy to attenuate CoV pathogenesis leading us to further investigate the function of the Ubl2 domain in MERS-CoV.

Unexpectedly, our recent cell-based studies using different truncated forms of MERS-CoV Ubl2 (MERS PLpro-ΔUbl2) suggest that the Ubl2 domain might not be as pivotal for PLpro enzyme activity as originally thought. In these cell-based assays, MERS-CoV PLpro-ΔUbl2 appeared to retain its multiple enzymatic functions and unlike SARS-CoV PLpro-ΔUbl2, preserved its ability to act as an IFN antagonist [94]. This intriguing discovery engendered the hypothesis for this study that the MERS-CoV Ubl2 domain is not required for PLpro catalytic function and stability *in vitro*. Here, we report a series of X-ray structural and kinetic studies on the MERS-CoV PLpro-ΔUbl2 construct to elucidate the importance of Ubl2 domain in MERS-

CoV stability, substrate specificity and enzymatic catalysis as well as to evaluate the efficacy of a reported MERS-CoV PLpro inhibitor [107].

2.2 Materials and Methods

2.2.1 Expression and Purification of MERS-CoV PLpro-Ubl2 and PLpro-ΔUbl2

The genes encoding MERS-CoV PLpro-Ubl2$_{1484\text{-}1802}$ and PLpro-ΔUbl2$_{1544\text{-}1802}$ were previously cloned into pEV-L8 expression vector [94] from a former member of the Mesecar lab, Dr. Yahira Báez. The amino acid sequences of PLpro are from MERS polyprotein 1ab from the strain human betacoronavirus 2c EMC/2012 (K4LC41). A stop codon was incorporated at the C-terminal end (at residue C319) of PLpro-ΔUbl2 by site-directed mutagenesis (Fwd, 5'-CTTCCGATTGATAAATTGGAAGTGGATAA-3'; Rev, 5'-CTTCCAATTTATCAATCGGAAGAGTATTT-3'). The resulting pEV-L8-PLpro-ΔUbl2 (Figure 2.1.B) plasmid DNA was subjected to DNA sequencing using the Purdue Genomics Core Facility to confirm that the correct construct was generated. Each construct pEV-L8-PLpro-Ubl2$_{1484\text{-}1802}$ or pEV-L8-PLpro-ΔUbl2$_{1544\text{-}1801}$ was transformed into *E. coli* BL21 (DE3) cells for protein expression, and a 10 mL LB starter culture (in a 50 mL falcon tube) containing 50 μg/mL kanamycin was inoculated with a colony and grown for 16 hours at 37 °C at 200 rpm using the ATR Biotech Multitron HT Infors Dual-Stack Incubator-Shaker until the OD600 reached ~4. One liter of super broth media [94] (in a 2.8 L Fernbach, unbaffled flask) supplemented with 0.2 % lactose, 0.6 % glycerol, and 0.05% glucose plus 50 μg/mL kanamycin was inoculated with the 10 mL starter culture and then grown for 24 hours at 25 °C at 140 rpm using the New Brunswick innova 44 until the OD600 reached ~4. Cells were harvested by centrifugation (6,750 x g, 4 °C, 20 minutes), and pellets from 2 L total cultures, weighing 16-18 g, were combined and frozen at -80 °C. On the day of purification, the frozen cell pellets were thawed and then resuspended in 74 ml lysis buffer (34-50 mL per 1 L culture). Lysis buffer contained buffer A (20 mM Tris, pH 7.5,

500 mM NaCl, 5 mM imidazole, 10 mM BME, and 5% glycerol) supplemented with 0.25 mg/ml lysozyme and 10 μg/ml DNase I. The resuspended cells were lysed on ice by sonication using a Branson Digital Sonifier (70% amplitude; 15 minutes, 10 s pulses, 10 s delays). The lysed cell debris was then removed by centrifugation (27,200 x g, 4 °C, 60 minutes).

The clarified lysate (Figure 2.2, lane 1) was then loaded onto a 5 ml His-Trap FF column (GE Healthcare) that was pre-charged with Ni^{2+} and equilibrated with 3% buffer B (20 mM Tris, pH 7.5, 500 mM NaCl, 400 mM imidazole, 10 mM BME, and 5% glycerol). After washing unbound proteins with 5 column volumes (25 mL) of 3% buffer B, PLpro was eluted with a 110 ml linear gradient (3% to 100%) of buffer B at a flow rate of 3 ml/min, and eluted fractions with active PLpro was pooled (Figure 2.2, lane 2). The His_8-tag was then removed by addition of tobacco etch virus (TEV) protease (1 mg TEV: 6 mg PLpro) after dialyzing against 1-2 L buffer A for 4 hours at 25 °C and then 4 °C overnight.

This sample was passed over the same His-Trap column and the unbound, untagged PLpro was collected in the flow-through. Untagged PLpro was then concentrated to 2-5 mg/ml using an Amicon Ultrafiltration Centrifugal device (10 kDa MW cutoff) while buffer exchanging into S75 buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 10 mM DTT, and 5% glycerol). The final purified protein was flash-frozen with liquid nitrogen in 100 μl aliquots and stored at -80 °C for activity assays and crystallization. The purity of this sample, referred to Ni His-trap 2 pool, is shown in Figure 2.2, lane 4 and another batch used for kinetic assays is shown in Figure 2.1.C.

2.2.2 Specific activity assays

Throughout purification, samples after each purification step were saved and tested for activity using 32 nM PLpro and either 250 nM Ub-AMC or Ub-Rho in assay buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1 mg/mL BSA, and 5 mM DTT. For lysate samples, to observe hydrolysis a higher amount of protein was used in the assay, approximately 2 μg is

sufficient. The assay was performed at a 50 µL volume in a full area, 96-well full area black microplate from Corning, and fluorescence was monitored using a BioTEK Synergy H1 multimode microplate reader at 25 °C. The initial rate of the reaction in the form of relative fluorescence units per min (rate; RFU/min) was converted to initial velocity (v; µM/min) or amount of hydrolyzed substrate per min by using the substrate extinction coefficient (Δε; RFU/µM) of product or the maximum amount of AMC that is released from the reaction. The specific activity (µM/min/mg) of each sample was determined using equation 2.1.

$$\text{Specific activity}\left(\frac{\mu\text{M}}{\text{min}\cdot\text{mg}}\right)=\frac{\text{Rate}}{\Delta\varepsilon\cdot\text{mg of enzyme in assay}}=\frac{\text{RFU}}{\text{min}}\cdot\frac{\mu\text{M}}{\text{RFU}}\cdot\frac{1}{\text{mg}} \quad (2.1)$$

The specific activity was used to calculate the fold purification shown in equation 2.2, where the lysate represented a fold purification of 1.0.

$$\text{Fold purification}=\frac{\text{Specific activity of sample}}{\text{Specific activity of lysate}} \quad (2.2)$$

In addition, the specific activity values were also used to calculate the amount of total enzyme units (µM/min) obtained at the purification step from the following equation 2.3.

$$\text{Total Units}\left(\frac{\mu\text{M}}{\text{min}}\right)=\text{Specific activity}\cdot\text{Total protein}=\frac{\mu\text{M}}{\text{min}\cdot\text{mg}}\cdot\text{mg} \quad (2.3)$$

The total enzyme units were then used to calculate the percent enzyme yield of recovery shown in equation 2.4, where the lysate represented a yield of 100%.

$$\text{Yield}\ (\%)=\frac{\text{Total Units of sample}}{\text{Total Units of lysate}}\times 100 \quad (2.4)$$

A summary of the total protein, total enzyme units, specific activity, fold purification, and percent enzyme yield from a purification batch after these steps is shown in Table 2.2 and Figure 2.2.

2.2.3 Crystallization and X-ray Structure Determination of MERS PLpro-ΔUbl2

For crystallization, the frozen, untagged PLpro was thawed on ice and loaded onto a HiLoad 26/60 Superdex 75 column (GE Healthcare) equilibrated with S75 buffer. Fractions containing active PLpro were then pooled and concentrated to 15 mg/ml for crystallization (Figure 2.2, lane 5). The purification summary from this preparation is shown in Table 2.2 using assay conditions 0.250 μM Ub-Rho, 32 nM PLpro in assay buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1 mg/mL BSA, and 5 mM DTT. A screen for initial crystallization conditions was performed using a Mosquito®Crystal liquid handling robot (TTP Labtech) in sitting drop mode and a series of sparse-matrix crystallization screens (Qiagen). Sitting drops were prepared by adding 100 nl of purified MERS-CoV PLpro-ΔUbl2 to 100 nl of reservoir solution. Three protein concentrations (5 mg/ml, 10 mg/ml, and 15 mg/ml) were setup in each of the three sub-wells in a 96-3 well sitting drop vapor diffusion plates (Greiner CrystalQuick crystallization plate). An initial crystallization hit from the cation suite containing 4.5 M Ammonium acetate and 0.1 M Tris-HCl, pH 8.5 was observed with Rigaku Minstrel® HT imaging robot after 1 week of incubation at 20 °C. Further optimization at 4 °C with drops containing 2 μl of purified MERS-CoV PLpro-ΔUbl2 at 10 mg/ml and 2 μl reservoir (5.5 M Ammonium Acetate, 0.1 M Tris-HCl, pH 8.5) yielded crystals with approximate dimensions of 0.05-0.1 mm after one week (Figure 2.2). Crystals were harvested using pins with nylon loops, transferred briefly to a cryo-protectant solution containing reservoir solution that was supplemented with 20% glycerol and then immediately flash-cooled by plunging into liquid nitrogen. Crystals were placed into SPINE pucks for transport to the Advanced Photon Source Synchrotron (APS), Argonne National Laboratory (ANL). X-ray data were collected on crystals using beamline 21-ID-F at the Life Sciences-Collaborative Access Team (LS-CAT).

X-ray data were collected on a single MERS-CoV PLpro-ΔUbl2 crystal using 1° rotations at 100 °K. X-ray data were indexed, processed, and scaled using HKL2000 [108]. To determine the initial phases for the structure, molecular replacement with Phaser was performed using the structure of MERS-CoV PLpro-Ubl2 apo (PDB code 4P16) as a search model [82]. Model building and refinement on the resulting structural solution containing one molecule in the asymmetric unit was completed using Coot [109] and Phenix Refine [110] using stimulating annealing for initial refinements to limit bias. Final data collection statistics and refinement parameters are shown in Table 2.1. Figures were generated with PyMOL (The PyMOL Molecular Graphics System, 1.8.0 Schrödinger, LLC).

2.2.4 Steady-State Kinetic Studies

The kinetic parameters of PLpro-Ubl2 and PLpro-ΔUbl2 for catalyzing the reaction of Ub-AMC (LifeSensors, Inc.), ISG15-AMC (Boston Biochem/R&D Systems), and a peptide substrate, Z-RLRGG-AMC (Bachem), were determined using a modified protocol in [94]. The release of the fluorophore, 7-amino-4-methylcoumarin (AMC) group ($\lambda_{ex} = 360 \pm 40$ nm, $\lambda_{em} = 460 \pm 40$ nm) from the substrate was monitored in the form of relative fluorescence units as a function of time (RFU/min) using a BioTEK Synergy H1 multimode microplate reader at 25 °C. The concentration of Ub-AMC was varied from 0.08 μM up to 30 μM. Reactions were initiated with 5.2 nM PLpro for low Ub-AMC (<10 μM) concentrations and 1.3 nM PLpro for high concentrations (≥10 μM) to ensure initial rates were captured. For the ISG15-AMC assay, the concentration of substrate was varied from 0.02 μM to 12 μM using 0.39 nM PLpro to initiate hydrolysis. The concentration of Z-RLRGG-AMC was varied from 1.6 μM up to 75 μM initiating peptide hydrolysis with 1.0 μM PLpro. The initial rates of the reaction were converted to initial velocity (v; μM/min) using the extinction coefficient (Δε; RFU/μM) of product or the maximum amount of AMC that is released from the reaction. The reaction rates (v/[E]; min^{-1}) measured in triplicate were plotted as a function

of substrate concentration, [S]. For saturating substrates, kinetic parameters, k_{cat} and K_m, were determined using the SigmaPlot (v12) enzyme kinetics module from the non-linear regression Michaelis-Menten equation (2.5):

$$\frac{v}{[E]} = \frac{k_{cat}[S]}{K_m + [S]} \tag{2.5}$$

where k_{cat} is defined as the number of substrate molecules hydrolyzed by PLpro per minute per active site and K_m represents the substrate concentration where the reaction rate is half-maximal. From these kinetic parameters, the catalytic efficiency (k_{cat}/K_m) of the enzyme was determined. The standard deviation of the k_{cat}/K_m was calculated using the following equation (2.6):

$$\Delta\left(\frac{k_{cat}}{K_m}\right) = \frac{k_{cat}}{K_m}\sqrt{(\frac{\Delta k_{cat}}{k_{cat}})^2 + (\frac{\Delta K_m}{K_m})^2} \tag{2.6}$$

where Δk_{cat} and ΔK_m are the associated errors from the k_{cat} and K_m values, respectively. For the nonsaturating peptide substrate, the apparent k_{cat}/K_m, values were approximated using a linear regression module in GraphPAD Prism6 from the following equation (2.7):

$$\frac{v}{[E]} = \frac{k_{cat}}{K_m}[S] \tag{2.7}$$

2.2.5 Ub_2 and Ub_4 Chain Cleavage Assays

A Ub_2 panel containing various isopeptide chain linkages (Boston Biochem) of Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, Lys63, and linear Ub_2 (UbiQ explorer planel) were incubated with 160 nM PLpro-Ubl2 or PLpro-ΔUbl2 at 25 °C for 2 hours in reaction buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM DTT). Reactions without enzyme served as a negative control. After 2 hours reaction mixtures were quenched with LDS sample buffer (Life Technologies) and then loaded onto a gradient (4-12%) SDS-PAGE at Ub_2 concentrations of 0.5 μg and 1.5 μg per well for Lys6-Lys63 and linear, respectively. Ub_4 cleavage assays for Lys48- and Lys63-linked chains (LifeSensors, Inc.) were performed using the same enzyme concentrations, buffer compositions,

and negative control as described above. The reaction was quenched at five different time points from 5 minutes to 2 hours and analyzed by SDS-PAGE loaded at 0.5 μg per well. To determine the relative amount of Ub species present in each time point reaction, we used semi-quantitative analysis in ImageJ 1.48v [111].

2.2.6 CD melting studies

Protein samples at 1-2 μM were loaded in a 10 mm quartz cell (Starna Cells) with magnetic stir bar in 2.5 ml of 0.1 M potassium phosphate (pH 7.5). The CD signal was measured at 220 nm as the temperature was increased at a step interval of 0.4 °C and rate of 1.0 °C/min as proteins were denatured using a Chirascan circular dichroism (CD) spectrometer (Applied Photophysics) equipped with a temperature control bath (Quantum Northwest Inc.). The CD ellipticity to increasing temperatures was normalized by using the following equation (2.4), where the CD_{min} represents the CD signal at 78 °C and CD_{max} represents the CD signal at 30 °C:

$$\text{Normalized ellipticity} = \frac{\text{CD}_{\text{experiment}} - \text{CD}_{\text{min}}}{\text{CD}_{\text{max}} - \text{CD}_{\text{min}}} \times 100 \quad (2.8)$$

The average T_m for each enzyme was calculated from three independent experiments by determining the maxima of the first derivative peak using Microsoft Excel and SigmaPlot (v12).

2.2.7 Determination of IC_{50} Values for inhibitors under Reducing and Non-reducing conditions

The exact compound **4** investigated by Lee *et al.* was purchased from Life Chemicals, Inc. (CAS # 2993-05-7), referred to as the company code name (**F2124-0890**) in this study [107]. Inhibition assays were performed in the presence and absence of 5 mM DTT for vUSPs, MERS-CoV PLpro, SARS-CoV PLpro, and MHV PLP2 (400 nM, 20 nM, and 3 μM final enzyme concentrations), at a 100 μL scale and three human USPs, USP7, USP17, and USP28 (1 nM, 5 nM, and 10 nM final enzyme concentrations), at a 30 μl scale. Assay conditions for PLpro-ΔUbl2 resembled those used for MERS-CoV PLpro. The enzymatic activities of SARS-CoV PLpro and

MHV PLP2 were monitored with 50 μM Z-RLRGG-AMC in assay buffer (50 mM HEPES, pH 7.5, 0.1 mg/ml BSA). The enzymatic activity of MERS-CoV PLpro was monitored with 75 μM Z-RLRGG-AMC in assay buffer described by Lee *et al.* [107]. The enzymatic activities of human USPs were monitored with 0.5 μM Ub-AMC in the same buffer [107]. The inhibitor was incubated with enzymes for at least 5 minutes before the reaction was initiated with substrate monitoring fluorescence using the BioTEK Synergy H1 multimode microplate reader at 25 °C. The % Inhibition was calculated across concentration ranges of inhibitor using equation (2.9), where the positive control represents reaction rate without addition of inhibitor. For the negative control, no background catalysis was observed without addition of enzyme and was therefore not used in this equation.

$$\% \text{ Inhibition} = \left[\frac{\text{Rate}_{\text{pos}} - \text{Rate}_{\text{sample}}}{\text{Rate}_{\text{pos}}}\right] \times 100 \tag{2.9}$$

For IC_{50} determination, data in non-reducing conditions were fit to the Hill equation (2.10):

$$\% \text{ Inhibition} = \frac{\% \text{ Inhibition}_{\text{max}}}{\frac{IC_{50}}{[\text{F2124} - 0890]^{n_H}} + 1} \tag{2.10}$$

where n_H represents the Hill coefficient. In the presence of reducing agent, data were either fit to the Hill equation, the Michaelis-Menten equation (2.11), or if the percent inhibition was under 50% at 200 μM **F2124-0890**, it was assumed that the IC_{50} value was >200 μM.

$$\% \text{ Inhibition} = \frac{\% \text{ Inhibition}_{\text{max}}[\text{F2124} - 0890]}{IC_{50} + [\text{F2124} - 0890]} \tag{2.11}$$

2.2.8 Biosensor assay

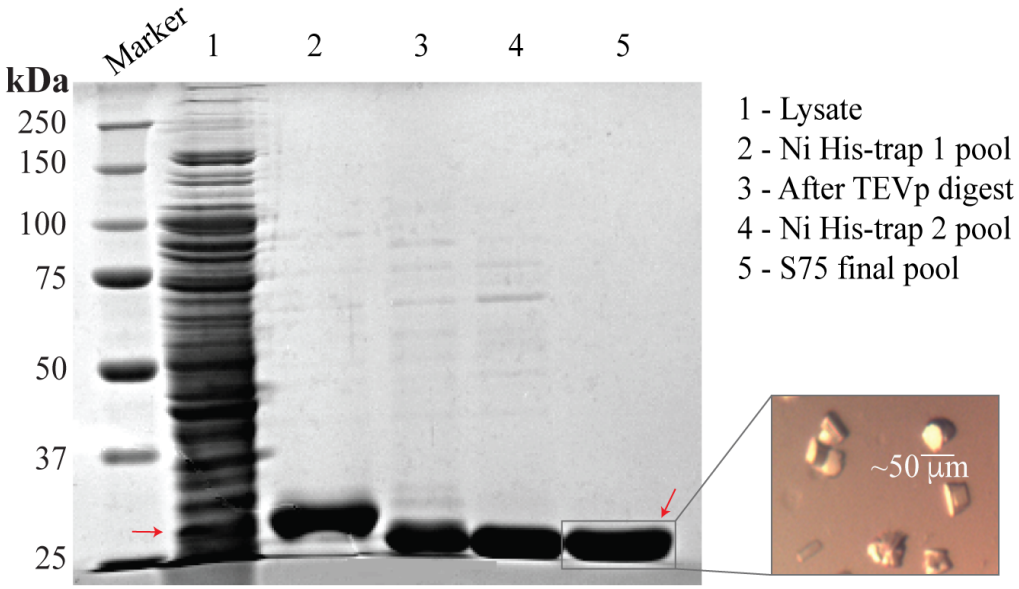
Our collaborators at Loyola University in Dr. Susan Baker's lab, a graduate student Robert Mettelman and a lab technician Amornrat O'Brien, tested the inhibition of compound F2124-0890 against MERS and SARS-CoV PLpro in cells. HEK293T cells were transfected with 150 ng pGlo-30F-RLKGG, 25 ng pRL-TK (Promega) and 150 ng of plasmid DNA expressing the indicated viral protease or empty vector. At 20 hours post-transfection, cells were lysed with 1X passive

lysis buffer (Promega) and 25 µl of lysate was assayed for luciferase activity using 96-well white bottom plates (Corning) and dual luciferase activating reagents (Promega). To evaluate expression of protein, western blotting for detection of the V5-epitope-tagged protease was performed. Briefly, 25 µl of lysate was mixed with 25 µl of 2X sample buffer, and proteins were separated by electrophoresis on 10% SDS-PAGE, transferred to PVDF membrane, and probed with mouse anti-V5 (Invitrogen) as previously described [112]. HRP-conjugated goat-anti-mouse (Southern Biotech) was used as the secondary antibody with detection using the Western Lighting Chemiluminescence Reagent Plus (Perkin Elmer) and visualized using a FluoroChemE Imager.

2.3 Results

2.3.1 X-ray Structure Determination of MERS-CoV PLpro without the N-terminal flanking Ubl2 domain

Since the functionality of MERS-CoV PLpro was observed to be independent of the Ubl2 domain in cellular assays, we sought to determine the structure of the segregated catalytic core.



Sample	Total Protein (mg)	Total Units (µM/min)	Specific Activity (µM/min/mg)	Purification Fold	Yield (%)
Lysate	1520	4328	2.9	1.0	100
Ni His-trap 1 Pool	20	1246	62	22	29
S75 final Pool	15	2007	134	47	46

Figure 2.2 SDS-PAGE (10%) and activity analysis summarizing the purification progress and optimized crystal of MERS PLpro-ΔUbl2

Top Right – The protein sample pool after each purification step is shown. Top Left – The final optimized crystal achieved from screening protein from the S75 final pool sample. Bottom – Purification summary table using 250 nM Ub-Rho.

The designed MERS-CoV PLpro-ΔUbl2 plasmid expresses only the catalytic core of PLpro (Figure 2.1.B). The 60 amino acids encoding for the Ubl2 domain at the N-terminus and a

single cysteine from the C-terminus of PLpro were removed. The cysteine was removed to aid in crystallization. PLpro-ΔUbl2 was expressed and purified using nearly identical conditions and procedures as those for MERS-CoV PLpro flanked with the N-terminal Ubl2 domain. The MERS-CoV PLpro-ΔUbl2 catalytic domain remained stable throughout purification with minimal precipitation and no activity loss. The resulting molecular weights for MERS-CoV PLpro-Ubl2 and PLpro-ΔUbl2 after removal of the octa-histidine tags were ~36 kDa and ~29 kDa, respectively (Figure 2.1.C). The first purification attempt of PLpro-ΔUbl2 is shown in Figure 2.2 and was analyzed as pure by SDS-PAGE and activity assays. The S75 final pool from this first attempt exhibited a specific activity of 134 μM/min/mg with a purification fold of 47 and a yield of 46%. The increase in yield (%) from the Ni His-trap 1 Pool and the S75 final pool could be due to the accuracy of the protein concentration determined by the Bradford assay of the cuvette-based Bio-Rad Bradford assay described in [102]. Approximately 15 mg of protein was obtained from a 2 L culture, and a diffractable quality crystal was produced from this protein preparation.

PLpro-ΔUbl2 crystallized in space group P 1 2 1 with one biologically active monomer in the asymmetric unit. X-ray data were collected to 1.95 Å, and the final X-ray data collection and refinement statistics are summarized in Table 2.1. The final X-ray structural model for MERS-CoV PLpro-ΔUbl2 has R-values of R_{work} = 16.6% and R_{free} = 18.8%. The MERS-CoV PLpro-ΔUbl2 structure contains only the catalytic domain of PLpro with its three subdomains: thumb, fingers, and palm (Figure 2.2.A). The secondary structure arrangement of the catalytic domain is identical to the structure described in [82] with seven α-helices (six in the thumb domain, one in the fingers domain), fourteen total β-strands (four in the thumb domain, four and two partial strands in the fingers domain, and four and two partial strands in the palm domain), and one 3_{10}-helix (η) in the fingers domain.

Table 2.1 Data-collection and structure refinement statistics for MERS PLpro-ΔUbl2

PDB entry	**5KO3 (MERS PLpro-ΔUbl2)**
Data-collection parameters	
Beamline	21-ID-F
Wavelength (Å)	0.98
Space group	P121
Unit cell dimensions:	
a, b, c (Å)	83.7, 30.5, 86.7
α, β, γ (°)	90, 116, 90
Resolution (Å)	100-1.95 (1.98-1.95)[a]
Number of reflections observed	351471
Number of unique reflections	29405
R_{merge} (%)[b]	7.1 (64.6)
R_{pim} (%)[c]	4.0 (35.8)
$CC_{1/2}$ (%) *in highest shell*	77.3
CC* (%) *in highest shell*	93.4
I/σI	26.0 (2.4)
% Completeness	98.8 (98.3)
Redundancy	4.1 (4.2)
Refinement	
Resolution range (Å)	43.2–1.95 (2.02–1.95)
No. of reflections in working set	29042
No. of reflections in test set	1470
R_{work} (%)[d]	16.6 (20.5)
R_{free} (%)[e]	18.8 (23.3)
Wilson B factor ($Å^2$)	29.5
Average B factor ($Å^2$)	39.6
RMSD from ideal geometry	
Bond length (Å)	0.015
Bond angle (deg)	1.36
Ramachandran plot	
Most favored (%)	95.8
Allowed (%)	3.8
Disallowed (%)	0.4

[a]Values in parentheses are for the last (highest resolution) shell.

[b]$R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of a given reflection, and $\langle I(hkl)\rangle$ is the mean intensity of symmetry-related reflections.

[c]$R_{pim} = \sum_{hkl}\sqrt{(\frac{1}{n}-1)}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i(hkl)$, where n is the multiplicity for multiplicity-weighted R_{merge}.

[d]$R_{work} = \sum_{hkl}\big||F_{obs}| - |F_{calc}|\big| \,/\, \sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[e]R_{free} was calculated using 5% of the data set chosen at random that were excluded from the refinement.

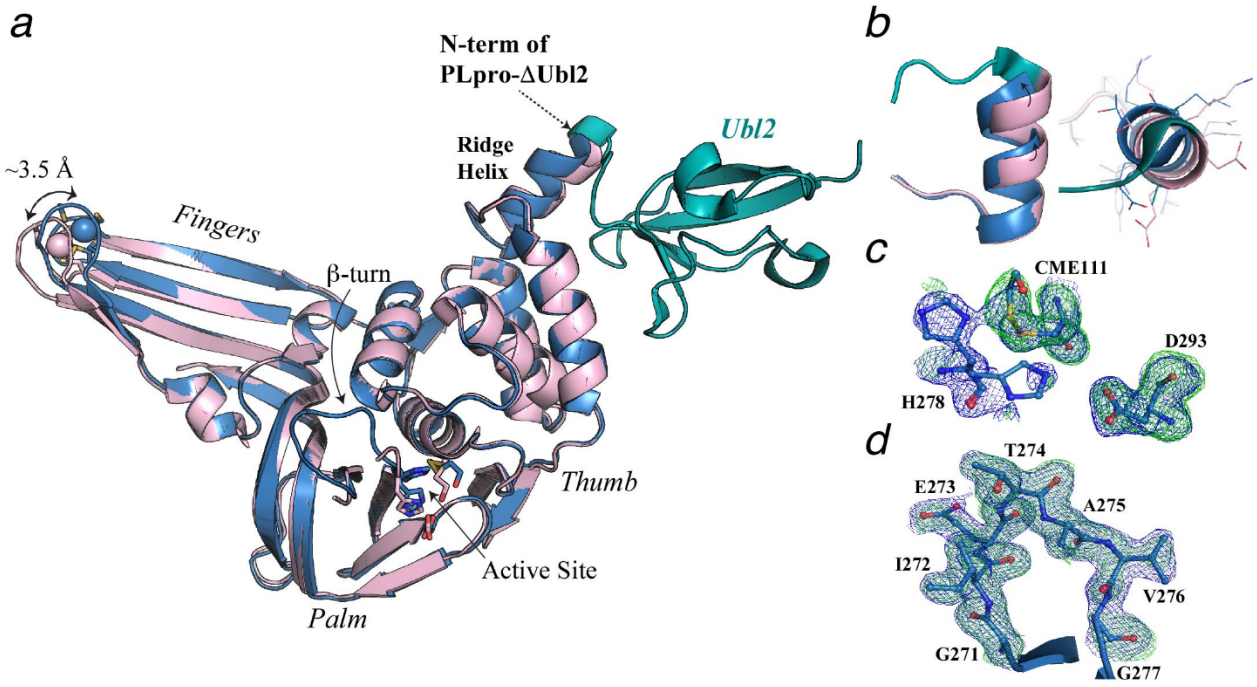


Figure 2.3 X-ray crystal structure of MERS-CoV PLpro-ΔUbl2 (blue, PDB code 5KO3) superimposed with the MERS-CoV PLpro (pink, PDB code 4P16)

(A) Overall structure of PLpro with its three subdomains and active site labeled. The Ubl2 domain of MERS-CoV PLpro is colored in green and the zinc atom of the fingers domain is shown as a sphere with zinc-coordinating cysteines represented as sticks. The arrow indicates the difference in the position of the zinc atom between the two structures. The ridge helix is indicated, and the position of where the Ubl2 domain was truncated at N-terminus of PLpro-ΔUbl2 is indicated with an arrow. The β-turn-substrate-binding loop of MERS-CoV PLpro-ΔUbl2, observed in this study, is also indicated as 'β-turn'. (B) Magnified views (elongated-left and helical wheel projection-right) of the ridge helix show the similarity between the two structures. (C) Electron density maps associated with the catalytic triad residues (Cys111, His278, Asp293) and β-mercaptoethanol (CME111) are shown in green mesh (F_o-F_c) and blue mesh ($2F_o$-F_c). F_o-F_c electron density omit maps, where the catalytic triad residues were omitted from the calculations, are contoured to 3σ. Final $2F_o$-F_c maps are contoured to 1σ. His278 was observed to reside in two positions after occupancy refinement – one at slightly higher occupancy (0.55). (D) Electron density maps associated with substrate-binding loop (residues 271-277) containing the β-turn are shown in green mesh (F_o-F_c) and blue mesh ($2F_o$-F_c) and are contoured to 3σ and 1σ, respectively. The entire loop could be modeled into the observable density and is represented as sticks. Atoms in Panels C-D are colored as follows; nitrogens (dark blue), oxygens (red), sulfur (yellow) carbons (light blue).

Similar to the MERS PLpro-Ubl2 structure, the catalytic Cys111 in the MERS-CoV PLpro-ΔUbl2 structure is also modified by β-mercaptoethanol (BME) with a partial occupancy of 0.83 (Figure 2.3.C) [82]. Weak electron density for the catalytic His278 is observed in final F_o-F_c omit maps (Figure 2.3.C). Due to the partially modified Cys111, His278 is observed to occupy at least

one alternative position in order to accommodate the bulky, modified Cys111 residue. On the other hand, and in contrast to previously reported unbound MERS-CoV PLpro structures, including Protein Data Bank (PDB) codes 4P16 [82], 4PT5 (unpublished), 4REZ [83] and 4RNA [107], strong (>3σ in F_o-F_c maps) and well-defined electron density is observed for the flexible loop encompassing residues 271-277 in the MERS-CoV PLpro-ΔUbl2 structure (Figure 2.3.D). The substrate-binding loop has been observed in the Ub-bound MERS-CoV PLpro structures [83, 113]. As a result, we were able to readily build and refine the entire loop that is responsible for substrate binding and inhibitor recognition in CoV PLpros and PLP2s [93, 94]. Altogether, the X-ray structural data suggests that the catalytic domain of MERS-CoV PLpro is highly stable in the absence of the Ubl2 domain.

2.3.2 Structural Comparison of PLpro with and without the Ubl2 domain

There are currently four structures of MERS PLpro flanked with the N-terminal Ubl2 domain that have been determined in the absence of any bound ligand [82, 83, 107]. To determine if the Ubl2 domain elicits an effect on the conformation of the MERS-CoV PLpro catalytic core, we superimposed the structures of PLpro-ΔUbl2 and PLpro-Ubl2 (Figure 2.3.A,B). The MERS-CoV PLpro catalytic domain is observed to adopt a conformation that is nearly identical to the structure with the Ubl2 domain intact. The resulting root-mean-square-deviation (RMSD) is 0.4 Å when the C_α of 254 residues in PLpro-ΔUbl2 are aligned with C_α of 258 residues in PLpro-Ubl2. The catalytic triad, Cys111-His278-Asp293, aligns well for both enzymes except for the His occupying the conformation near the modified cysteine group. Truncation of the Ubl2 domain appears to cause only slight deviations in the ridge helix of the thumb domain due to the loss of two helix residues, Thr61 and Ala62. We also observe some variation in the position of the zinc atom in the fingers domain, which is reminiscent of the open and closed conformations observed in the Ub-bound complex [83]. This observation suggests that the zinc-fingers binding motif has

high flexibility, which may provide an explanation as to why there is weaker electron density and increased B-factors associated with this region that includes residues 225-230.

2.3.3 Enzymatic Activity of MERS-CoV PLpro is unaffected by the loss of the Ubl2 domain

With the goal of testing the enzymatic activities of MERS-CoV PLpro with and without the Ubl2 domain, more untagged PLpro was purified using the protocol described in 2.2.1, which included a Ni His-trap 1, TEVp digest, followed by a Ni His-trap 2 column. Several protein batches of each construct were expressed and purified. A representative purification table summarizes the purification progress for PLpro-Ubl2 and PLpro-ΔUbl2 is shown in Table 2.1. Typically, for the Ni His-trap 2 Pool, the final specific activity of PLpro-ΔUbl2 construct using 250 nM Ub-AMC ranged from ~200-400 μM/min/mg while a specific activity of ~500 μM/min/mg was typically obtained for the PLpro-Ubl2 construct. One difference that was observed between constructs throughout purification was the efficiency of the TEVp digest. While His_8-tag was completely removed for PLpro-ΔUbl2 after overnight incubation at 4 °C, only ~50% of PLpro-Ubl2 tag was cleaved even after an extended incubation. To troubleshoot the inefficient cleavage, a second TEVp digest and Ni His-trap 2 column was performed on the uncleaved tagged PLpro-Ubl2 construct (after eluting this sample with 5 CV of 100% buffer B off the Ni His-trap 2 column). The untagged PLpro-Ubl2 construct Ni His-trap 2 Pool #1 (first round) and Ni His-trap 2 Pool #2 (second round) have similar specific activities (Table 2.2). Therefore, to increase yield of the untagged form of the PLpro-Ubl2 construct, any tagged PLpro-Ubl2 remaining after the first Ni His-trap 2 can be re-dialyzed and purified with a second Ni His-trap 2 column.

Table 2.2 Purification summary of MERS PLpro-Ubl2 and PLpro-ΔUbl2 from 2 L culture of E. coli BL21-DE3 using 250 nM Ub-AMC as a substrate

Sample	Total Protein (mg)	Total Units (μM/min)	Specific Activity (μM/min/mg)	Purification Fold	Yield (%)
PLpro-Ubl2					
Lysate	525	16078	31	1.0	100
Ni His-trap 1 Pool	46	32527	707	23	202
Ni His-trap 2 Pool #1	9	4169	474	15	26
Ni His-trap 2 Pool #2	8	4215	562	18	26
PLpro-ΔUbl2					
Lysate	1079	9707	9	1.0	100
Ni His-trap 1 Pool	24	2350	99	11	24
Ni His-trap 2 Pool	5	1821	372	41	19

To determine if the enzymatic activity of PLpro is dependent on the Ubl2 domain, we determined the steady-state kinetic parameters of the PLpro-ΔUbl2 and PLpro-Ubl2 catalyzed hydrolysis of three different Ub-based substrates, Ub-AMC, ISG15-AMC and Z-RLRGG-AMC, which are commonly used to assess PLpro DUB, deISGylating and proteolytic activities. The kinetic assays for each substrate were performed side-by-side with each enzyme in the same assay plates under identical assay conditions. The kinetic response of each enzyme to increasing concentrations of substrate are shown in Figure 2.4. We were unable to reach saturation with the Z-RLRGG-AMC peptide substrate up to concentrations of 75 μM, a concentration that begins to approach the concentration range whereby the inner filter effect for the AMC fluorophore can confound the assay [94]. Therefore, this first-order range of the kinetic data were fit to a line to obtain the slope, which is the apparent k_{cat}/K_m or catalytic efficiency. In contrast, both enzymes could be saturated with ISG15-AMC and Ub-AMC substrates, and the kinetic data were fit to the

Michaelis-Menten equation to obtain individual k_{cat} and K_m values. The resulting kinetic parameters for both enzymes against all three substrates are summarized in Table 2.3.

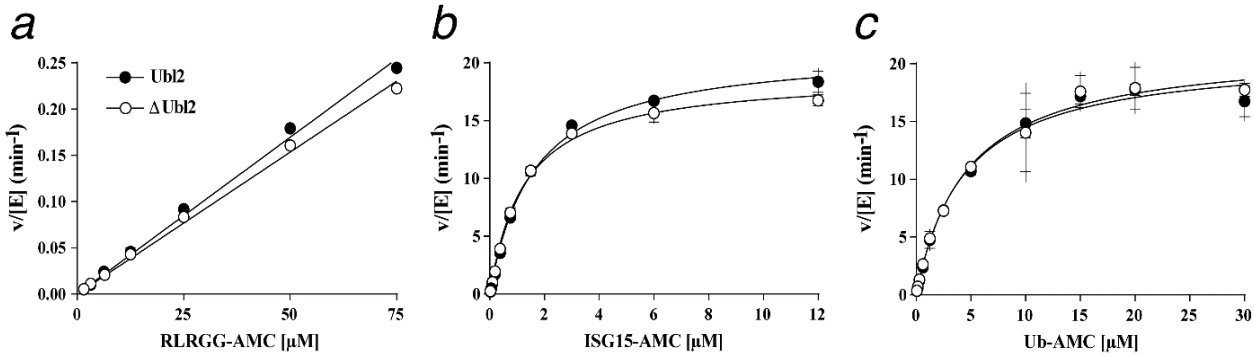


Figure 2.4 The kinetic response of MERS-CoV PLpro-Ubl2 (black circles) and PLpro-ΔUbl2 (white circles) to the increasing concentrations of three different ubiquitin-based substrates

(A) RLRGG-AMC. (B) ISG15-AMC. (C) Ub-AMC. Data in panel a failed to reach saturation and were therefore fit to a line whereas panels B and C were fit to the Michaelis-Menten equation. All data were measured in triplicate, and the error bars represent the standard deviations from the triplicate data.

As suggested by our previous work in cells, the kinetic parameters for PLpro-Ubl2 and PLpro-ΔUbl2 catalyzed hydrolysis of Ub-AMC, ISG15-AMC and Z-RLRGG-AMC substrates are nearly identical for the two enzymes. Compared to the k_{cat}/K_m value for the Z-RLRGG-AMC peptide, both enzymes hydrolyze Ub-AMC (~1,600 times) and ISG15-AMC (~5,000 times) more efficiently. MERS-CoV PLpro with and without the Ubl2 domain hydrolyzes ISG15-AMC substrate ~3-fold more efficiently than Ub-AMC although the turnover numbers, k_{cat}, are identical. The higher catalytic efficiency is mainly due to the lower K_m value observed for ISG15-AMC. Assuming that the $K_d \cong K_m$, PLpro may bind ISG15 3-fold tighter compared to Ub. The results of these kinetic studies support our hypothesis that the Ubl2 domain is not required for MERS-CoV PLpro DUB, deISGylating, and proteolytic activities *in vitro*.

Table 2.3 Kinetic parameters for PLpro-Ubl2 and PLpro-ΔUbl2 using three different FRET Ub-based substrates

Enzyme	**Substrate**		
Kinetic Parameter	**RLRGG-AMC**	**Ub-AMC[a]**	**ISG15-AMC[a]**
PLpro-Ubl2			
k_{cat}/K_m (μM^{-1} min^{-1})	0.003[b]	4.8 ± 0.4	13.4 ± 0.7
k_{cat} (min^{-1})	N/A	20.8 ± 0.5	21.2 ± 0.3
K_m (μM)	N/A	4.4 ± 0.4	1.6 ± 0.1
PLpro-ΔUbl2			
k_{cat}/K_m (μM^{-1} min^{-1})	0.003[b]	4.7 ± 0.6	15.1 ± 0.8
k_{cat} (min^{-1})	N/A	21.4 ± 0.8	19.0 ± 0.3
K_m (μM)	N/A	4.6 ± 0.6	1.3 ± 0.1

[a]Steady-state values reported as a mean ± standard deviation, were determined from a minimum of triplicate measurements (best-fit slopes shown in Figure 2.4). [b]Value of k_{app} with nonsaturating substrate approximates k_{cat}/K_m.

2.3.4 MERS-CoV PLpro Ub chain specificity and poly-Ub processing are not dependent on the Ubl2 domain

Recent studies demonstrated that MERS-CoV PLpro with an intact Ubl2 domain has broad Ub chain specificity based on cleavage of various diubiquitin (Ub_2) chains with isopeptide linkages [105]. MERS-CoV PLpro has been proposed to use a single monoubiquitin recognition sub-site, S1, for Ub binding to processes all Ub chains [94, 105]. It is possible that the MERS-CoV Ubl2 domain could function to assist PLpro in its ability to discriminate between different Ub_2 linkages. We therefore tested the ability of MERS-CoV PLpro with and without an intact Ubl2 domain to process different Ub_2 isopeptide linkages, including Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63, and linear Ub_2, which is linked via the amino-terminal Met residue. We incubated each of the substrates with MERS-CoV PLpro-Ubl2 or PLpro-ΔUbl2 for 2 hours and analyzed the cleaved products by SDS-PAGE. The results are shown in Figure 2.5, and they indicate that the Ubl2 domain does not impact the ability of PLpro to recognize different isopeptide linkages.

MERS-CoV PLpro is capable of recognizing and processing all Ub_2 linkages except the peptide linkage (linear). Linkages that were efficiently cleaved to nearly all monoubiquitin after the time course were Lys11, Lys48, and Lys63. However, linkages with partially reacted or unreacted Ub_2 species were Lys6, Lys27, Lys29, and Lys33. These results are also consistent with recent findings that MERS-CoV PLpro without the Ubl2 domain was still able to be modified by a K48-linked Ub_2 warhead supporting the hypothesis that the Ubl2 domain is not involved in substrate recognition [105].

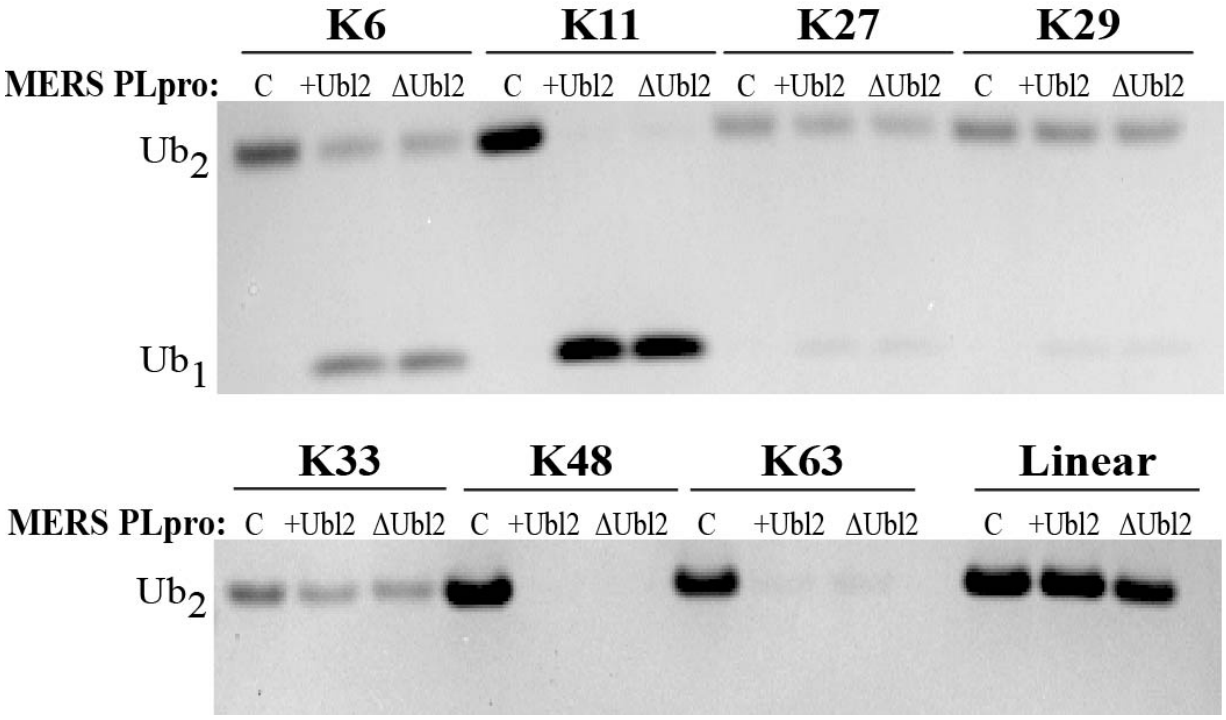


Figure 2.5 The Ubl2 domain is not required for MERS PLpro Ub2-processing specificity
The cleavage of different Ub_2 linkages (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, Lys63, and linear) mediated without addition of enzyme (C) or by 160 nM PLpro-Ubl2 (+Ubl2) or PLpro-ΔUbl2 (ΔUbl2) for 2 hours analyzed by SDS-PAGE.

Since Lys48- and Lys63-linked polyubiquitin chains are preferentially utilized in host innate immune response pathways [66, 68], we further evaluated the kinetics of hydrolysis of Lys63- and Lys48-linked tetraubiquitin (Ub_4) chains by PLpro-Ubl2 or PLpro-ΔUbl2 to evaluate if the Ubl2 domain is involved higher-order polyubiquitin chain processing. The cleavage assay

was performed over a 2 hour time course, and reaction products were analyzed at various time points from 5 minutes to 2 hours using SDS-PAGE (Figure 2.6.A,B). Ub_4 substrates without addition of enzyme served as the negative control. Progress curves for PLpro catalyzed hydrolysis of Ub_4 were constructed by quantifying the amount of each ubiquitin species for individual bands, i.e. monoubiquitin (Ub_1), diubiquitin (Ub_2), triubiquitin (Ub_3) and unreacted tetraubiquitin (Ub_4), at each time point based on their band intensities normalized to 100% Ub_4 or fully liberated Ub_1.

The cleavage assays for Lys63-Ub_4 (Figure 2.6.A) and Lys48-Ub_4 (Figure 2.6.B) are nearly identical for the reactions catalyzed by MERS-CoV PLpro with and without the Ubl2 domain.

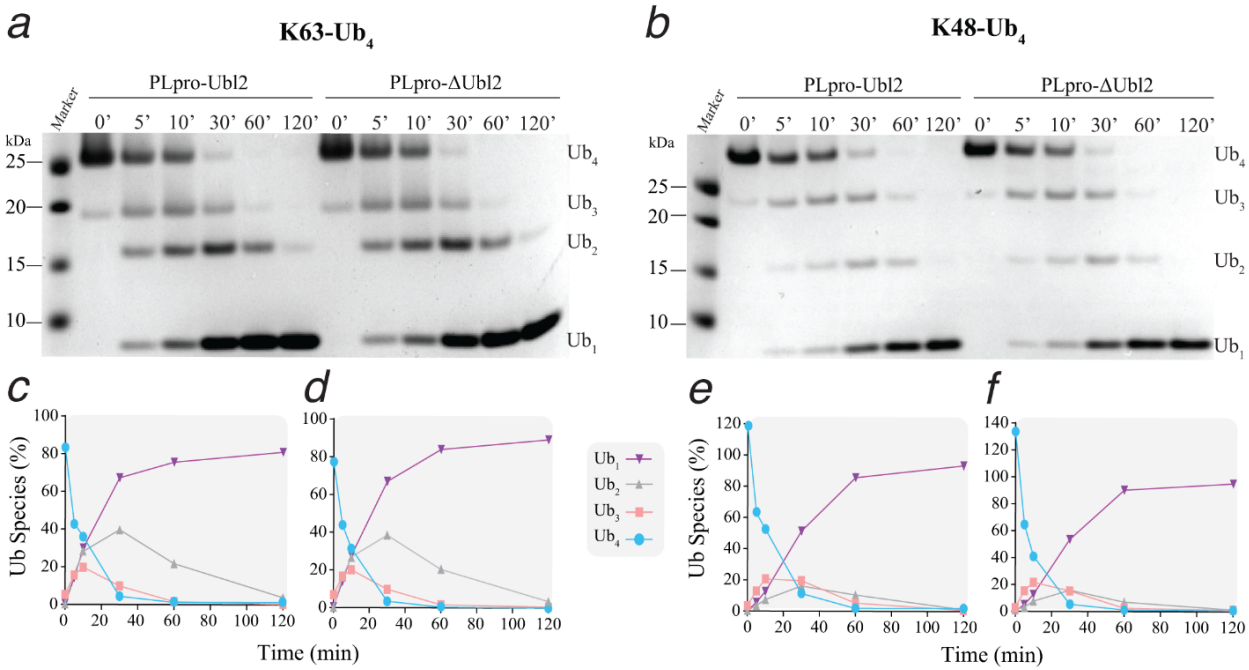


Figure 2.6 Ubl2 domain is not required for MERS PLpro Ub4 processing

(A,B) Time course of cleavage processing of Lys63-Ub_4 (A) and Lys48-Ub_4 (B) by 160 nM PLpro-Ubl2 and PLpro-ΔUbl2. Samples were quenched with sample buffer at the indicated time points and products were analyzed by SDS-PAGE. The negative control (time point '0') was incubated without addition of enzyme. Time course hydrolysis curves for Lys63-Ub_4 (C,D) and Lys48-Ub_4 (E,F) representing the relative amount of each Ub species at each reaction time point catalyzed by PLpro-Ubl2 (C,E) and PLpro-ΔUbl2 (D,F). A total intensity summation or total amount of substrate loaded was averaged across all time points and used to estimate the amount of Ub present for individual bands. At time point zero, Lys48-Ub_4 was a greater intensity compared to the average, hence >100%.

Both substrates are readily converted into monoubiquitin species after 2 hours consistent with previous reports [94, 105]. Based on the progress curves, for Lys48-linked chains, a steady increase of each Ub species was observed with no apparent accumulation of other Lys48-Ub forms. Unexpectedly, after 30 minutes there was approximately two times more unreacted Lys63-Ub_2 (Figure 2.6.C,D) compared to Lys48-Ub_2 (Figure 2.6.E,F) indicating that the MERS-CoV PLpro may be less efficient at processing Lys63-linked compared to Lys48-linked chains. Together, the results suggest that the Ubl2 domain of MERS-CoV PLpro is not involved in any significant recognition and cleavage of polyubiquitin chain substrates and that a second ubiquitin recognition subsite S2 (SUb2) does not seem to exist for MERS-CoV PLpro in contrast to SARS-CoV PLpro [77].

2.3.5 MERS-CoV PLpro-ΔUbl2 is thermally stable

Although MERS-CoV PLpro-ΔUbl2 and MERS-CoV PLpro have nearly identical substrate recognition patterns and kinetic parameters at room temperature, it is possible that the Ubl2 domain may alter the thermostability of the catalytic domain at higher temperatures. To test this possibility, we performed circular dichroism (CD) melting studies on each enzyme by monitoring the CD signal as a function of temperature and then determining the thermal melting temperatures (T_m). The average T_m values from three independent experiments were found to be 61.2 ± 0.3 °C for MERS-CoV PLpro-Ubl2 and 60.7 ± 0.2 °C for PLpro-ΔUbl2. The normalized CD melting curves are shown in Figure 2.7. The change in thermal melting temperature (ΔT_m) of only 0.6 °C indicates that MERS-CoV PLpro is a structurally stable enzyme even in the absence of the N-terminal Ubl2 domain. In fact, MERS-CoV PLpro-ΔUbl2 remains folded at higher temperatures compared to MHV PLP2, which unfolds at temperatures less than 50 °C [101].

Therefore, we conclude that the MERS-CoV Ubl2 domain does not stabilize or destabilize the catalytic domain *in vitro*.

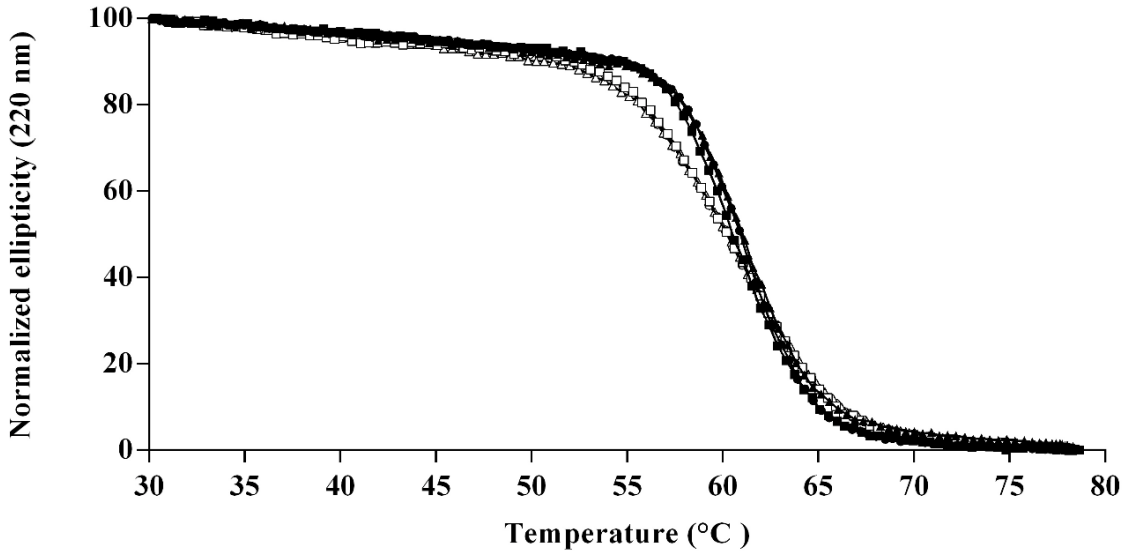


Figure 2.7 Normalized CD melting curves of MERS PLpro-Ubl2 (shaded circles, squares, triangles) and PLpro-ΔUbl2 (white circles, squares, triangles)

Data from three independent experiments for both enzymes are plotted and used to determine the thermal melting temperatures. Each replicate is plotted with a different shape.

2.3.6 Evaluation of compound F2124-0890, a purported inhibitor of MERS-CoV PLpro

Recent studies by Lee *et al.* reported that compound **4**, commercial code **F2124-0890** (Life Chemicals), inhibits MERS-CoV and SARS-CoV PLpro activity with IC_{50} values in the low micromolar range [107]. We performed an independent analysis of the ability of compound **F2124-0890** to inhibit MERS-CoV PLpro and MERS-CoV PLpro-ΔUbl2 in addition to other viral and cellular USPs. First, we varied concentrations of compound **F2124-0890** and measured the percent inhibition of viral papain-like proteases, including MERS-CoV PLpro-Ubl2, MERS-CoV PLpro-ΔUbl2, SARS-CoV PLpro and MHV PLP2 both in the absence and presence of 5 mM DTT a reducing agent (Figure 2.8). Compound **F2124-0890** equally inhibits both MERS-CoV PLpro-Ubl2 and MERS-CoV PLpro-ΔUbl2 both in the absence and presence of reducing agent and

supports our general observation that the Ubl2 domain does not influence MERS-CoV PLpro catalytic function. The resulting IC_{50} values are given in Table 2.4.

Next, we performed the same experiment with three different human USPs, USP7, USP17 and USP28, both in the absence and presence of reducing agent (5 mM DTT). What is immediately apparent from the data presented in Figure 2.8 is that compound **F2124-0890** is capable of significant inhibition of all of the viral and human USP enzymes but only in the absence of reducing agent. In the presence of reducing agent, the inhibition of these enzymes is either eliminated or significantly reduced. Our results are summarized in Table 2.4, and they stand in strong contrast to those obtained by Lee *et al.* who reported that compound **F2124-0890**, compound **4** in their studies, strongly inhibits MERS-CoV (IC_{50} = 6.2 μM) and SARS-CoV (IC_{50} = 10.9 μM) PLpro in the presence of reducing agent, 5 mM DTT or 2 mM GSH [107]. Another striking observation from the data presented in Table 2.4 is that compound **F2124-0890** is non-selective, i.e. it is promiscuous, potently inhibiting multiple USP homologs.

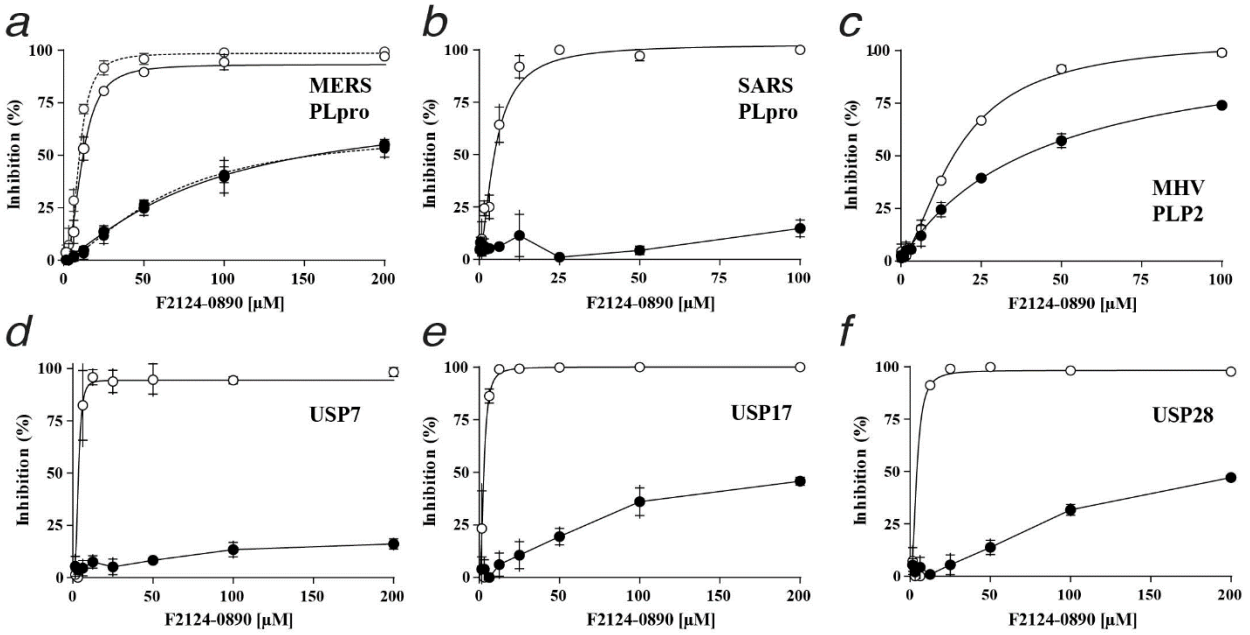


Figure 2.8 In vitro analysis of F2124-0890 with proteases under nonreducing conditions reveals its lack of specificity and lack of potency under reducing conditions

Percent (%) inhibition was plotted as a function of increasing inhibitor concentrations in the absence of DTT (white circles) and presence of DTT (black circles) against viral and human proteases, MERS PLpro-Ubl2 (solid line, A) and PLpro-ΔUbl2 (dotted line, A), SARS PLpro (B), MHV PLP2 (C), USP7 (D), USP17 (E), and USP28 (F). Error bars represent the standard deviations obtained from triplicate data.

We further evaluated **F2124-0890** using a recently described cell-based assay, named the pGlo biosensor assay, for CoV PLpro activity [112]. HEK-293T cells were transfected with plasmid DNA expressing an inactive form of luciferase and either the wild-type or a catalytic mutant (Cys111Ala) of PLpro. We found that expression of SARS-CoV and MERS-CoV PLpro activate the biosensor and that a selective SARS-CoV PLpro inhibitor that we developed, compound **3e**[93, 112], blocks SARS-CoV PLpro activity. In contrast, addition of **F2124-0890** had no effect on either MERS-CoV (Figure 2.9.A) or SARS-CoV (Figure 2.9.B) PLpro activity in the biosensor assay, which supports and confirms our *in vitro* results that compound **F2124-0890** loses potency in physiological reducing environments.

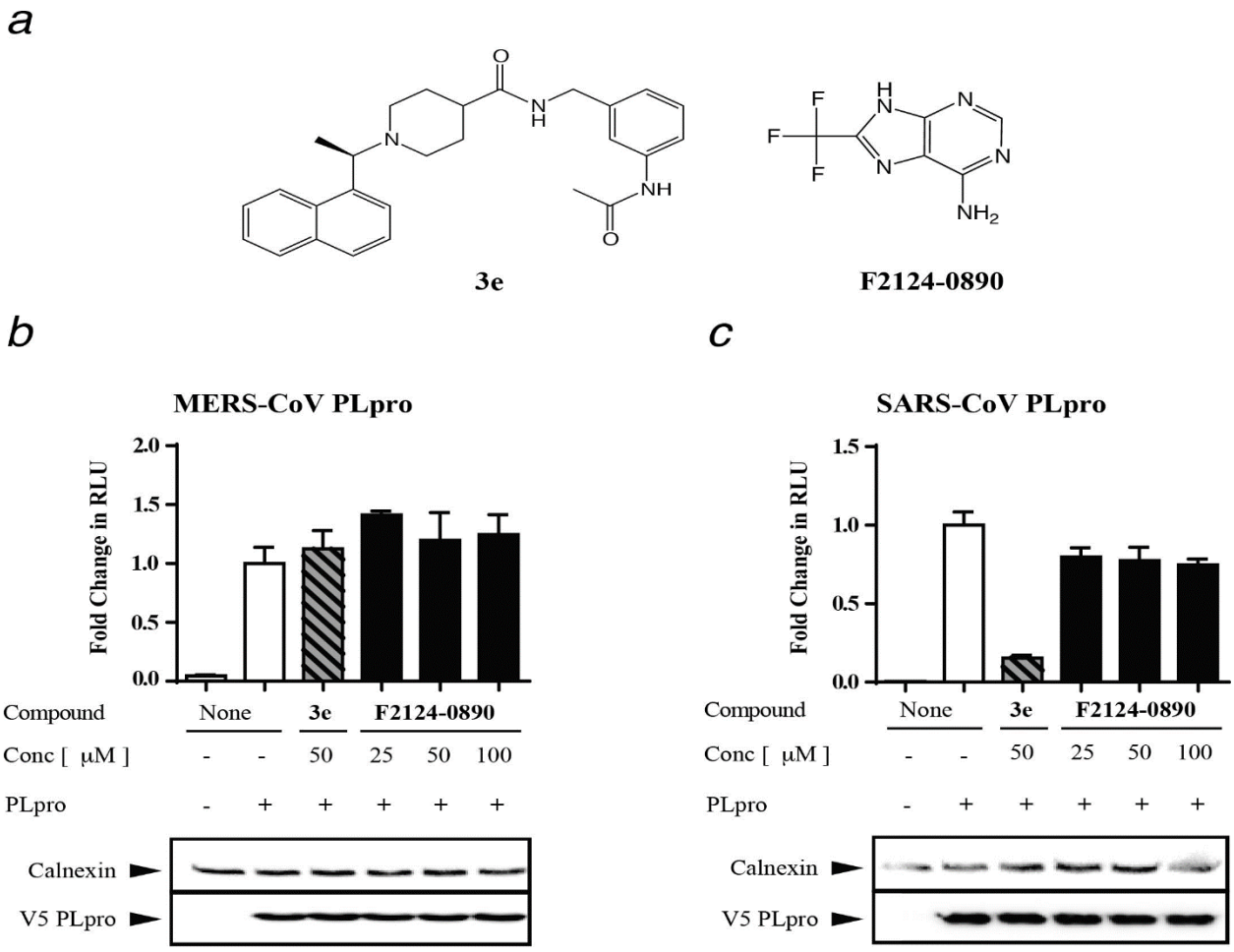


Figure 2.9 Luciferase-based biosensor assay reveals lack of inhibitory potency of F2124-0890 against MERS-CoV and SARS-CoV PLpro

(A) Structure of SARS-CoV PLpro inhibitor **3e** and putative inhibitor **F2124-0890**. (B,C) A biosensor endpoint assay was used to determine if compound **F2124-0890** inhibits SARS- or MERS-PLpro activity in cells. The graph shows representative results of each compound with error bars depicting standard deviation. Relative expression of PLpro and the cellular protein calnexin were determined by western blot using monoclonal antibodies specific for V5 (PLpro) and calnexin. Each compound concentration was evaluated in triplicate and experiments were performed three times.

Table 2.4 Effect of Reducing agent on the Nonselective Inhibition of F2124-0890 towards Viral and Human Proteases

	Enzyme	Lee *et al.*[b]	- DTT		+ DTT	
			IC_{50} (μM)[c]	**Hill Coefficient (n_H)**	**IC_{50} (μM)[a]**	**Hill Coefficient (n_H)**
Viral	MERS PLpro-Ubl2	6.2 ± 0.9	11.8 ± 0.4	2.5 ± 0.2	87.6 ± 24.9[c]	1.3 ± 0.2
	MERS PLpro-ΔUbl2		8.7 ± 0.3	2.6 ± 0.2	64.6 ± 10.9[c]	1.5 ± 0.2
	SARS PLpro	10.9 ± 0.9	4.9 ± 0.4	1.6 ± 0.2	-	
	MHV PLP2		18.0 ± 0.9	1.6 ± 0.1	43.1 ± 3.9[d]	
Human	USP7		3.8 ± 0.8	3.7 ± 1.5	-	
	USP17		2.7 ± 0.2	2.2 ± 0.3	>200	
	USP28		4.0 ± 0.2	2.2 ± 0.1	>200	

[a]-, no inhibition. [b]IC_{50} values previously reported from ref [107]. [c]Data fit to the Hill Equation. [d]Data fit to the Michaelis-Menten Equation.

2.4 Discussion

We investigated the function of the Ubl2 domain of MERS-CoV PLpro in substrate recognition and catalysis, structural stability and inhibition by a purported small molecule inhibitor. We also report the first X-ray crystal structure of a CoV PLpro or PLP2 without its flanking Ubl2 domain. We found using X-ray crystallography that removal of the Ubl2 domain from MERS-CoV PLpro did not alter the structure of the catalytic domain significantly nor did it change the structural stability as determined by melting temperatures derived from CD melting curves. Steady-state kinetic studies revealed that the Ubl2 domain associated with MERS-CoV PLpro is not required for its enzymatic function, including DUB, deISGylating, and proteolytic activities. In addition, examining MERS-CoV PLpro-mediated catalysis towards different polyubiquitin substrates with different isopeptide linkages revealed that the Ubl2 domain does not influence high-order Ub chain processing or Ub chain specificity. Overall, these studies indicate that the core catalytic domain of MERS PLpro is a robust enzyme that can be used in cell-based and *in vitro* assays making it highly amenable for high-throughput screening to evaluate potential inhibitors.

Whether other domains within the MERS-CoV nsp3 participate in the discrimination of different Ub chain linkages still remains unclear. Our findings suggest, however, that the Ubl2 domain is likely not involved in this function. MERS-CoV PLpro is capable of cleaving a variety of different chain linkages, though some chains are more favored. The substrate recognition of MERS-CoV PLpro is therefore similar to MHV PLP2 [54]. Our findings support that the recognition and cleavage of these Ub_2 chains by MERS-CoV PLpro is independent of the presence of the Ubl2 domain.

MERS-CoV PLpro can also hydrolyze both Lys48- and Lys63-linked polyubiquitin chains to monoubiquitin at equal rates either with or without the Ubl2 domain. The rapid processing of

both chains to monoubiquitin suggests that MERS-CoV PLpro, similar to MHV PLP2, utilizes monoubiquitin recognition at a single S1 sub-site (SUb1) to cleave all Ub chains and ISG15. In contrast, SARS-CoV PLpro prefers to utilize Ub_2 recognition to hydrolyze substrates. For example, SARS-CoV PLpro hydrolyzes Lys48-linked chains more efficiently by using two Ub binding sites across S2-S1 sub-sites as opposed to Lys63-linked chains, which are recognized by the single SUb1 sub-site [77]. The structural basis for Ub_2 recognition was recently revealed via the crystal structure of SARS-CoV PLpro in complex with Lys48-linked Ub_2 supporting the previous models [77, 81, 94, 105]. In the SUb2 binding pocket, the ridge helix of SARS-CoV PLpro, which is immediately adjacent to the Ubl2 domain, actively engages with the distal Ub of the K48-linked Ub_2 substrate [81]. We propose that the ridge helix of MERS-CoV PLpro may not actively participate in substrate binding as with SARS-CoV PLpro. However, we can only conclude from the two residues removed from the ridge helix that these specific residues are not involved in ISG15 or polyubiquitin catalysis, as no further mutagenesis was done in this region. Altogether, our findings, coupled with the aforementioned studies, clearly show that MERS-CoV and SARS-CoV PLpro utilize different mechanisms when recognizing and cleaving host proteins, and specifically, MERS-CoV PLpro likely does not contain the SUb2.

Our original X-ray structure of SARS-CoV PLpro revealed for the first time the presence of Ubl domains in CoV nsp3s, and it established the first known vUSP defining a new class [63]. This seminal work on vUSPs set the stage for new discoveries, including a bioinformatics study on the homologies of human USPs which revealed that the Ubl fold, resembling the β-grasp architecture of Ub, is predicted to be present in at least 16 human USPs [114]. The locations of these Ubl domains can be found at either their N- and C-termini or embedded within their catalytic core. A plethora of Ubl domains residing within human USPs suggest that they may play a significant functional role in the tightly orchestrated process of protein degradation as well as other

critical signaling processes. Thus far, the function of Ubl domains in USPs have been attributed to the alteration of enzyme catalysis and specificity, or the recruitment of binding partners to mediate processes, including cellular localization, trafficking, and signal transduction [114].

Although extensive studies have characterized Ubl domains from human USPs, the function of viral Ubl domains in biological systems remains elusive. Initially, the function of the Ubl2 domain was evaluated in cell-culture with SARS-CoV PLpro where the Ubl2 domain was found to be essential for PLpro's ability to act as an IFN antagonist and inhibit IRF3 phosphorylation or NF-κB signaling [76]. Interestingly, when the Ubl2 domain was removed from PLpro, SARS-CoV PLpro-ΔUbl2 maintained its DUB and protease activity. However, further characterization of SARS-CoV PLpro-ΔUbl2 *in vitro* was not pursued due to its instability during expression and purification. Interestingly, in a recent study both SARS-CoV and MERS-CoV PLpro-ΔUbl2 constructs were expressed and purified *in vitro*. Both enzymes maintained their catalysis towards a K48-linked Ub_2 warhead as compared to PLpro with the Ubl2 domain [105]. These results suggest that the Ubl2 domain may not be involved in the mechanism of substrate recognition or catalysis for PLpros. However, the Ubl2 domain was found to be important for MHV PLP2 thermal stability [101]. Previous work evaluated the effect of a conserved single Val787Ser mutation in the Ubl2 domain, which was found to decrease PLP2 enzymatic activity at physiological temperatures and attenuate mutant virus in mice. We propose that a similar mutation could provoke an "unraveling effect" if introduced in other Ubl2 domains adjacent to PLpro. However, in the case of MERS-CoV PLpro, we show that we can completely remove the Ubl2 domain of MERS-CoV PLpro, and the resulting PLpro catalytic core is still able to maintain its stability and catalysis *in vitro* and in cells [94].

We also investigated whether a compound, **F2124-0890**, which was recently reported to be a potent and selective inhibitor of both MERS-CoV and SARS-CoV PLpro [107], could also

inhibit MERS-CoV PLpro without its Ubl2 domain. This purine analogue was first synthesized in 1958 as a potential anticancer agent, and in the late 1980s and early 1990s, the compound was used as a reactant for designing arrhythmia and antiviral drugs as well as compounds set to regulate plant growth [115, 116]. For over 20 years, **F2124-0890** was seldom reported in the literature until in 2014 when it was identified as an inhibitor of SARS-CoV 3CLpro by Lee *et al.* [117]. In that study, **F2124-0890**, referred to as compound **14,** was identified as an inhibitor of SARS-CoV 3CLpro via a high-throughput screen. **F2124-0890** was found to inhibit SARS-CoV 3CLpro with mixed-type inhibition (IC_{50} of 13.9 μM). In 2015, Lee *et al.* also performed a similar HTS study this time against MERS-CoV and SARS-CoV PLpro and they identified the same compound, **F2124-0890**, which they referred to in that study as compound **4**. It was reported that compound **4** inhibited PLpro from both CoVs with low micromolar IC_{50} values and, based on the mechanisms of inhibition of each enzyme, was predicted to act as a competitive inhibitor against MERS-CoV PLpro and an allosteric inhibitor of SARS-CoV PLpro. The binding mechanism was described to take place at either the active site pocket or an unknown allosteric site.

In contrast to the aforementioned studies of Lee *et al.*, we demonstrate that **F2124-0890** (a.k.a. compound **4** or **14**) is non-selective under non-reducing conditions; inhibiting all viral and human cysteine proteases tested and confirming that it is a pan-assay interference compound (PAIN) [118]. We also evaluated the inhibitory ability of **F2124-0890** under reducing conditions by either placing DTT in the biochemical assays or using the natural reducing environment of the cell. We found that reducing agent either greatly diminished or eliminated the ability of **F2124-0890** to inhibit MERS-CoV and SARS-CoV PLpro and the USPs tested. Possible explanations as to why the compound may only show efficacy under non-reducing conditions could be that the compound binds in a non-specific manner to viral and human USPs and promotes reversible-oxidation of the active site cysteine. Another explanation could be that chemical reducing agents

may directly compete against the inhibitor for binding to the active site. Reducing agents, such as DTT and BME, can modify active site cysteines. The bulky modified cysteine formed by the BME reducing agent observed in the crystal structure supports the fact that the inhibitor may not be able to bind under the reducing conditions due to the encumbered active site pocket. However, **F2124-0890** is unable to inhibit MERS-CoV or SARS-CoV PLpro under the natural reducing conditions of a cell indicating that inhibition by this compound is complex and likely non-specific. It is clear from our data and the data presented in the literature that **F2124-0890** has poor selectivity among cysteine proteases, lacks inhibitory potency in cell-based assays, and has greatly reduced or no inhibitory potency in *in vitro* assays in the presence of reducing agents. Therefore, **F2124-0890** is likely a PAIN that should not be pursued further as a lead compound for therapeutic development or other uses.

In summary, the catalytic core of MERS-CoV is stable and highly active without its Ubl2 domain. MERS PLpro-ΔUbl2 exhibits the same substrate specificity profile of MERS-CoV with an intact Ubl2 domain suggesting that the Ubl2 domain is not necessary for normal MERS-CoV PLpro function. MERS PLpro-ΔUbl2 is highly amenable to enzyme inhibitory studies, and it easily forms crystals that diffract to high resolution. Overall, the properties of MERS PLpro-ΔUbl2 suggest that it may be an ideal construct for structure-based inhibitor design efforts.

CHAPTER 3. STRUCTURAL AND KINETIC INSIGHTS INTO SARS CORONAVIRUS PAPAIN-LIKE PROTEASE INTERACTION WITH UB VERSUS ISG15

A version of this chapter has been published in the journal article [119] in collaboration with Dr. Scott Pegan's lab at University of Georgia.

3.1 Introduction

The SARS-CoV outbreak of 2002 illustrated that CoV infection can result in severe respiratory disease in humans, far beyond mild cold-like symptoms. To permit the design of antivirals and/or vaccines for intervention, beta-CoV SARS-CoV PLpro has been widely studied. SARS-CoV PLpro has viral polyprotein processing activity for CoV replication and deubiquitinating (DUB) and deISGylating (deISG) function for the suppression of the innate-immune response, specifically through antagonizing IFN, chemokine, and cytokine production [68, 119].

Over the past decade, there has been extensive work focused on understanding the recognition of SARS-CoV PLpro with the Ub substrate but less of a focus on ISG15. Several SARS-CoV PLpro structures bound with Ub has been solved, including Ub-aldehyde [77] and the modified K48-linked di-Ub substrate [81]. In contrast, no structure has been solved with SARS-CoV bound with full-length ISG15 to date. Based on mutagenesis experiments, a structural model of SARS-CoV PLpro bound with ISG15 has been proposed [77], but its ISG15 specificity is still obscure.

In this Chapter, kinetic data of SARS-CoV PLpro mutants are discussed with altered DUB and deISG activity profiles compared to the wild-type. Our collaborators in Dr. Scott Pegan's lab, provide insight into SARS-CoV PLpro interaction with ISG15 by determining SARS-CoV PLpro in complex with the C-terminal domain of ISG15 [119]. These kinetic and structural data may help in the design of SARS-CoV PLpro mutants for live-attenuated CoV design.

3.2 Materials and Methods

3.2.1 Site-directed mutagenesis, expression, and purification of SARS PLpro mutants

SARS-CoV pET-15b-PLpro mutants (residues 1541-1855 of the SARS-CoV viral polyprotein) were generated using site-directed mutagenesis and the QuickChange® approach (Agilent) by Dr. Yahira Báez [37]. The resulting plasmids were transformed and then isolated from electrocompetent XL1-Blue cells and sent for sequencing at the Purdue Genomics Core Sequencing Facility to confirm the correct single point transformants were generated in SARS-CoV PLpro.

The expression and purification for the SARS-CoV PLpro wild-type and mutants were previously described [120]. BL21 (DE3) cells were transformed with pET-15b-PLpro mutant plasmid via electroporation, and a colony was used to inoculate a 5 mL (in a 50 mL falcon tube) LB culture containing 50 μg/mL carbenicillin, which was grown as described in section 2.2.1. The large scale 1 L super broth culture (media described in 2.2.1) with 50 μg/mL carbenicillin was inoculated with 5 mL of the starter culture and harvested by centrifugation (3,000 x g, 4 °C, 22 minutes) after shaking at 25 °C for 24 hours.

Pellets from a 1 L culture (ranging from 12-16 g) were resuspended in 50 mL buffer A (20 mM Tris, pH 7.5, 500 mM NaCl, 10 mM imidazole) plus a pinch of lysozyme and DNase I. Resuspended cells were sonicated using a Branson Digital Sonifier (70% amplitude; 10 minutes, 5.5 s pulses, 9.9 s delays), and lysate was clarified by centrifugation (29,000 x g, 4 °C, 22 minutes). Approximately 50 mL of lysate was loaded onto a 5 ml His-Trap FF column (GE Healthcare) that was pre-charged with Co^{2+}, which was equilibrated with buffer A at a flowrate of 2 mL/min. The column was washed with 5 CV of buffer A before eluting tagged PLpro with a linear gradient from 0% to 100% buffer B (20 mM Tris, pH 7.5, 500 mM NaCl, 500 mM imidazole) over 100 mL at a flow-rate of 2 mL/min, collecting 5 mL fractions. Active tagged PLpro was pooled based on

specific activity using the 50 μM RLRGG-AMC substrate. A final enzyme concentration of 0.00158 mg/mL or 0.04 μM PLpro was used in this assay with buffer containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, 150 mM NaCl, 2.5 mM DTT. Tagged PLpro was concentrated and buffer exchanged into storage buffer (20 mM Tris, pH 7.5, 20% glycerol, 10 mM DTT) using Amicon Ultrafiltration Centrifugal device (30 kDa MW cutoff) to approximately 1.4 – 16 mg/mL, and snap frozen with liquid nitrogen in 100-250 μL aliquots before storage in the -80 °C for kinetic experiments.

3.2.2 Functional studies of SARS-CoV PLpro Mutants

The steady-state kinetic parameters of SARS-CoV PLpro wild-type and mutants were determined for three different Ub-based fluorescent substrates, utilizing 7-amino-4-methylcoumarin (AMC), commonly used to assess the protease, deubiquitinating, and deISGylating activity of PLPs, including a small peptide substrate, Z-RLRGG-AMC (Bachem), Ub-AMC (LifeSensors, Inc.), and ISG15-AMC (Boston Biochem/R&D Systems). Kinetic assays with Ub-AMC and ISG15-AMC were performed on the same day and side-by-side in the same assay plate to directly compare the enzymatic activity of SARS-CoV PLpro to that of each of the mutants. The steady-state kinetic studies were also repeated for the wild type and mutants approximately 5 months apart, and the resulting duplicate data were combined for analysis. Kinetic assays with the peptide substrate were also performed in triplicate. For all experiments, the assay conditions (i.e., buffering conditions and assay volume, etc.) were set up as previously described [24]. The exception was that the stock substrates purchased from the vendors had different lot numbers. The steady-state kinetic data obtained from separate experiments performed on different days and with different substrate lot numbers helped ensure that the trends in the resulting kinetic parameters were reproducible.

The enzymatic activity of PLpro-mediated hydrolysis of the fluorophore AMC group was determined using a BioTEK Synergy H1 multimode microplate reader at 25 °C with an excitation wavelength of 360 nm (bandwidth = 40 nm) and an emission wavelength of 460 nm (bandwidth = 40 nm). The change in the relative fluorescence as a function of time (RFU/min) was monitored over a sufficient time period to allow the determination of the enzymatic rate in the steady-state region. For the ISG15-AMC assay, the substrate concentrations were varied from 0.2 μM up to 19.2 μM. The reactions were initiated by the addition of enzyme with the final enzyme concentrations as follows: 0.48 nM WT, 0.23 nM Gln233Glu, 0.23 nM Met209Ala, or 7.3 nM Arg167Glu. For the Ub-AMC assay, substrate concentrations were varied from 0.5 μM to 17.6 μM. The final enzyme concentrations were 3.7 nM WT, 7.3 nM Gln233Glu, 7.3 nM Met209Ala, or 0.23 nM Arg167Glu. For the Z-RLRGG-AMC assay, the concentrations of substrate were varied from 0.8 μM to 50 μM, and the final concentration of the wild-type enzyme was 0.14 μM. To capture the initial rate of peptide hydrolysis for the M209A mutant, we used a lower enzyme concentration of 25 nM. As is typical for SARS-CoV PLpro, the enzyme could not be saturated with the Ub-AMC and Z-RLRGG-AMC substrates. As such, the kinetic response of the enzyme to these substrates was linear, and thus, the data were fit to a line to approximate the catalytic efficiency (k_{cat}/K_m) for each enzyme. For the ISG15-AMC assays, the data were fit to the Michaelis–Menten equation to determine the associated kinetic parameters (k_{cat}, K_m, and k_{cat}/K_m) for each enzyme [121]. Saturation was not attained with the Arg167Glu mutant enzyme for ISG15-AMC, and therefore, these kinetic data were also fit to a line to determine the apparent kcat/Km. The errors associated with each kinetic parameter were obtained from the best-fit line or curves for each mutant. All data, from separate experiments, were included in the fits to arrive at the final errors (Table 4.1).

3.3 Results

3.3.1 X-ray structure of SARS-CoV PLpro-CISG15 complex

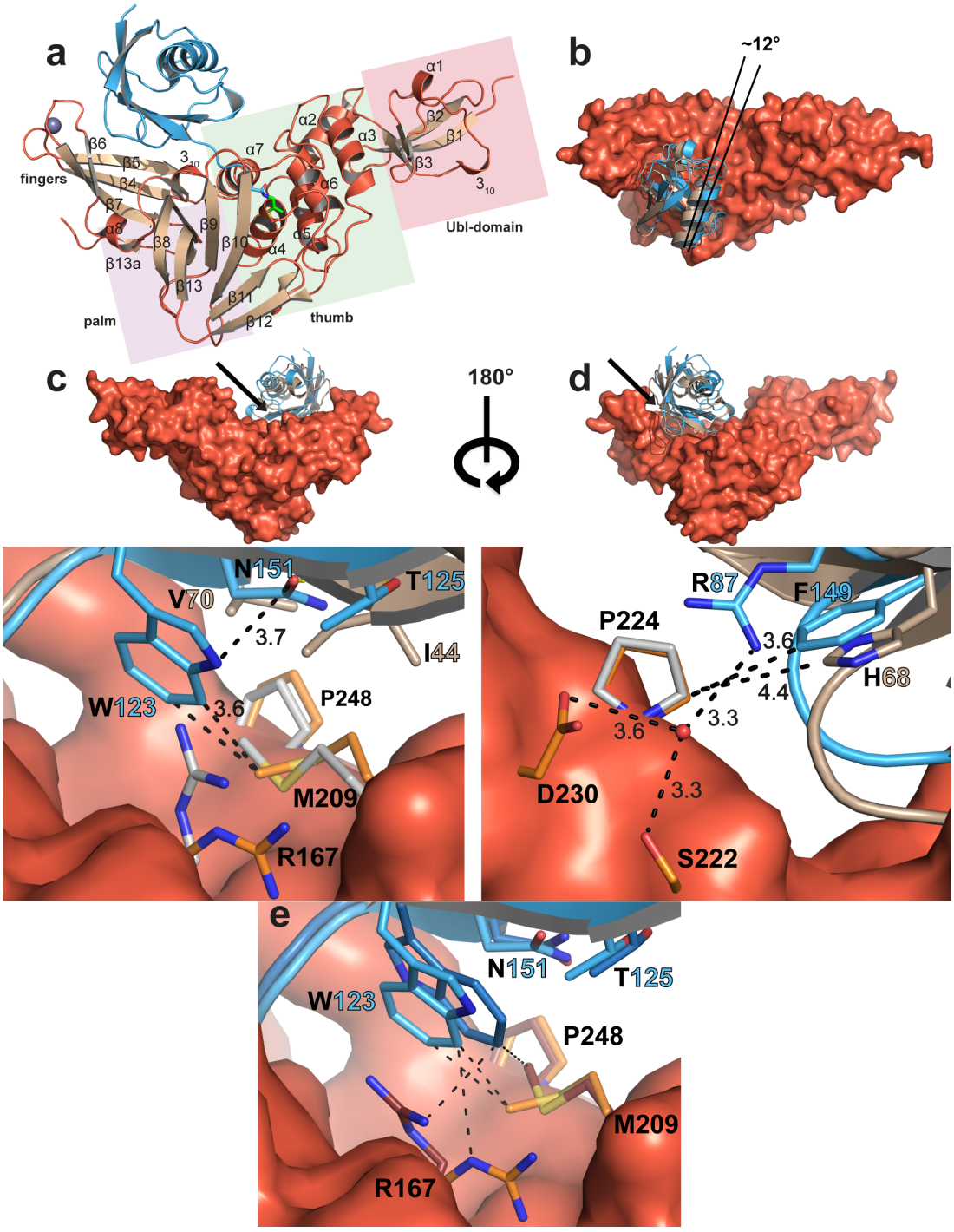
While X-ray crystal structures have been solved for SARS and MERS PLpro bound to Ub, our understanding of the interactions of these enzymes with ISG15 have been limited to enzymatic, mutational, and computational modeling studies [77, 122]. To gain molecular insight into the specific interactions between PLPs and ISG15, our collaborators in Dr. Scott Pegan's lab determined the X-ray structure of the SARS-CoV PLpro in complex with the C-terminal domain of human ISG15 (ChISG15) to 2.62 Å with two complete copies within the asymmetric unit. Specifically, they utilized a form of ChISG15 modified with propargylamine at the C-terminus (ChISG15-PA) to form a covalent modification with the active site cysteine. Attempts at co-crystallization of the full-length human ISG15 with SARS-CoV PLpro, as either a covalent adduct or non-covalently bound species were unsuccessful. Data collection and refinement statistics for the SARS-CoV PLpro-ChISG15 are reported in [119], as well as the methods for crystallization. The structure of SARS-CoV PLpro contained the classic tertiary fold associated with PLPs consisting of the -N-terminal ubiquitin-like 2 (Ubl2) and the thumb-palm-finger domains (Figure 3.1).

To unveil the key differences in how SARS-CoV PLpro may engage with Ub and ISG15, the structure of SARS-CoV PLpro bound with CISG15 was superimposed with mono-Ub (PDB entry: 4MM3). Besides a 180° flip of Trp123, both copies of the SARS-CoV PLpro-ChISG15 complex aligned with Ub in a similar manner (Figure 3.1.E). Globally, ChISG15 was shifted by approximately 12° compared to Ub (Figure 3.1.B) and distinct hydrophobic interactions were observed. Specifically, Arg167, Met209, and Pro248 of SARS-CoV PLpro are involved in Ub and ISG15-binding. While the putative hydrophobic patch of Ub, consisting of Ile44, Val70, and Leu8 [77], is used for Ub interaction, ISG15 lacks the Ub hydrophobic patch. Instead, the protease

interacts with hydrophobic residues of ISG15 in two regions. Met209 of the palm domain and Pro224 of the fingers domain recognizes Trp123 and Phe149 of ISG15 (Figure 3.1.C,D). The latter interaction is not adopted in the SARS-CoV PLpro bound to full-length hISG15, which was recently solved by Dr. Renata Everett (unpublished data). Additionally, less water-mediated interactions were observed in our structure compared with Ub [77]. This may suggest that water-mediated interactions play less of a role with hISG15 binding or potentially less waters are resolved in the crystal structure.

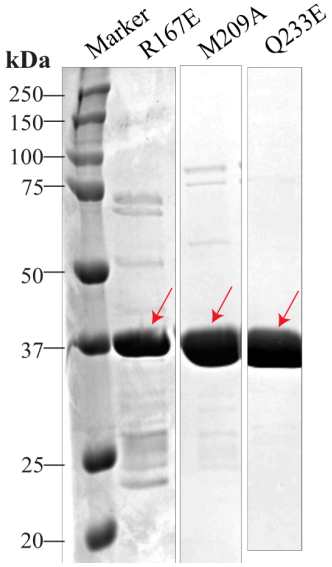
Figure 3.1 Analysis of SARS-CoV PLpro in complex with CISG15 versus Ub

(A) Cartoon representation of the SARS-CoV PLpro complexed with ChISG15 (blue). Secondary structure of the PLpro is designated, with helices and loops rendered in reddish orange and β-sheets rendered in wheat. The structural domains of the PLpro are identified, consisting of the fingers (white), palm (purple), thumb (green), and Ubl2 (red) labeled with colored boxes. The vinyl thioether propargylamine linker is colored green, and the Zn(II) ion in light purple. (B) Overlay of ChISG15 (blue) bound to SARS PLpro (reddish orange) compared to a Ub (light brown)-bound structure (PDB entry: 4MM3) based on a secondary structure alignment of the respective PLpros. The approximate degree of shift in the orientation of analogous α-helices is indicated. (C) Comparison of the interaction of ChISG15 (blue) versus Ub (light brown) with the hydrophobic patch of SARS-CoV PLpro, with the site of interaction within the overall structure indicated by an arrow. Side-chain conformations of the PLpro are colored orange for the ChISG15-bound structure, gray for the Ub-bound structure. Residues E127–D133 of ChISG15 and A46–T55 of Ub were removed for clarity. Inter- and intramolecular distances (Å) are shown by black dashes. (D) Comparison of ChISG15 versus Ub at the site of the additional hydrophobic interaction in the ChISG15-bound structure. Colored as in (C). Intermolecular distances (Å) are shown as black dashes. (E) Interchain variability of ChISG15 Trp123 in binding to SARS-CoV PLpro. The two copies within the asymmetric unit were overlaid based on a secondary structure alignment of the PLpros. Side chains of the PLpro colored in orange correspond to the ChISG15 colored light blue, and the PLpro side chains colored burgundy to the ChISG15 colored a darker blue. Intermolecular distances are colored in black. Residues E127–D133 of ChISG15 are omitted for clarity.


a
fingers
palm
thumb
Ubl-domain
b
~12°
c
180°
d
N151
T125
V70
3.7
I44
W123
3.6
P248
M209
R167
R87
F149
P224
3.6
4.4
H68
3.3
3.6
D230
3.3
S222
e
N151
T125
W123
P248
M209
R167

3.3.2 Purification summary of SARS-CoV PLpro mutants

To gain a better understanding of the residues mediating specificity for Ub or ISG15, mutations targeting the ubiquitin binding site 1 (SUb1) of SARS-CoV PLpro were designed with potential to modulate the recognition of these substrates. An initial kinetic characterization was done by Dr. Yahira Báez, and mutant enzymes Arg167Glu, Met209Ala, and Gln233Glu were sought for a complete kinetic characterization to confirm trends from the initial study [37]. These residues are located in the thumb, palm, and fingers domain, respectively.



Sample	Total Protein (mg)	Total Units (µM/min)	Specific Activity (µM/min/mg)	Purification Fold	Yield (%)
R167E	10	44244	4424	80.6	86.5
M209A	58.8	727721	12376	32.6	95.4
Q233E	92.4	845133	9146	1.8	8.4

Figure 3.2 Summary of SARS-CoV PLpro mutants' final purification pools and activity

Left image – 12.5% SDS-PAGE analysis of final protein pools of tagged PLpro after the first Co His-trap 1 column. Protein was loaded at 5 µg. Right table – PLpro mutant purification summary table of final pools assayed with 50 µM RLRGG-AMC. Activity of the lysate for each mutant was used to set the yield and purification fold at 100% and 1.0, respectively.

These tagged PLpro mutants at ~38 kDa were purified using a single Co^{2+} His-trap column. Since mutations were designed to target the recognition sites of Ub and ISG15 rather than near the active site, the small peptide Z-RLRGG-AMC was used to track activity throughout purification. The SDS-PAGE gel and activity analysis of the final pool, summarized in Figure 3.2, indicated that the sample was pure enough for kinetic experiments. The Gln233Glu sample expressed very

well and even the lysate exhibited a similar specific activity compared to the final pool, which is why the purification fold was not as high compared to the other mutants. Lower total protein (10 mg) was obtained for the Arg167Glu mutant while the other mutants were purified at much higher total protein yield (>50 mg from a 1 L cell culture). The specific activity of the Arg167Glu was also notably lower than the Met209Ala and Gln233Glu using the Z-RLRGG-AMC peptide while Met209Ala appeared to have a higher specific activity. However, these final pools were not assayed in the same assay plate for direct comparison.

3.3.3 Kinetics of SARS PLpro mutants with altered Ub and ISG15 specificities

To fully characterize the function of SARS-CoV PLpro mutants, Arg167Glu, Met209Ala, and Gln233Glu hydrolysis towards various substrates were determined. Fluorogenic Ub- and ISG15-7-amido-4-methylcoumarin (AMC) group, were used to determine the kinetic parameters of the mutants for DUB and deISG activity, respectively, while a small peptide Z-RLRGG-AMC substrate displaying the consensus recognition sequence was used to probe for changes in catalytic activity. The kinetic curves from the complete kinetic characterization and the extracted kinetic parameters from these curves are compiled in Figure 3.3 and Table 4.1, respectively. The catalytic efficiencies of the SARS-CoV PLpro mutants with a structural explanation is summarized in Figure 3.4 [119].

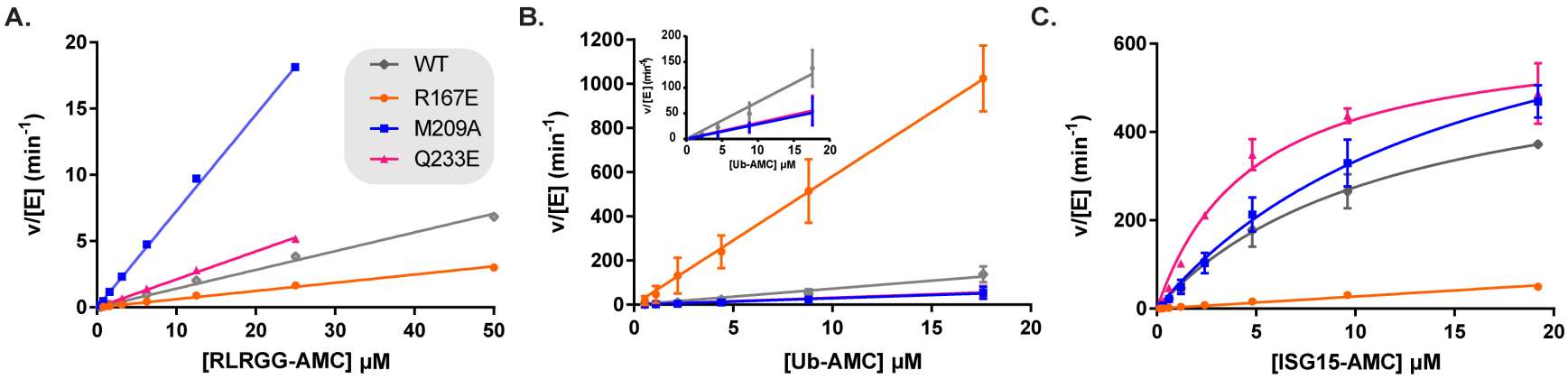


Figure 3.3 Kinetic curves of SARS PLpro wild-type and mutants for Ub-based substrates

(A-C) The turnover number as a function of substrate concentration was plotted for PLpro WT (gray), R167E (orange), M209A (blue), and Q233E (pink) for RLRGG-AMC (A), Ub-AMC (B), and ISG15-AMC (C). Points at 50 µM RLRGG-AMC for M209A and Q233E were omitted from the linear fit for the calculation of apparent k_{cat}/K_m due to curvature.

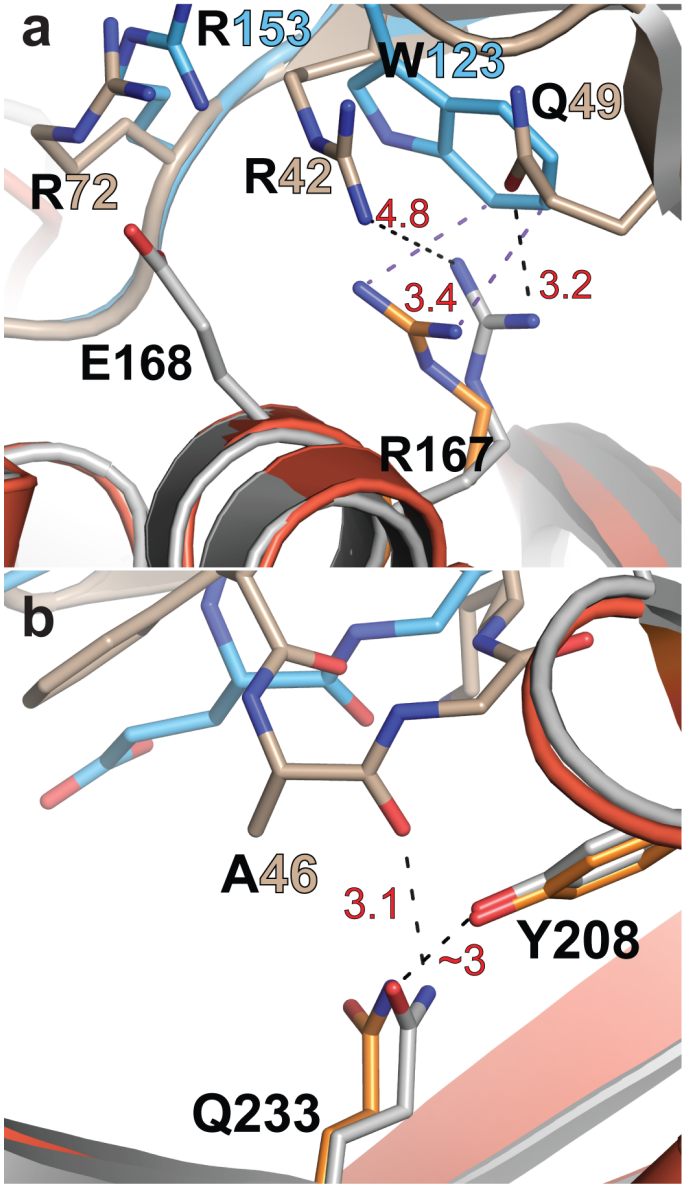
Table 3.1 Kinetic parameters of SARS PLpro WT and mutants with different Ub-based fluorescent substrates

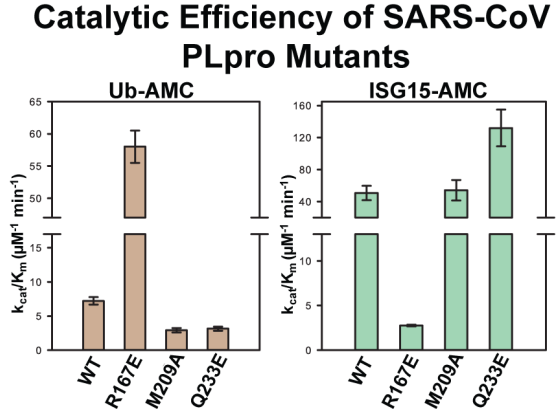
Substrate	**SARS PLpro Enzymes**			
Kinetic Parameter	**WT**	**R167E**	**M209A**	**Q233E**
RLRGG-AMC[a]				
k_{cat}/K_m (μM^{-1} min^{-1})	0.141 ± 0.002	0.062 ± 0.001	0.727 ± 0.007	0.211 ± 0.002
Ub-AMC[a]				
k_{cat}/K_m (μM^{-1} min^{-1})	7.22 ± 0.56	58.0 ± 2.5	2.90 ± 0.32	3.14 ± 0.32
ISG15-AMC[b]				
k_{cat}/K_m (μM^{-1} min^{-1})	50.7 ± 9.0	2.76 ± 0.10[a]	54.3 ± 12.8	132 ± 23
k_{cat} (min^{-1})	602 ± 49	-	865 ± 100	634 ± 40
K_m (μM)	11.9 ± 1.9	-	15.9 ± 3.3	4.80 ± 0.77

[a]Best-fit slope values derived k_{app} for nonsaturating substrates and approximates k_{cat}/K_m
[b]Steady-state values calculated from the Michaelis-Menten equation are reported from duplicate measurements

Figure 3.4 Structural basis and kinetic data of SARS-CoV PLpro mutants

(A, B) Sites within the ChISG15- and Ub-bound SARS-CoV PLpro structures corresponding to the mutations causing increased Ub-AMC activity (top) versus increased ISG15-AMC activity (bottom). Overlay of SARS-CoV PLpro bound to ChISG15 (blue, PDB entry: 5TL6) versus Ub (light brown, PDB entry: 4MM3)-bound structure was done based on a secondary structure alignment of the respective PLpros. Side-chain conformations of the PLpro are colored in orange for the ChISG15-bound structure and gray for the Ub-bound structure. Intermolecular distances to indicate proximity for SARS-CoV PLpro–Ub are shown as black dashes with the ones for SARS-CoV PLpro–ChISG15 in purple. (C) Catalytic efficiency of activity of SARS-CoV PLpro mutants toward Ub- and ISG15-AMC. Corresponding data found in Table 3.1.


a
R153
W123
Q49
R72
R42
4.8
3.4
3.2
E168
R167
b
A46
3.1
~3
Y208
Q233


C
Catalytic Efficiency of SARS-CoV PLpro Mutants
Ub-AMC
ISG15-AMC
kcat/Km (μM-1 min-1)
WT
R167E
M209A
Q233E

Consistent with the results throughout the purification, Met209Ala mutant exhibited increased apparent catalytic efficiency towards Z-RLRGG-AMC, roughly 5-fold enhancement compared to the wild-type levels, while its activity for ISG15- and Ub-AMC were for the most part retained and within 2-fold. From a structural prospective, the slight reduction of Ub-AMC activity (2.5-fold reduction) observed is not surprising, as Met209 engages with Ile44-centered hydrophobic patch of Ub. The alanine substitution would therefore shrink the hydrophobic pocket normally occupied by the methionine, but overall, this affect is minor. The observed increase of the activity toward Z-RLRGG-AMC is less straightforward. Met209 lacks direct interaction with the last 5 a.a. of Ub or ISG15. However, its replacement by alanine could sterically open up a space that is more accommodating for the artificial Z-adduct of the peptide.

While the change in activity of the Arg167Glu mutant compared to the wild-type was more pronounced, both Arg167Glu and Gln233Glu were found to have increased catalytic efficiencies for processing either Ub- or ISG15-AMC with a corresponding reduction in activity for the other. The net effect alters the selectivity of SARS-CoV PLpro enzyme, which normally prefers ISG15 to Ub by 7-fold. Arg167Glu mutant is over 8 times more efficient than the wild-type enzyme at hydrolyzing Ub-AMC but is about 20 times less efficient at hydrolyzing ISG15-AMC (Figure 3.4.C). In contrast, Gln233Glu is nearly 3-fold more efficient than wild-type at hydrolyzing ISG15 and 2-fold less efficient at hydrolyzing Ub. Based on catalytic efficiencies, Gln233Glu favors ISG15 to Ub over 40-fold while Arg167Glu favors Ub to ISG15 over 20-fold.

From the previously determined SARS-CoV PLpro Ub-bound structure, residue Arg167 has been shown to interact with Gln49 of Ub, but is close in proximity (~4.8 Å) to the charged Arg42 of Ub [77]. The charge flip with Arg167Glu may therefore introduce a more favorable electrostatic interaction with Arg42 of Ub while also maintaining its interaction with Gln49 of Ub. From the SARS-CoV PLpro ChISG15-bound structure, Arg167 is in close proximity to the

interaction between Trp123 and Met209. The glutamate substitution may affect Arg167 π–cation interaction with Trp123 of ChISG15 (Figure 3.1.E and 3.4.A). Additionally, the charge flip could disrupt the salt-bridge between Arg153 of ISG15 and Glu168 of PLpro. Arg153 of ChISG15 is also the P5 arginine of the RLRGG sequence, so this could also explain why Arg167Glu exhibited slightly lower Z-RLRGG-AMC activity compared to the wild-type. However, this reduction was only about 2-fold. Overall, Arg167 is important for ISG15 recognition and must accommodate Arg42 for Ub binding.

In comparison to Arg167Glu, the Gln233Glu mutant showed less deviations compared to the wild-type. The negatively charged group at position 233 may introduce charge repulsion with the backbone carbonyl of Ala46 in Ub to reduce Ub-AMC activity (Figure 3.4.B). In the ChISG15 structure, however, Gln233 of PLpro does not interact with ISG15 to explain the increase in ISG15-AMC activity. The full-length ISG15 bound structure may be necessary to determine if Gln233 is involved in ISG15 binding.

3.4 Summary

Chapter 3 provides structural and kinetic insights into how SARS-CoV PLpro recognizes and accommodates Ub versus human ISG15. The structure of SARS-CoV PLpro bound to the C-terminal, principle binding domain of ISG15 was determined and several mutants including Arg167Glu, Met209Ala, and Gln233Glu were characterized with altered activity profiles. These mutants targeted the SUb1 of SARS-CoV PLpro, the region near or at the mono-Ub and CISG15 binding sites. Many Ub-binding proteins including viral USP-like PLPs have been known to recognize Ub by its hydrophobic patch composed of several residues surrounding Ile44. The SARS-CoV PLpro-ChISG15 complex reveals that ISG15s have a conserved Trp123 that may be utilized by ISG15-binding proteins. From our mutagenesis experiments, Arg167 appears to be an important residue of SARS-CoV PLpro to engage with ISG15 at this region. While SARS-CoV

PLpro wild-type, prefers to hydrolyze ISG15 over Ub by 7-fold, we show that a single mutation can flip the selectivity of the protease. The Arg167Glu mutant prefers to hydrolyze Ub over ISG15 by 20-fold suggesting that a glutamate at this region of the Ub-binding site is preferred for DUB activity.

Further structural and mutagenesis studies that fully characterize the binding of SARS-CoV PLpro with the N-terminal domain of ISG15 at its second ubiquitin-binding site (SUb2) are necessary to identify additional sites that dictate ISG15 specificity. Modelling work by our collaborators and new finding from our lab, suggest that charged and polar residues of SARS-CoV PLpro ridge helix mediate interaction at the N-terminal lobe of ISG15 in contrast to the hydrophobic residues involved in recognizing distal Ub_2. These preliminary results are exciting avenues to explore to fully understand how SARS-CoV PLpro distinguishes between Ub-chains versus ISG15.

Overall, these SARS-CoV PLpro mutants may help develop PLpros with directed shifts in substrate specificities. PLpro mutants that disrupt specific protein-protein interactions may be useful tools to modulate the IFN-antagonism activity of SARS-CoV PLpro and for attenuating SARS-CoV infection.

CHAPTER 4. X-RAY STRUCTURE OF MERS PAPAIN-LIKE PROTEASE BOUND WITH ISG15 FACILITATES DESIGN OF PLPS WITH ATTENUATED OR ENHANCED SUBSTRATE SPECIFICITIES

The papain-like protease (PLpro) domain is an attractive target that is encoded in the RNA genome of coronaviruses, such as the infectious Middle East respiratory syndrome (MERS) coronavirus. PLpro acts as a protease by processing the large viral polyprotein to promote virus replication. In addition, PLpro acts as a deubiquitinating (DUB) and deISGylating (deISG) enzyme by removing ubiquitin (Ub) and interferon-stimulating gene 15 (ISG15) from cellular proteins. The latter activities are involved in cloaking the virus from the immune system. However, the precise role of DUB versus deISG activity in antagonizing the innate immune response is unclear, partly due to the difficulties in separating each activity. In this study, we determine the first structure of PLpro in complex with the full-length human ISG15 to 2.3 Å resolution. We utilize our newly determined PLpro-ISG15 and the previous PLpro-Ub complexes as molecular guides to design PLpro mutants that lack either or both DUB/deISG activities. We test 13 different PLpro mutants and the wild-type for protease, DUB, and deISG activity using fluorescence-based assays. Results show that we can selectively modulate DUB activity at positions 1649 and 1653 while mutation of Val1691 or His1652 of PLpro to a positive charged residue completely impairs both DUB/deISG activities. Overall, our study provides functional tools to delineate the importance of DUB versus deISG activity in virus-infected cells as well as potential candidates for attenuating the MERS virus for modified vaccine design efforts.

4.1 Introduction

Interferon-stimulated gene 15 (ISG15) is a 17-kDa antiviral ubiquitin-like (Ubl) protein rapidly induced by type I interferons (IFNs) in response to viral infection [123, 124]. ISG15 contains two Ubl-folds connected by a short interdomain linker or hinge region, and like other Ubl proteins, is conjugated to target proteins through an isopeptide linkage. The conjugation of ISG15 to target proteins is known as ISGylation, which is one of the critical mechanisms necessary for

its antiviral activity. ISG15ylation involves the consecutive catalytic function of three interferon-induced enzymes: the E1 activating enzyme Ube1L, the E2 conjugating enzyme UbCH8, followed by the E3 ligase human HERC5. Although the biochemical functions associated with ISGylation are still under investigation, a number of reports have now shown that ISG15 conjugation targets a broad range of viruses in an attempt to inhibit viral replication, including influenza A and B virus, Sindbis virus, HIV-1, herpes simplex-1, and murine herpesvirus [125-127]. One model for ISGylation involves a cotranslational process where HERC5 conjugates free ISG15 onto newly synthesized viral proteins at the ribosome in order to disrupt their viral function [124].

To help control ISG15 and its antiviral effector system, ISGylation can be reversed through the process of enzymatic deISGylation. The host deISGylase, ubiquitin-specific protease 18 (USP18), catalyzes the removal of ISG15 from target proteins in the cell, a process known as deISGylation. Recently, it was found that in humans, ISG15 deficiency actually increases resistance to viral infection. This is in contrast to what has been observed in mice, where ISG15 deficient mice are more susceptible to viral infection [128]. Reports suggest that in humans free extracellular ISG15 may be required to stabilize USP18 levels for temporal regulation of IFN-α/β signaling, a mechanism that does not exist in mice [128, 129]. Interestingly, the expression level of USP18 was found to be elevated in patients infected with chronic hepatitis C virus (HCV), suggesting the possibility that HCV can modulate host-derived machineries to promote viral infection [130].

Other single-stranded RNA viruses, such as coronaviruses and arteriviruses, have evolved evolved their own viral-encoded deISGylating (deISG) mechanisms to counteract the host ISG15/USP18 regulation system. For example, it was discovered that coronavirus (CoV) papain-like proteases (PLpro) encode deISG and deubiquitinating (DUB) catalytic activities in addition to their normal catalytic function, which is to cleave the viral polyprotein at three sites between

between nsp1-2, nsp2-3, and nsp3-4. The additional DUB and deISGylating activities of PLpros are thought to contribute to the suppression of the innate immune response by acting on IFN-β and NF-κB signaling pathways [77, 79, 85, 87]. However, the precise role for each of these activities in antagonizing the innate immune response remains unclear due in part to the difficulty in separating these activities and characterizing each function in cell culture and *in vivo*.

In this study, we focus on PLpro from MERS-CoV because this highly pathogenic *betacoronavirus* still persists in the human population in contrast to the shorter-lived SARS-CoV [10]. Since the initial outbreak in 2012, cases of MERS-CoV continue to be reported across 27 counties with the largest number of cases centralized in Saudi Arabia, United Arab Emirates, and the Republic of Korea [131]. To date, ~35% of the reported patients infected with MERS-CoV have succumbed to the fatal respiratory disease [132]. As there are no antivirals or vaccines currently available, there is a critical need to understand the structural and molecular basis of virus-host interactions to facilitate the design of treatments against MERS-CoV infection.

Since PLpro is multifunctional in nature, it is an attractive target for protein engineering studies. For example, selectively removing PLpro DUB/deISG activities while maintaining its polyprotein cleavage activity is one strategy that could potentially be used for modified vaccine design through the development of an attenuated coronavirus. Designing different PLpro mutants that are DUB, deISG, or both DUB/deISG deficient would also be advantageous for functional studies in virus-infected cells to delineate how each activity contributes to virus replication and pathogenesis.

Our group first showed that it is possible to selectively modulate the SARS-CoV PLpro DUB and/or deISG activities without affecting its ability to cleave the *trans* nsp2-3 site by using structure-guided engineering [77]. Subsequently, our approach was translated to MERS-CoV PLpro using the structure of MERS-CoV PLpro in complex with Ub as a molecular guide [83]. In

this study, mutations were introduced into the PLpro domain that disrupted DUB activity in biochemical assays and in cell culture, but its protease activity remained near wild-type levels. However, since the PLpro mutants engineered in that study were only tested for DUB activity, it is still unclear if the deISG activity of PLpro was also affected and if both DUB as well as deISG of PLpro contributed to the activation observed in the IFN antagonism assays [83].

Another gap in our current knowledge exists in understanding the specificity of MERS-CoV PLpro for the full-length ISG15 protein. The structure of PLpro with a full-length ISG15 protein has yet to be determined. An X-ray structure of MERS-CoV PLpro bound to the C-terminal domain of human ISG15 (CISG15) was recently reported but it lacks the N-terminal domain which is known to be important in binding SARS PLpro [133]. The X-ray structure of mouse USP18 in complex with full-length mouse ISG15 has been determined but USP18 has low sequence identity to PLpros and it strictly recognizes ISG15 and has no recognition of Ub [74].

Here, we determine the first structure of a viral USP bound to the full-length human ISG15 protein to better differentiate the structural basis of MERS-CoV PLpro activity towards ISG15 versus Ub. We provide a detailed comparison between the ISG15 and Ub-bound complex as well as mutagenesis studies of key residues that may shed light on potentially eliminating both DUB/deISG activities of PLpro as well as selectively alternating PLpro substrate preference.

4.2 Materials and Methods

4.2.1 Expression and purification of MERS-CoV PLpro-ΔUbl2

The MERS-CoV PLpro catalytic domain of nsp3 (residues 1544-1801 of the MERS viral polyprotein 1ab) which is missing the Ubl domain was expressed from plasmid pEVL8-his8-TEVp-MERS-CoV PLpro-ΔUbl2 in *E. coli* BL21(DE3) by autoinduction. Cell were grown for 24 hours at 25 ºC at a one liter scale after inoculation with a 10 mL starter culture as previously described [103]. Cells from four one-liter cultures were pelleted by centrifugation (6,440 x g, 4°C,

30 min), and the resulting pellets from two one-liter cultures were distributed into two different 50 mL falcon tubes. Each pellet weighed approximately 16 g and 10 g, respectively, and were stored in the -80°C until purification where each pellet was purified separately. On the day of purification, each pellet was thawed from the -80°C and resuspended in 5 mL lysis buffer / 1 g of cells [103]. After resuspension, the cells were sonicated on ice with the same protocol as before but a shorter lyse time of 10 min [103].

The lysed cells were clarified by centrifugation (26,200 x g, 4°C, 45 min), and the supernatant was passed over a 5 ml HisTrap FF column (GE Healthcare) that was pre-charged with Ni^{2+} and equilibrated with 5 column volumes (CV) of 3% buffer B at 3 mL/min [103]. Weakly bound proteins were eluted from the column using 12 CV of 3% buffer B. Bound proteins, including his8-TEVp-MERS-CoV PLpro-ΔUbl2, were eluted from the column over a period of 33 minutes using a linear gradient of 20 CV starting at 3% buffer B and progressing to 100% buffer B while collecting 5 mL fractions. The fractions that showed absorbance at 280 nm on the chromatograms were analyzed by SDS-PAGE and for PLpro enzymatic activity. Those fractions judged to be most pure were combined and the final pools had volumes of 30 ml and 20 ml, respectively.

The N-terminal octa-his-tag was removed by dialyzing each pool against 2 L of buffer A supplemented with tobacco etch virus (TEV) protease at a 1 mg TEV: 6 mg PLpro ratio for 96 hours at 4°C. During dialysis, precipitation was observed, which predominantly contained contaminating proteins as assessed by SDS-PAGE. The dialyzed pools from each purification were passed separately and in tandem over a 5 ml HisTrap FF pre-charged Ni^{2+} column. The eluates containing the now untagged PLpro-ΔUbl2 in the flow-through were collected and the column was then washed using 5 CV of Buffer A. The flow-through and wash samples were combined and concentrated to ~2 mL at 6.3 mg/mL using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa

MW cutoff). Concentrated MERS PLpro-ΔUbl2 was then buffer exchanged into reaction buffer containing 10 mM Tris, pH 7.5, 5% glycerol, 10 mM BME using a PD-10 desalting column (GE Healthcare). The untagged and fresh (i.e. unfrozen and unstored) PLpro-ΔUbl2 was used immediately after purification for generating a fresh PLpro-ISG15 complex.

4.2.2 Expression and purification of human ISG15-PA

ISG15 propargylamine (ISG15-PA) was synthesized using the chitin intein-fusion method previously described by Wilkinson et al. [134]. Plasmid pTYB2-ISG15$_{1–156}$Cys78Ser was expressed in *E. Coli* BL21(DE3) cells by autoinduction for 24 hours at 25 ºC. A one liter LB culture were inoculated with 10 mL of transformed *E. Coli* BL21(DE3) cells. After centrifuging (3011 x g, 20 min, 4 ºC) the culture and discarding the supernatant, the harvested cells (~7 g) were frozen at –80 ºC. Cells were thawed and then resuspended in 50 mL chitin-column buffer (50 mM MES pH 6.5, 350 mM sodium acetate), supplemented with 50 μM PMSF, before lysed by sonication using a Branson Digital Sonifer (65% amplitude; 15 minutes, 5.5 s pulses, 5.5 s wait intervals). Cell debris was removed by centrifugation (28,960 x g, 30 minutes, 4 ºC) and the supernatant (clarified lysate) was saved.

Chitin resin (20 mL in 20% ethanol, New England Biolabs) was transferred to a 150 mL fritted funnel and was equilibrated with 200 mL cold chitin-column buffer using vacuum filtration. The equilibrated chitin resin was transferred to a 100 mL beaker, combined with the clarified lysate (50 mL), and allowed to incubate with gentle stirring for 30 minutes at 4 ºC. The chitin resin was then transferred back to the fritted funnel, and unbound proteins were washed using vacuum filtration with ten volumes of 50 mL column buffer. The resin was resuspended in 50 mL cold reaction buffer (50 mM MES pH 6.0, 350 mM sodium acetate, 122 mM sodium 2-mercaptoethanesulfonate (MESNa)), transferred to a beaker, and allowed to incubate at 4 ºC overnight with gentle stirring.

The next day, the ISG15-MESNa that formed after the reaction was separated from the resin using vacuum filtration. The resin was then resuspended and washed with two 50 mL portions of column buffer to recover any residual ISG15-MESNa. The elution and washes were combined and concentrated to approximately 20 mL using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MW cutoff). Approximately 0.3 g propargylamine HCl (Sigma-Aldrich) was added to the concentrated ISG15-MESNa at a 250 mM final concentration and 2.5 mL of 2 M NaOH was added dropwise to adjust the solution to pH 10. To generate ISG15-PA, the reaction was allowed to precede overnight at room temperature and the final reaction products were centrifuged (28,960 x g, 10 min, 4 ºC) to remove any precipitation. The resulting supernatant was concentrated to approximately 2 mL and was buffer exchanged into reaction buffer using a PD-10 desalting column (GE Healthcare). Freshly generated ISG15-PA was used immediately for generating fresh PLpro-ISG15 complex. Ub-PA was also generated with a similar procedure as described above for probe reactivity assays. The only difference was that Ub-intein-CBD was expressed using *E. Coli* BL21(DE3) + RIPL cells, and protein was concentrated using 3 kDa MW cut-off Amicon device.

4.2.3 Generation and Purification of MERS-CoV PLpro-ΔUbl2 in complex with human ISG15-PA

An incubation reaction containing 12.7 mg of PLpro-ΔUbl2 and 75 mg of ISG15-PA (1:10 molar ratio) was allowed to react overnight at 4°C in reaction buffer. Several attempts were used to separate the complex from free PLpro by passing the reaction over an 8 mL MonoQ 10/100 GL column (GE Healthcare), but regrettably free PLpro was not separated from the PLpro-ISG15 complex. Our complexed PLpro was assayed for activity with 250 nM Ub-AMC. The PLpro-ISG15 complex still had some remaining activity (60 μM/min/mg) compared to the untagged PLpro sample (212 μM/min/mg). Alternatively, an additional 21 mg of ISG15-PA was added to the reaction mixture. The goal was to achieve 100% reacted PLpro-ISG15 complex. The reaction

mixture was kept at 4°C for 72 hours in MonoQ buffer A (10 mM Tris, pH 9.0, 5% glycerol, 10 mM BME) to ensure the complex was negatively charged for the final purification step. The reaction was passed over a final MonoQ column equilibrated with MonoQ buffer A, and unbound protein was washed with 3 CV of MonoQ buffer A. The PLpro-ISG15 complex was eluted with a linear gradient from 0 to 20% MonoQ buffer B (10 mM Tris, pH 9.0, 500 mM NaCl, 5% glycerol, 10 mM BME) for the duration of 64 minutes (20 CV), collecting 3 mL fractions at a flow rate of 2.5 mL/min. The final pool, shown in Figure 4.3, was buffer exchanged into storage buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 5% glycerol, 10 mM DTT) using a PD-10 desalting column and concentrated to 15.3 mg/mL using an Amicon Ultra-15 Centrifugal Filter Unit (30 kDa MW cutoff).

4.2.4 Crystallization and structure determination of PLpro-ISG15 complex

Freshly prepared PLpro-ΔUbl2-hISG15 complex was used to screen for initial crystallization conditions. A series of sparse-matrix crystallization screening solutions from Anatrace (Midwest Center for Structural Genomics (MCSG) screens-1-4) were used and sitting drops were formed using a Mosquito®Crystal liquid handling robot (TTP Labtech). Three protein concentrations (5 mg/mL, 10 mg/mL, 15.3 mg/mL) were screened in 96-3 well sitting drop vapor diffusion plates (Greiner CrystalQuick crystallization plate) at 1:1 ratio (100 nL of purified PLpro-ΔUbl2-hISG15 complex to 100 nL of reservoir solution). The 200 nL protein drops were allowed to equilibrate against 50 µL reservoir solution at 20°C in a Rigaku Minstrel® HT plate hotel and photos of each drop were recorded daily.

An initial crystal hit from the MCSG-3 screen, containing 0.2 M potassium citrate tribasic, pH 8.3, 20% (w/v) PEG3350, was observed after 2 days. Initial crystals grew as thin plates with a hexagonal morphology (150 µm). This condition was optimized further using a 24-well sitting drop format with drops containing 1 µl of purified PLpro-ΔUbl2-hISG15 complex at 15.3 mg/ml and 1 µl reservoir. Optimization reagent, 2 M potassium citrate tribasic pH 8.3, was purchased

from Anatrace, Inc. while a 50% PEG3350 solution was made from powder. Large crystals (>500 µm) formed within 2 days in only a small subset of conditions, but many appeared fused or twinned. Large isolated crystals did form in higher citrate concentrations, which were harvested. A crystal

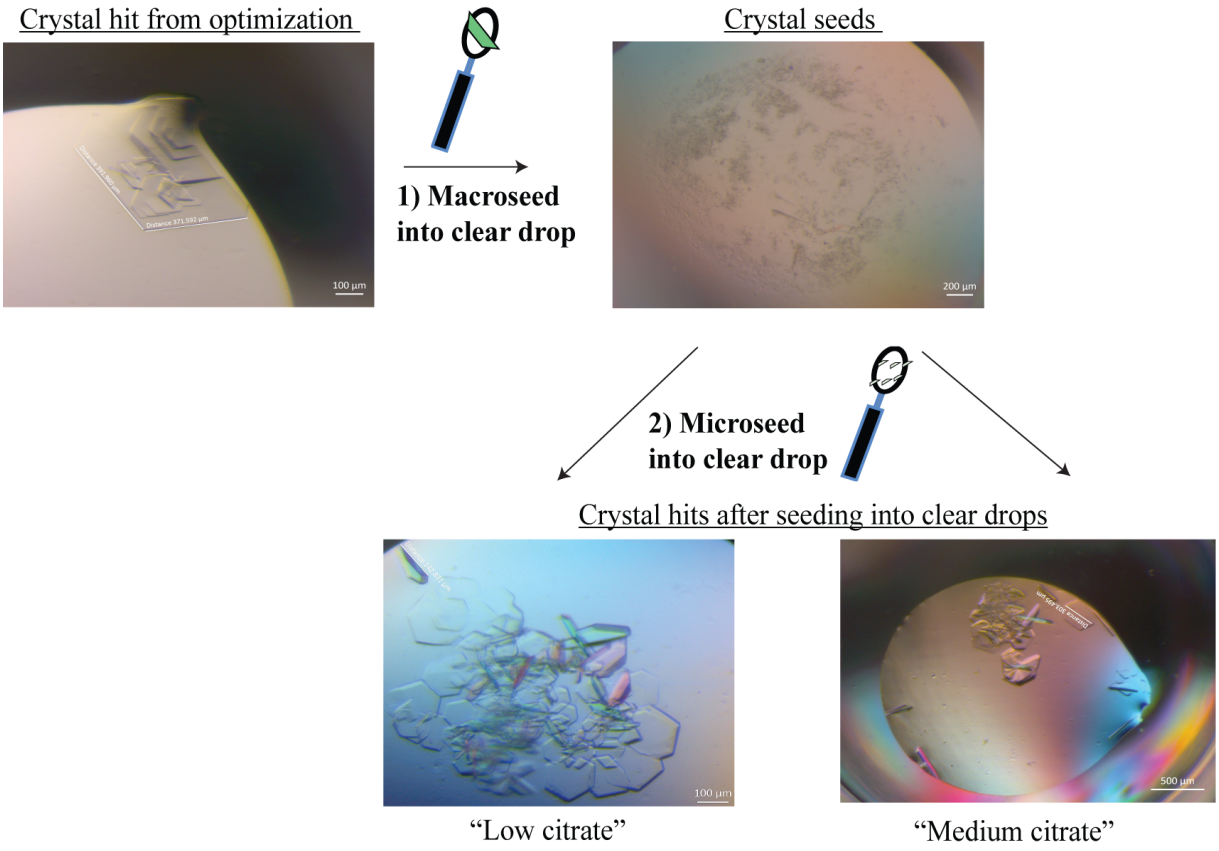


Figure 4.1 Procedure used to generate crystals of MERS PLpro-ISG15 in clear drops. Top left – Image of the crystal hit used for macroseeding. A single crystal slice was looped and transferred to a clear drop. Top Right – Image of the macroseeded clear drop. Small microcrystals formed almost rapidly after macroseeding. The seeds were disturbed with a nylon loop and briefly dipped into subsequent clear drops. Bottom images - Crystal morphology of crystals produced from this method is shown under polarized light. Crystallographic data was collected on crystals from "low citrate" and "medium citrate" wells, and the former dataset was submitted in the PDB.

in the condition containing 0.35 M potassium citrate tribasic, pH 8.3, 18% (w/v) PEG3350 was screened and data was collected in house using the Rigaku R-AXIS IV++ detector (180 mm detector distance) at 100 K with from 0 – 150° with a 12 min exposure time on the home source "Moe" at Purdue University. Another crystal termed "high citrate" in reservoir containing 0.35 M

potassium citrate tribasic, pH 8.3, 22% (w/v) PEG3350 was harvested and X-ray data was collected at the Advance Photon source (APS) at Argonne National Laboratory (ANL).

To obtain more crystals for data collection, crystals were further optimized by macroseeding crystals into the clear drops in the same tray (Figure 4.1). A small crystal was cut from the layered crystal sheet in reservoir containing 0.35 M potassium citrate tribasic, pH 8.3, 22% (w/v) PEG3350 and transferred to a clear drop, which contained the same reservoir solution. Small microcrystals or crystal seeds rapidly formed in the macroseeded drops and were subsequently used for transferring into another clear drop. Briefly, the microcrystals were looped with a 0.05-0.1 mm loop and quickly submerged into different clear drops. After transferring these microcrystals, single isolated hexagonal crystals (50 μm) formed within a few minutes and continued to grow to ~100-300 μm in size after 24-48 hour incubation at 20°C. High quality crystals were grown using this technique (Figure 4.1). Crystals were grown in the same initial condition containing 0.2 M potassium citrate tribasic, pH 8.3, 20% (w/v) PEG3350, which was termed “low citrate” and also at a slightly higher citrate concentration (0.25 M potassium citrate tribasic, pH 8.3, 20% (w/v) PEG3350), which was termed “medium citrate”. Crystals were harvested using nylon loops, briefly soaked in cryoprotectant solution containing 20% glycerol, before flash-cooled into a vial submerged in liquid nitrogen. Vials were harbored in SPINE pucks during cryo-freezing and transported to the APS at ANL.

X-ray data were collected on crystals of the PLpro-ISG15 complex at 100 K using a Dectris Eiger 9M detector (a2°/s spindle rate; frame rate of 4; snap delta 0.25°) on beamline 21-ID-D at the Life Sciences-Collaborative Access Team (LS-CAT). Data were collected on three different crystals from the three different conditions described above and each was indexed, processed, and scaled separately using HKL-3000 [135]. HKL-2000 was used for the in-house dataset.

To determine the initial phases for the PLpro-ISG15 complex, the PLpro domain from the bound PLpro-Ub complex (PDB accession code: 4RF1) was used as a search model and the programs Phaser [136] and Phenix were used to identify a molecular replacement solution. Free human ISG15 [123] with the PDB accession code of 1Z2M was manually positioned into the corresponding residual electron density using the program Coot [109]. The model was next refined using ridged body refinement and the program Phenix.Refine [110]. Additional rounds of manual modeling building followed by coordinate and B-factor refinement were performed using the programs Coot and Phenix.Refine. Table 4.2 shows the data collection and refinement statistics for the in-house dataset. The initial phases for the datasets collected at APS were also solved by using molecular replacement with programs Phaser [136] and Phenix, but the refined model built from dataset collected in house was used as a search model. Again, additional rounds of manual modeling building followed by coordinate and B-factor refinement were performed using the programs Coot and Phenix.Refine. Solvent molecules were added towards the end of model building and refinement which was ultimately complete when the R_{free} and R_{work} values plateaued at their minimums.

The atomic coordinates and structure factors for the MERS PLpro-ISG15 complex that crystalized at the lowest citrate concentration are deposited in the PDB with the accession code of 6BI8. Table 4.3 shows the final X-ray data collection and refinement statistics from this dataset, and also the other two datasets collected at APS. Figures were generated with Pymol (The Pymol Molecular Graphics System, 1.8.0 Schrödinger, LLC).

4.2.5 Expression and purification of MERS-CoV PLpro mutants

A series of 13 site-directed mutants were generated using the QuickChange™ site-directed mutagenesis system [137] from Stratagene (La Jolla, CA). The wild-type pEVL8-his8-TEVp-MERS-CoV PLpro-ΔUbl2 plasmid was used as a template and primers were designed and

synthesized using Integrated DNA Technologies (IDT). Primer used for site-directed mutagenesis are shown in Table 4.1. The mutagenesis protocol, outlined in Figure 4.2, was designed based on the recommended cycling instructions for Phusion polymerase (NEB), which was performed with a Bio-Rad MJ Mini Personal Thermal Cycler. Upon successful amplification of the plasmid confirmed by 1% agarose gel, the PCR product was subject to Dpn1 digest and PCR clean-up before the linear DNA was transformed into XL1 Blue electrocompetent cells (Agilent) to circularize and generate more plasmid DNA.

Correct mutations in the PLpro gene were confirmed by Sanger-based sequencing, which was performed by the Purdue Genomics Core Sequencing Facility. Unsuccessful mutagenesis attempts, which contained long primer insertions, were troubleshooted by performing PCR at varying primer concentrations from 0.5 – 0.03 μM, keeping the template concentration at 10 ng. Plasmids that were generated using lower primer concentrations were less likely to have insertions present in the PLpro genes, and this eventually allowed for all mutants to be successfully generated.

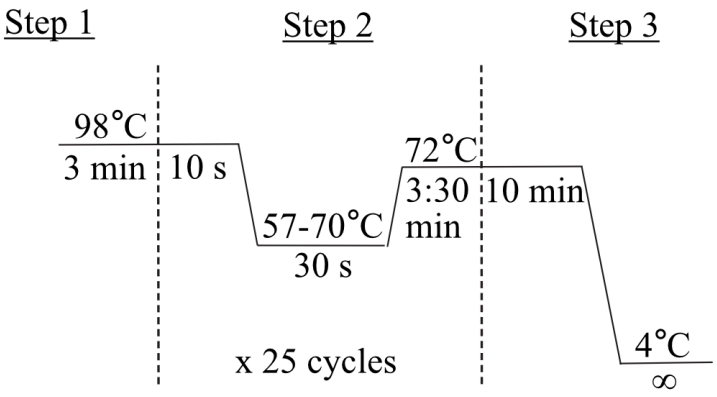


Figure 4.2 PCR Cycling instructions for site directed mutagenesis of MERS PLpro mutants.

Expression and purification of MERS-CoV PLpro mutant enzymes were performed by a modified procedure described above for the wild-type enzyme. For each of the mutant enzymes Arg1649Ala, His1652Ala, His1652Phe, Val1691Arg, the pellets (12-15 g) from two one-liter

cultures were harvested and a two-day purification procedure was done using a Ni^{2+}-charged, 5 mL HisTrap FF column. The pooled fractions obtained from the Ni Hi-trap column were subjected to TEV protease digest at a 1:10 ratio of TEV to PLpro while dialyzing against 1 L of buffer A for 16 hours at 4°C in order to remove the N-terminal his-tag. The next day, the dialyzed sample was passed over the 5 ml Ni^{2+}-charged HisTrap FF column, and the flow-through as well as the wash were combined and concentrated the same way as the wild-type. Instead of using a desalting column, the final pool was buffer exchanged into storage buffer using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MW cutoff) and then concentrated to ~6-14 mg/mL. The final purified mutants were flash-frozen with liquid nitrogen at 100 K and stored in the -80°C until their kinetic characterization.

Table 4.1 Primers used for site-directed MERS PLpro Ub/ISG15 binding mutants

Mutation	Forward Primer (5' – 3')	Reverse Primer (5' – 3')
K1550A	CACTGGCAGAGCTGTACGGTCCGGTTGATC	CAGCTCTGCCAGTGCCTTAGTCTCGTC
K1550F	CACTGTTTGAGCTGTACGGTCCGGTTGATC	CAGCTCAAACAGTGCCTTAGTCTCGTCATTGG
R1649A	GAGCGCTCTGCTGCACACCGTACTGGCAAAAG	GTACGGTGTGCAGCAGAGCGCTCGCATCATCAG
R1649E	GAGCGAACTGCTGCACACCGTACTG	GTGTGCAGCAGTTCGCTCGCATCATCAG
H1652A	GTCTGCTGGCAACCGTACTGGCAAAAGC	AGTACGGTTGCCAGCAGACGGCTCGCATCATC
H1652R	GTCTGCTGCGTACCGTACTGGCAAAAGC	AGTACGGTACGCAGCAGACGGCTCGCATCATC
H1652F	GTCTGCTGTTCACCGTACTGGCAAAAGC	AGTACGGTGAACAGCAGACGGCTCGCATCATC
T1653R	CTGCACCGTGTACTGGCAAAAGCGGAACTG	CTTTTGCCAGTACACGGTGCAGCAGACGGCTC
T1653F	CTGCACTTCGTACTGGCAAAAGCGGAAC	CAGTACGAAGTGCAGCAGACGGCTCGCATC
V1691R	GTTACCGTGGCGTGCAGACCGTAGAAGACC	GGTCTGCACGCCACGGTAACAACAAGC
V1691K	GTTACAAAGGCGTGCAGACCGTAGAAGACC	GGTCTGCACGCCTTTGTAACAACAAGC
V1691F	GTTACTTTGGCGTGCAGACCGTAGAAGACC	GGTCTGCACGCCAAAGTAACAACAAGC
V1691S	GTTACAGTGGCGTGCAGACCGTAGAAGACC	GGTCTGCACGCCACTGTAACAACAAGC

The remaining mutants were expressed by autoinduction at a 500-mL scale in super broth media. Mutants were purified in parallel 2-4 mutants at a time. Harvested pellets weighing ~3-5 g were resuspended in 50 mL lysis buffer, and the cells were lysed at a lower intensity and duration (5-minute protocol at 65% amplitude with 6.6 s pulses and 9.9 s delays) before clarified by centrifugation (26,200 x g, 4 °C, 25 min). The clarified lysates were passed over a Bio-Rad Econo-Pac® gravity column packed with a 3 mL HisPur™ Ni-NTA resin (Thermo Fisher Scientific), which was equilibrated with buffer A. Unbound proteins were washed from the column with a batch wash using 25 mL of buffer A and then with 25 mL of a lower percentage wash ranging from 12-30% buffer B. Finally, the mutant PLpro was eluted with 25 mL of 50% of buffer B. All washes and elutions were tested for absorbance at 280 nm with a BioTek Take3™ Multi-Volume Plate and were pooled based on the presence of his-PLpro band at 31 kDa on the SDS-PAGE gel. As done with the other PLpro mutants, TEV protease was added to the pool samples before the pool was transferred to dialysis tubing (10 kDa MW cutoff), and dialyzed against 1 L buffer A for 16 hours at 4°C. To separate TEV protease from the untagged PLpro mutant, the dialyzed sample was again passed over the gravity column equilibrated with buffer A. The flow-through containing untagged PLpro was collected as well as a 15-20 mL wash of Buffer A. The flow-through and the wash were combined and buffer exchanged into storage buffer using the Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MW cutoff) before protein was concentrated to between 1-9 mg/mL. As done with the other PLpro mutants, samples were flash-frozen with liquid nitrogen and stored at -80 °C.

4.2.6 Steady-state kinetic characterization of MERS PLpro wild-type and mutant enzymes

An initial screen to assess the catalytic activity of all 13 mutant PLpros and wild-type enzyme was performed at a single substrate concentration on three commercially available fluorogenic substrates; 50 μM Z-RLRGG-AMC (Bachem), 1 μM Ub-AMC (Boston

Biochem/R&D Systems) and 1 μM ISG15-AMC (LifeSensors, Inc.). Each of these substrates contains the 7-amino-4-methylcoumarin (AMC) fluorescence group at the C-terminus. Fluorescence due to the release of the AMC group from the substrates was measured as described previously using a BioTEK Synergy H1 multimode microplate reader at 25°C [94, 103]. The reaction conditions i.e. plate-type and volume for each assay has been previously described in [94]. Peptide hydrolysis was initiated with 0.75 μM of PLpro (final concentration) in reaction buffer (50 mM HEPES, 0.1 mg/mL BSA, 5 mM DTT). For the Ub-AMC and ISG15-AMC assays, reactions were initiated with 2.5 nM PLpro (final concentration) with the exception of His1652Arg, Val1691Arg, and Val1691Lys, which were initiated with a higher enzyme concentration of 0.83 μM due to lower turnover rates in cleaving both substrates which was observed during purification. The k_{cat}, i.e turnover number, for each mutant and the wild-type enzyme were measured in triplicate for each fluorogenic substrate and at each substrate concentration. Rates were normalized to the wild-type as a relative percentage (%). Data from the initial kinetic screen are shown in Figure 4.10.A.

To more fully characterize each mutant that had observable differences in catalytic rates, the steady-state kinetic parameters of the MERS-CoV PLpro wild-type and six mutants were determined for ISG15-AMC and Ub-AMC. Substrate concentration was varied from 0.19 μM to 6 μM for ISG15-AMC and from 0.38 μM to 12 μM for Ub-AMC. The reactions were initiated with 1.7 nM PLpro in both assays. The initial reaction rates were measured in duplicate and the averaged rates were plotted as a function of substrate concentration. For mutant enzymes showing observable saturation at higher substrate concentrations, these data were then fit to the Michaelis-Menten equation using the enzyme kinetics module in SigmaPlot (version 12; Systat Software, Inc.). The kinetic parameters (k_{cat}, K_m) resulting from the fit of the data as well as the calculated value of k_{cat}/K_m for each enzyme were tabulated. The Val1691Ser enzyme could not be saturated

with either of the Ub-AMC and ISG15-AMC substrates up to the highest concentrations tested. Saturation was also not observed with the Thr1653Arg enzyme using Ub-AMC as the substrate. In these cases, the initial reaction rates were plotted as a function of substrate concentration, and the data were fit to a line with linear regression in GraphPAD Prism6 and the apparent k_{cat}/K_m was determined (Table 4.5).

4.2.7 Ub-PA and ISG15-PA probe reactivity assays with wild-type and mutant MERS PLpros

Probe reactivity assays were set-up at a 10-µL scale with pure MERS PLpro enzymes and PA probes. The reactivity of MERS PLpro wild-type was assessed with both PA probes at varying molar ratios. Ub-PA and human ISG15-PA were diluted in storage buffer containing fresh 10 mM DTT at final concentrations ranging from 5-25 µM and 25-100 µM, respectively. Each reaction was initiated with 5 µM MERS PLpro wild-type and allowed to react for 1 hour at room temperature before quenching with 5X loading buffer (250 mM Tris-HCl pH 6.8, 50% glycerol, 10% SDS, 0.02% bromophenyl blue, and fresh 400 mM DTT). The reactivity of MERS PLpro wild-type and mutants were evaluated at single molar ratios of 5 µM PLpro and either 25 µM Ub-PA (1:5 molar ratio) and 100 µM ISG15-PA (1:20 molar ratio), which were quenched at different time points (2, 5, 30 min) with loading buffer. The entire reaction was loaded onto 4-20% Mini-PROTEAN® TGX™ precast SDS-PAGE gels from BioRad, and proteins were visualized with Coomassie-Brilliant Blue staining.

4.3 Results

4.3.1 X-ray Structure Determination of the MERS PLpro-ISG15 complex

In Chapter 2, we determined that the core MERS-CoV PLpro-ΔUbl2 enzyme, i.e. without its adjacent N-terminal ubiquitin-like 2 (Ubl2) domain, is thermally stable, retains full kinetic activity and substrate specificity profile and has an increased ability to readily form crystals that

diffract to high resolution compared to the enzyme with the Ubl2 domain [103]. We therefore used the MERS-CoV PLpro-ΔUbl2 construct for the crystallization of the MERS-CoV PLpro-ΔUbl2-ISG15 complex, which was generated by reacting MERS-CoV PLpro-ΔUbl2 with the full-length, human ISG15 protein that was modified with a reactive, propargylamide (PA) group at its C-terminus (Figure 4.3). The PA group on ISG15 reacts with the catalytic cysteine in the active site that can be readily purified. From this point on, we will refer to the MERS-CoV PLpro-ΔUbl2 construct as simply "MERS-CoV PLpro" for simplicity.

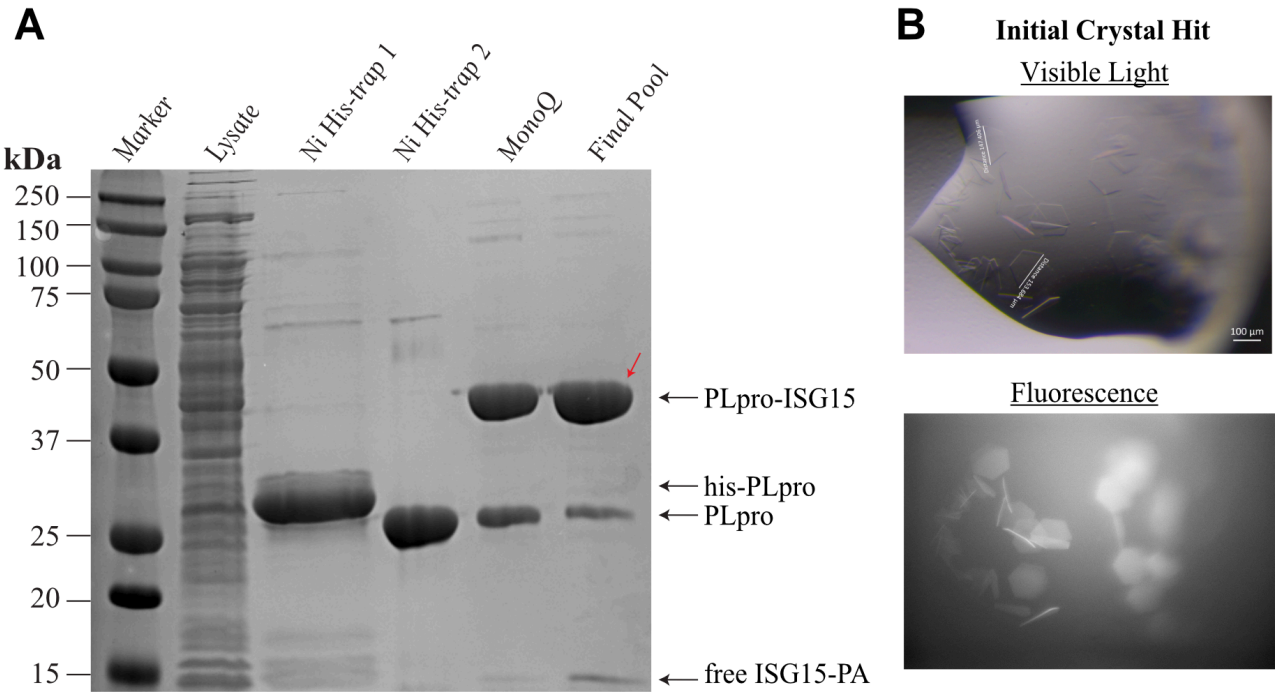


Figure 4.3 Purification summary and initial crystal of the MERS-CoV PLpro-ISG15 complex (A) 12.5% SDS-PAGE gel analysis of purified MERS-CoV PLpro-ΔUbl2 (lanes 2-4), and purified PLpro-ISG15 complex. The covalent PLpro-ISG15 complex was generated with ISG15-PA and purified via consecutive MonoQ anion exchange columns (lanes 5-6). Final pool (loaded at 10 µg) is predominantly the covalent PLpro-ISG15 complex with minor traces of free PLpro and ISG15-PA in the sample. (B) Top – Image of the hexagonal crystals observed under visible light. Bottom – Fluorescence after exposure to UV light. Images were taken using Rock Imager from Formulatrix.

By SDS-PAGE analysis, the final pool contained predominantly MERS PLpro-ISG15 complex with minimal free untagged PLpro and free ISG15-PA (Figure 4.3.A). This sample

produced thin hexagonal crystals, which were fluorescent upon exposure to UV light (Figure 4.3.B). After optimizing from this initial condition from a 200 nL drop to a 2 µL drop size, larger crystals were obtained, some that were nearly >500 µm in size (Figure 4.4). An initial dataset was collected on a crystal slice in house, which diffracted to 3.28 Å resolution. Representative diffraction patterns are shown in Figure 4.4 and the data collection statistics are reported in Table 4.2. The refined structure (statistics also in Table 4.2) served as a nice search model for molecular replacement for data sets collected at APS.

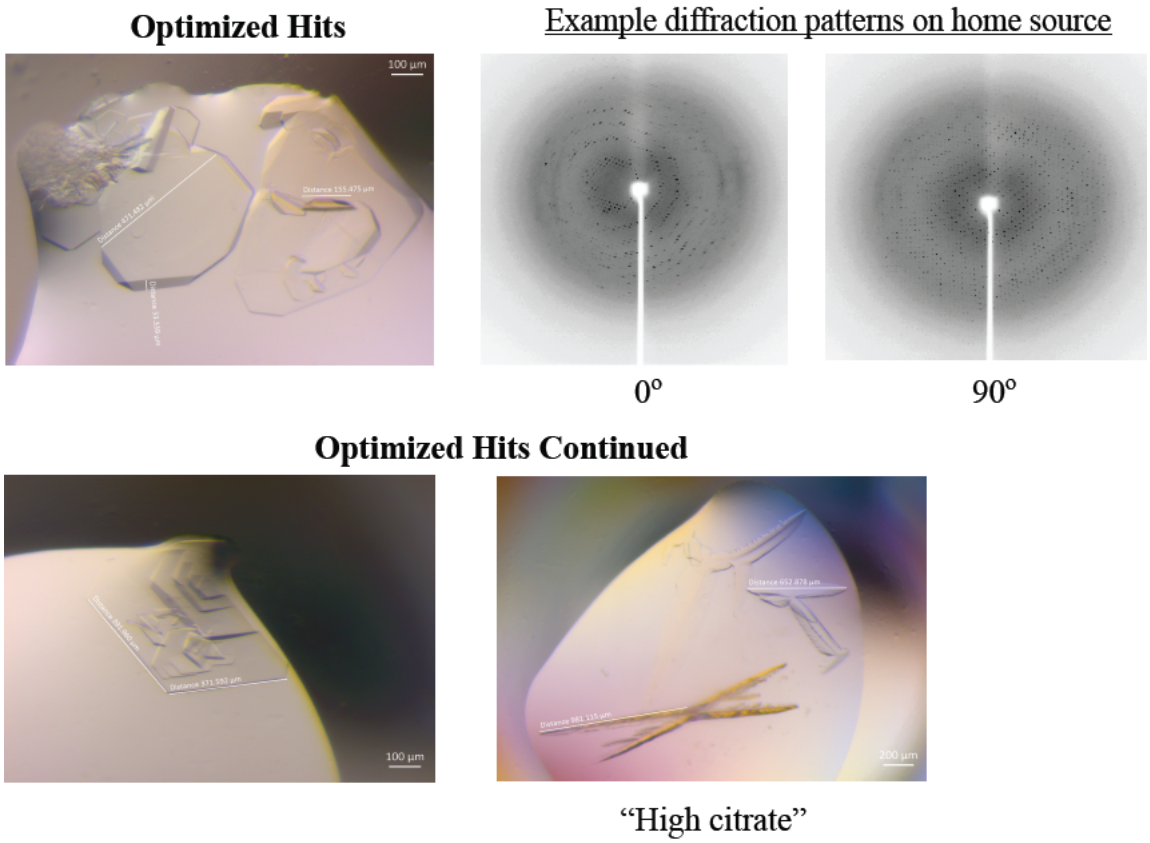


Figure 4.4 MERS PLpro-ISG15 complex crystals and representative diffraction patterns

Top left – Image of well where crystals were harvested and screened on the home source. Top right – Diffraction pattern images collected on 200-300 µm crystal at two different angles. A full data set was collected, and statistics are shown in Table 4.2. Bottom – Other images of crystals formed in same tray with unique crystal morphologies. A full dataset was collected at APS on a crystal from the "high citrate" condition.

Table 4.2 Data-collection and structure refinement statistics for MERS PLpro-ISG15 crystal collected in house

Data-collection parameters	
Beamline	In house ("Moe")
Wavelength (Å)	1.54
Space group	*R3*
Unit cell dimensions:	
a, b, c (Å)	*147.2, 147.2, 128.077*
α, β, γ (°)	*90, 90, 120*
Resolution (Å)	25-3.28 (3.40-3.28)[a]
Number of reflections observed	358332
Number of unique reflections	15762
R_{merge} (%)[b]	6.1 (23.2)
R_{pim} (%)[c]	4.1 (17.2)
$CC_{1/2}$ (%) *in highest shell*	90.5
CC* (%) *in highest shell*	97.5
I/σI	10.5 (2.5)
% Completeness	93.4 (93.9)
Redundancy	2.6 (2.3)
Refinement	
Resolution range (Å)	25.0–3.28 (3.36–3.28)
No. of reflections in working set	20193
No. of reflections in test set	2064
R_{work} (%)[d]	21.8 (28.3)
R_{free} (%)[e]	25.3 (32.4)
Wilson B factor (Å^2)	62.66
Average B factor (Å^2)	69.7
RMSD from ideal geometry	
Bond length (Å)	0.009
Bond angle (deg)	0.889
Ramachandran plot	
Most favored (%)	87.6
Allowed (%)	8.44
Disallowed (%)	3.97

[a]Values in parentheses are for the last (highest resolution) shell.

[b]$R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of a given reflection, and $\langle I(hkl)\rangle$ is the mean intensity of symmetry-related reflections.

[c]$R_{pim} = \sum_{hkl}\sqrt{(\frac{1}{n}-1)}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i(hkl)$, where n is the multiplicity for multiplicity-weighted R_{merge}.

[d]$R_{work} = \sum_{hkl}\big||F_{obs}| - |F_{calc}|\big| \,/\, \sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[e]R_{free} was calculated using 10% of the data set chosen at random that were excluded from the refinement.

Three complete X-ray datasets were collected on crystals of the MERS-CoV PLpro-ISG15 complex at APS that were grown at three different citrate concentrations (0.2, 0.25 and 0.35 M). Images of wells that contained these crystals are shown in Figure 4.1 and Figure 4.4 and are labeled low, medium, and high citrate, respectively. A summary of the X-ray data collection and refinement statistics from all three structures are provided in Table 4.3. Unexpectedly, analysis of the electron density surrounding the Zn-finger motif for the PLpro-ISG15 complex crystals grown from 0.35 M citrate revealed no observable electron density for the Zn^{2+} atom in the Zn-finger motif even though these crystals diffracted to the same resolution and grew in the same space group as the crystals grown at the two lower citrate concentrations (Figure 4.5). At citrate concentrations of 0.2 M and 0.25 M, the Zn^{2+} atom in the Zn-finger motif is clearly visible in electron density suggesting that the higher concentration of citrate chelated the Zn^{2+} and thereby depleted it from the Zn-finger. Therefore, the X-ray structure of the MERS-CoV PLpro-ISG15 complex grown from crystals at 0.2 M citrate was determined to a resolution of 2.3 Å with a final R_{work} = 17.1% and R_{free} = 21.5% and was deposited into the PDB (accession code 6BI8). The enzyme crystallized in space group R3 with two identical copies of the PLpro-ISG15 complex in the asymmetric unit. The PLpro chains from each dimer were virtually identical with a RMSD of 0.08 Å over Cα atoms from 255 residues while human ISG15 exhibited slightly more conformational diversity with a RMSD of 0.8 Å over 151 Cα positions.

Table 4.3 Data-collection and structure refinement statistics for MERS PLpro-ISG15

PDB code	**6BI8**		
Potassium citrate concentration	Low(0.20 M)	Medium (0.25 M)	High (0.35 M)
Data-collection parameters			
Beamline	21-ID-D	21-ID-D	21-ID-D
Wavelength (Å)	0.987	0.987	0.987
Space group	*R3*	*R3*	*R3*
Unit cell dimensions:			
a, *b*, *c* (Å)	*148.035, 148.035, 134.189*	*148.254, 148.254, 134.237*	*147.903, 147.903, 133.758*
α, *β*, *γ* (°)	*90, 90, 120*	*90, 90, 120*	*90, 90, 120*
Resolution (Å)	100-2.30 (2.34-2.30)[a]	100-2.30 (2.34-2.30)[a]	100-2.30 (2.34-2.30)[a]
Number of reflections observed	2377975	4887313	3624708
Number of unique reflections	49135	48960	49239
R_{merge} (%)[b]	10.0 (56.0)	13.5 (87.9)	11.1 (63.1)
R_{pim} (%)[c]	5.8 (33.7)	7.4 (48.9)	6.1 (33.8)
$CC_{1/2}$ (%) in highest shell	73.2	70.3	79.3
CC (%) in highest shell*	91.9	90.9	94.1
I/σI	20.2 (2.79)	16.6 (1.80)	18.0 (1.98)
% Completeness	98.6 (99.0)	99.3 (100.0)	98.8 (96.5)
Redundancy	3.8 (3.6)	4.2 (4.1)	4.2 (4.2)
Refinement			
Resolution range (Å)	42.734-2.291 (2.348-2.291)	42.805-2.295 (2.353-2.295)	42.696-2.281 (2.338-2.281)
No. of reflections in working set	48445	48588	48654
No. of reflections in test set	2006	2003	2009
R_{work} (%)[d]	17.15 (20.84)	18.17 (24.92)	18.54 (26.03)
R_{free} (%)[e]	21.50 (27.12)	23.05 (26.44)	22.81 (30.23)
Wilson B factor (Å^2)	30.81	34.75	35.11
Average B factor (Å^2)	44.9	49.4	58.2
RMSD from ideal geometry			
Bond length (Å)	0.013	0.010	0.010
Bond angle (deg)	1.06	1.06	1.07
Ramachandran plot			
Most favored (%)	94.88	94.17	94.15
Allowed (%)	4.38	4.86	4.51
Disallowed (%)	0.73	0.97	1.34

[a]Values in parentheses are for the last (highest resolution) shell.

[b]$R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of a given reflection, and $\langle I(hkl)\rangle$ is the mean intensity of symmetry-related reflections.

[c]$R_{pim} = \sum_{hkl}\sqrt{(\frac{1}{n}-1)}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i(hkl)$, where n is the multiplicity for multiplicity-weighted R_{merge}.

[d]$R_{work} = \sum_{hkl}\big||F_{obs}| - |F_{calc}|\big| / \sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[e]R_{free} was calculated using 4% of the data set chosen at random that were excluded from the refinement.

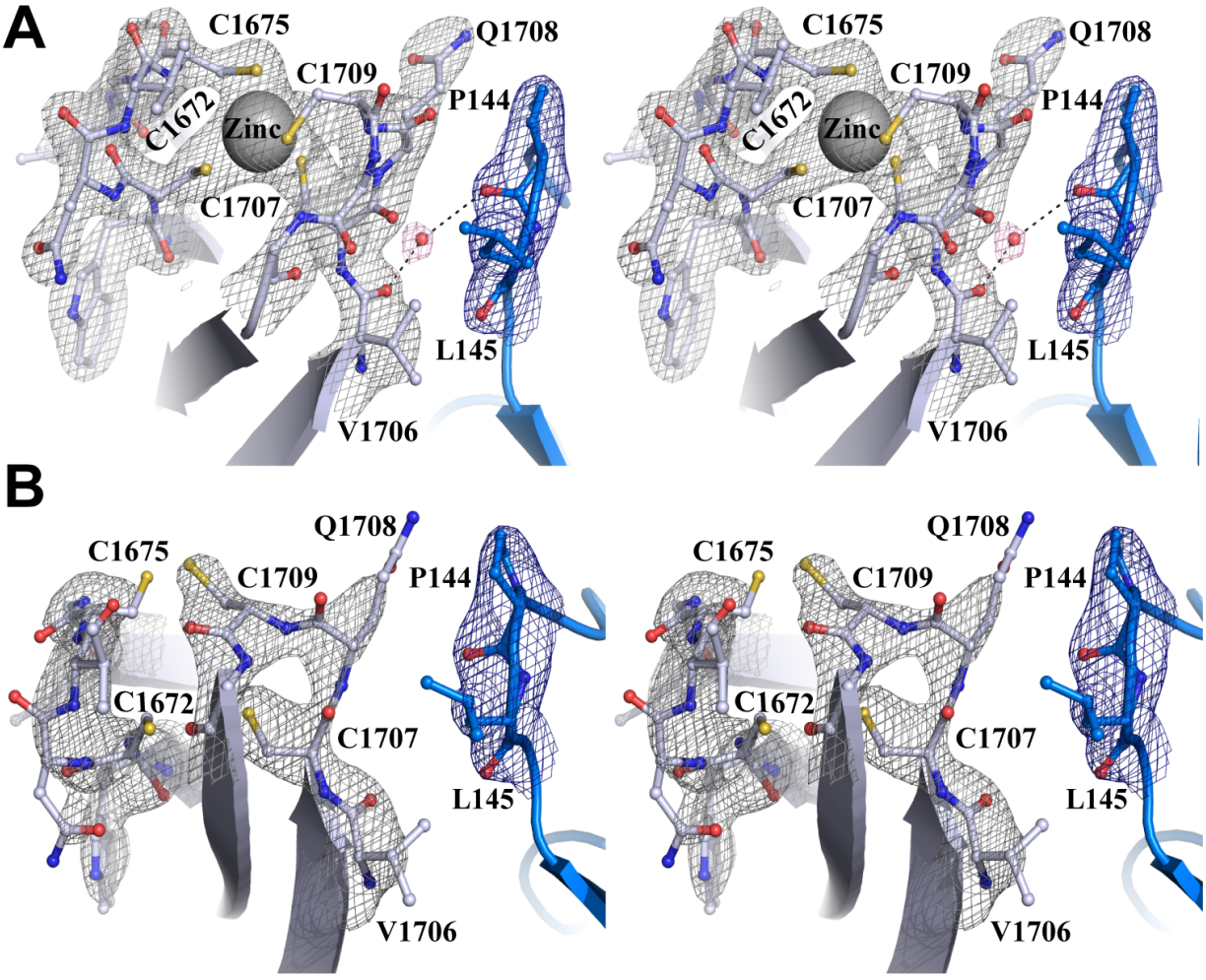


Figure 4.5 Structural zinc chelation upon increasing citrate concentration.

Stereoview of the $2F_o$-F_c electron density maps (contoured to 1 σ) of MERS-CoV PLpro-ISG15 complex data collected from crystallization conditions containing 0.25 M (A) versus 0.35 M (B) potassium citrate tribasic. Electron density maps are represented as gray, blue, and pink mesh for PLpro, ISG15, and a water molecule, respectively. Cysteine ligands (PLpro, gray) and interacting residues (PLpro, gray; ISG15, blue) are shown as sticks. Panel A shows the zinc atom represented as a gray sphere, and backbone carbonyls coordinate a water molecule (red sphere) with hydrogen bonds represented by black dashed lines.

4.3.2 Interactions of MERS-CoV PLpro with ISG15

The interactions between MERS-CoV PLpro and ISG15 are predominantly centralized to the ubiquitin recognition subsite S1 (SUb1) of PLpro [77] which is composed of residues at the interface between the fingers and thumb domains (Figure 4.6). Both water-mediated and

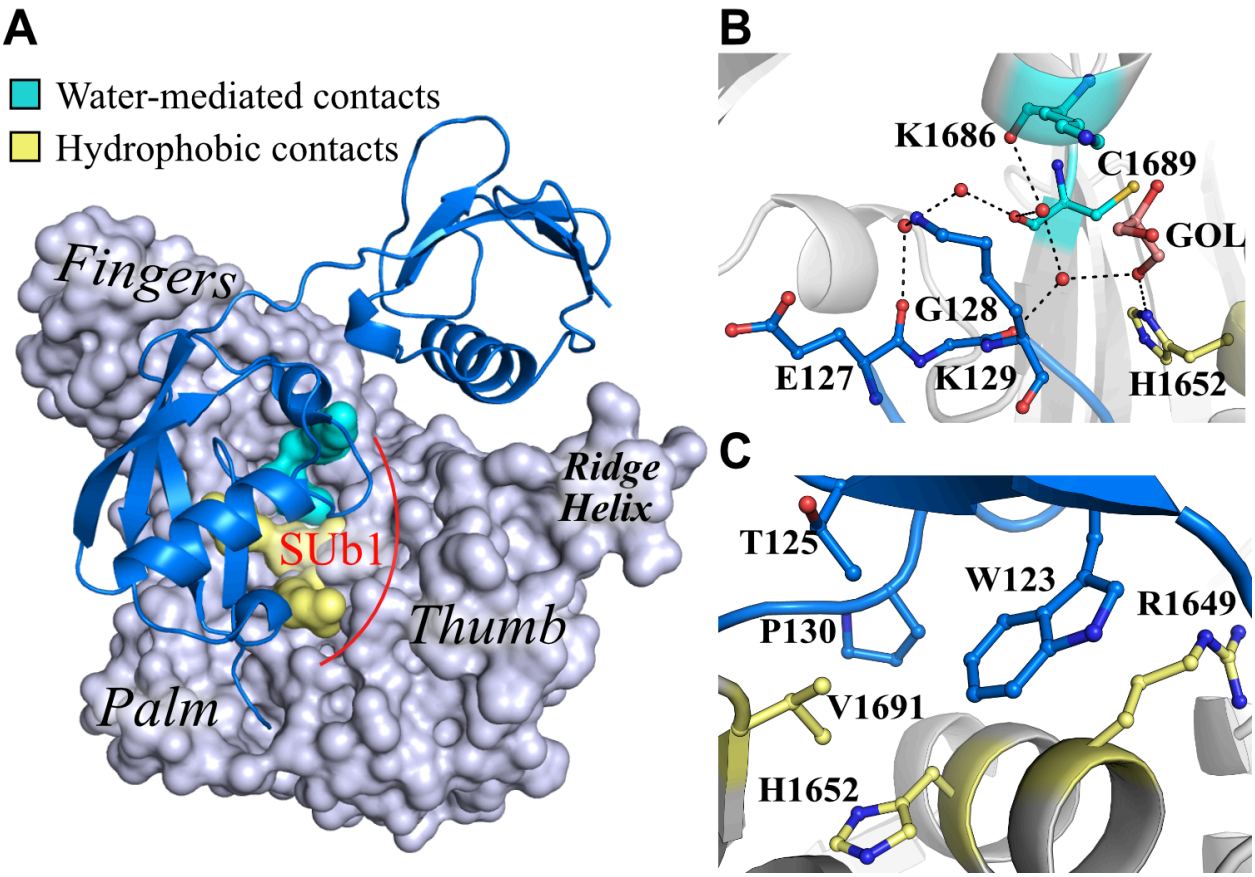


Figure 4.6 PLpro S1 subsite and ISG15 interactions involve both water-mediated and hydrophobic contacts

(A) Structure of PLpro-ISG15 complex. PLpro is shown as a gray surface representation with the thumb-palm-fingers domains labeled as well as the N-terminal ridge helix of the thumb domain. ISG15 is show in a blue cartoon representation. Water-mediated and hydrophobic surface residues of the PLpro SUb1 are colored in cyan and yellow, respectively where contacts are made with the C-terminal domain of ISG15. (B) Close-up view of water-mediated network (cyan) shows peptide backbone and side-chains of Lys129 of ISG15 (blue sticks) and His1652 of PLpro (yellow sticks) coordinating water molecules (red spheres) and a glycerol (GOL, pink sticks) molecule between PLpro and ISG15. (C) Magnified view of hydrophobic contacts (yellow) where no solvent molecules are present. Trp123, Pro130, and Thr125 of ISG15 (blue sticks) are accommodated to form a hydrophobic pocket with MERS-CoV PLpro. Atoms in Panels B-C are colored as follows; nitrogens (dark blue), oxygens (red), sulfur (yellow) carbons (light blue).

hydrophobic interactions are observed in the contacting regions (Figure 4.6.A). The water-mediated network is predominately formed by main-chain interactions between the carbonyl groups of Lys1686 and Cys1689 in the fingers domain of PLpro and the carboxyl side chain groups of Glu127 and Gly128 of ISG15 (Figure 4.6.B). The side-chain of Lys129 in ISG15 also participates in a water-mediated interaction with MERS-PLpro.

Van der Waals interactions between the Cβ atom of His1652 and residues Pro130 and Trp123 are also observed forming a small hydrophobic pocket (Figure 4.6.C). Trp123 of ISG15 is sandwiched between the aliphatic chain of Arg1649 of PLpro and the hydrophobic side-chain of Val1691, which also interacts with the alkyl group of Thr125 of ISG15. Apart from Thr125, which is unique to human ISG15, the proline and tryptophan residues are conserved among species of ISG15, and appear to be utilized for interactions by a wide range of deISGylases, both in the USP and ovarian tumor family (vOTU), suggesting that this region of PLpro may be critical for ISG15 recognition [74, 119, 138].

Surprisingly, there are no strong interactions between the N-terminal domain of ISG15 and the ridge helix of PLpro, which is the first helix of the thumb domain, adjacent to that of the Ubl2 domain in PLpros [77]. After data refinement, residue Lys1550 of the ridge helix of MERS-CoV PLpro and Gln34 of ISG15 appeared to be within hydrogen bonding distance (Figure 4.7.B). However, electron density associated with these side-chains, especially Gln34 of ISG15 is weak suggesting that these residues may not interact (Figure 4.7.B). To test this hypothesis, Lys1550 was mutated to an alanine residue as well as the equivalent phenylalanine residue of the ubiquitin recognition subsite S2 (SUb2) in SARS PLpro and no differences were observed in the k_{cat} values for the response of MERS-CoV Lys1550Ala and Lys1550Phe to the substrate ISG15-AMC compared to wild type (Figure 4.10.A).

To look closer at the flexibility of ISG15 in general, we aligned ISG15 from our MERS-CoV PLpro bound structure to other free and bound ISG15 structures comparing the alignment of the entire substrate versus just the C-terminal domain (Table 4.4). This was done by matching atoms C_α using least squares fit in COOT. The free mouse ISG15 was excluded from the structural alignment because it has already been shown to have adopted an alternative tertiary structure different from that of free human ISG15 and other bound structures solved to date [119].

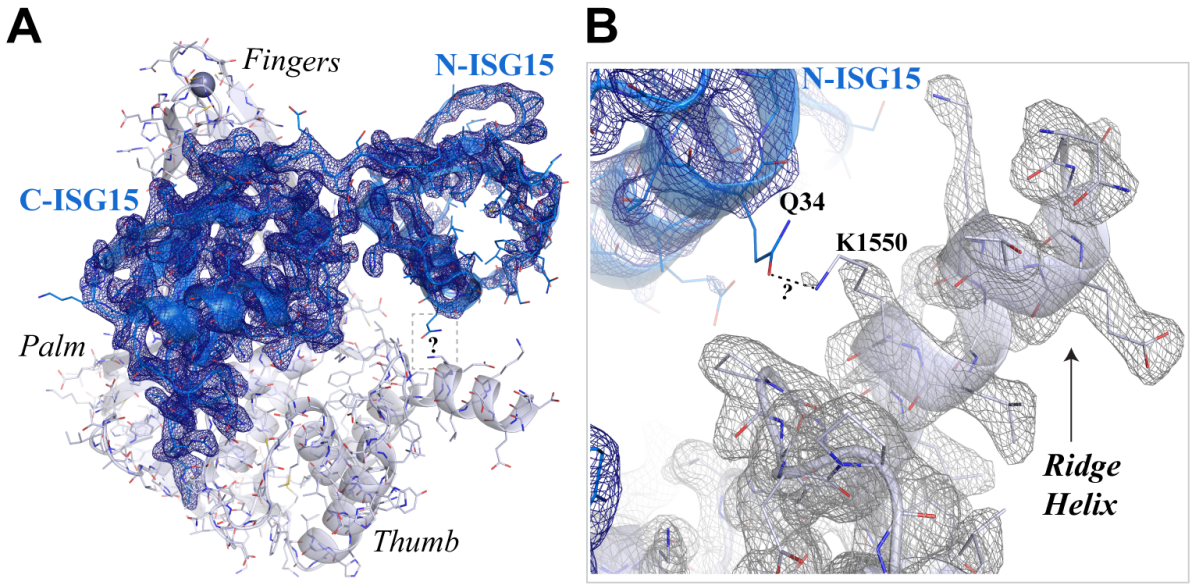


Figure 4.7 Electron density maps reveal that the N-terminal domain of ISG15 (blue) has more flexibility and weakly associates with MERS-CoV PLpro (gray)

(A) Overall structure of PLpro with the three subdomains labeled bound to ISG15. $2F_o$-F_c electron density map of the complete ISG15 (blue mesh) is contoured to 1σ, which show that N-terminal domain of ISG15 (N-ISG15) has a weaker overall density map compared to the C-terminal domain of ISG15 (C-ISG15). PLpro and ISG15 are shown as a cartoon with residues represented as lines. A gray dotted box highlights the region of interest. (C) Magnified view of 'Ridge Helix' (arrow), and its corresponding $2F_o$-F_c electron density map (gray mesh) contoured to 1σ, and N-ISG15 reveals one interaction in hydrogen bonding distance (dotted lines), which was observed after data refinement. Weak side-chain densities pose the question on if there is a direct interaction, which is illustrated with a question mark.

Table 4.4 Structural comparison of human ISG15 bound to MERS-CoV PLpro with other full-length and C-terminal domain ISG15 structures

	MERS PLpro-hISG15 (chain C)			
	PDB/chain ID	RMSD (Å)	C_α	References
Free hISG15	1Z2M/A	1.88	151	[123]
vOTU-hISG15	3PSE/C	1.87	148	[139]
USP18-mISG15	5CHV/C	4.01	151	[74]
USP18-mISG15	5CHV/D	3.77	151	[74]
NS1-hISG15	3SDL/C	1.51	151	[140]
NS1-hISG15	3SDL/D	1.59	151	[140]
MERS PLpro-hISG15	6BI8/D	0.82	151	This study
	MERS PLpro-hISG15 (chain C)			
Free hISG15	1Z2M/A	0.47	73	[123]
vOTU-hISG15	3PSE/C	0.60	73	[139]
USP18-mISG15	5CHV/C	0.69	73	[74]
USP18-mISG15	5CHV/D	0.67	73	[74]
NS1-hISG15	3SDL/C	0.45	73	[140]
NS1-hISG15	3SDL/D	0.44	73	[140]
NS1-hCISG15	3PHX/B	0.54	73	[138]
SARS PLpro-hCISG15	5TL6/A	0.53	73	[119]
SARS PLpro-hCISG15	5TL6/C	0.52	73	[119]
SARS PLpro-mCISG15	5TL7/A	0.63	73	[119]
SARS PLpro-mCISG15	5TL7/C	0.50	73	[119]
MERS PLpro-hISG15	6BI8/D	0.17	73	This study

Overall human ISG15 deviates the most with mouse ISG15 (mISG15) bound to USP18 with a RMSD of nearly 4 Å while free and bound human ISG15 (hISG15) structures align relatively well in comparison (RMSD of 1.5-1.8 Å). Across the board, the C-ISG15 shows even less deviations (RMSD < 0.7 Å) compared to the full-length structures. In fact, C-ISG15 of free ISG15 aligns well with C-ISG15 (RMSD = 0.47 Å) from our bound structure. This suggest that the conformation of C-ISG15 does not significantly change upon binding to deISGylases. Overall, C-ISG15 has more defined electron density and secondary structure elements while N-ISG15 appears have more looping regions in our structure. These results and our electron density maps

support that N-ISG15 is more flexible than the C-ISG15, and this may explain why deISGylases recognize the latter. Interestingly, NS1 from influenza virus only contacts the former as a mechanism to inhibit ISG15 conjugation in cells [140]. Therefore, it is possible that other ISG15-binding proteins may stabilize the N-ISG15 domain.

4.3.3 Structural differences in recognition of ISG15 and Ub by MERS-CoV PLpro

To identify potential residues of MERS-CoV PLpro that dictate substrate specificity for ISG15 versus Ub, we superimposed the MERS-CoV PLpro-ISG15 structure was superimposed with the Ub-bound structure (PDB accession code 4RF1), and the result is shown in Figure 4.8. The C-terminal domain of ISG15 (residues 82-156) and Ub (residues 1-75) superimpose quite well with a RMSD of 0.95 Å over 75 Cα atoms, and the PLpro structures superimpose with a RMSD of 1.16 Å over 252 Cα atoms (Figure 4.8.A). As expected, PLpro interacts with RLRGG residues at the C-terminus of both ISG15 and Ub substrates in a nearly identical manner by utilizing an intricate hydrogen-bonding network mainly composed of main-chain interactions (Figure 4.8.B-D). These interactions were described previously described for Ub [83] with one minor difference in the orientation of Arg155 of ISG15 (equivalent to Arg74 of Ub) as shown in Figure 4.8.D. Arg155 of ISG15 corresponds to X in the P3 position of the LXGG substrate recognition sequence of PLpro. In the Ub-bound structure, the side-chain η-amino group of Arg74 is hydrogen-bonded to the main-chain carbonyl of Thr1755, which is found in the substrate-binding loop of MERS-CoV PLpro, while in ISG15, triethylene glycol (PGE) is bound in place of the Arg74 side-chain (Figure 4.8.B-C). Instead, Arg155 of ISG15 undergoes a ~122° flip and interacts with the main-chain carbonyl of Pro1644 in the adjacent loop of the thumb domain between PLpro helix α6-α7. This alternative conformation of Arg155 was also observed in the structure of MERS-CoV PLpro

in complex with only the C-terminal domain of ISG15 [133]. Therefore, there is observable flexibility for Arg in the P3 position in adopting different contacts with PLpro.

In both ISG15- and Ub-bound structures, the guanidinium group of Arg1649 in PLpro forms a stacked interaction with the guanidium group of Arg153 (equivalent to Arg72 of Ub), which is P5-Arg of the RLRGG motif (Figure 4.8.E, region #1). However, the structural differences are observed with residues surrounding Arg1649. In the Ub-bound complex, Arg42 of Ub positions itself to alleviate charge-charge repulsion with Arg1649 while in ISG15, Trp123 forms a more favorable interaction with Arg1649 by packing against its aliphatic chain. Another difference observed between ISG15 and Ub-bound structures is that Thr1653 only forms direct hydrogen-bond interactions with Gln49 (equivalent to Pro130 of ISG15) and Glu51 (equivalent to Glu132 of ISG15) in the Ub-bound complex whereas in the ISG15-bound complex Thr1653 forms no direct interactions (Figure 4.8.E, region #1). In the structure of MERS-CoV with full-length ISG15, Glu132 of ISG15 does not appear to make strong contacts with PLpro. Though Glu132 was noted as a 'principle electrostatic interaction' in the structure of MERS-PLpro with N-terminally truncated ISG15, it forms a fairly weak electrostatic interaction (~3.8 Å) with Lys1657 in the full-length ISG15 structure, which was only adopted in a single monomer. This, the presence of the N-terminal domain in full-length ISG15 influences interactions of the C-terminal domain of ISG15 with PLpro.

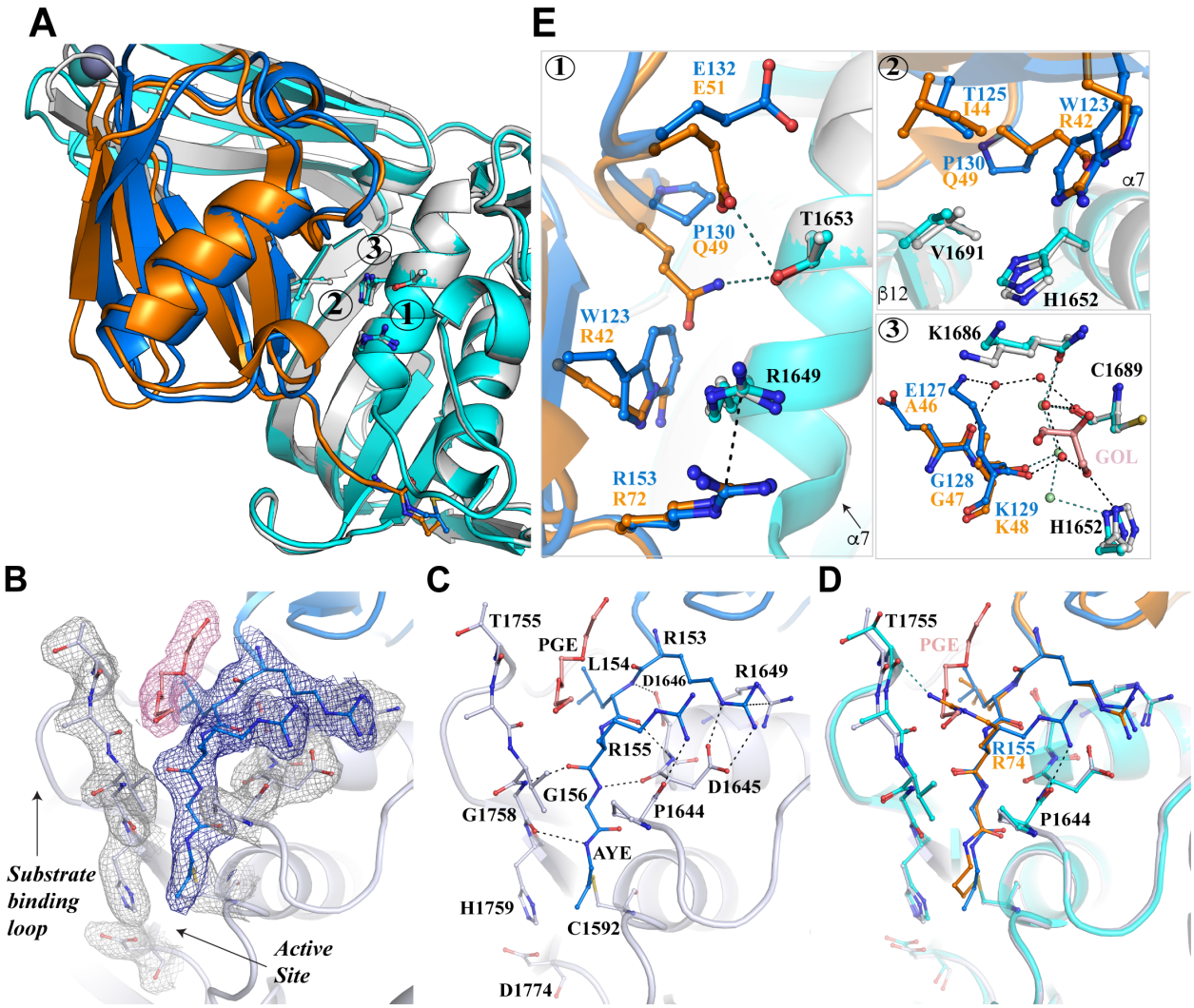
Another common feature between the ISG15 and Ub bound MERS-CoV structures is that Val1691 of PLpro forms Van der Waals contacts with Ile44 in the Ub hydrophobic patch and Thr125 in ISG15 (Figure 4.8.E, region #2). However, in ISG15 this hydrophobic 'patch' appears to be extended by Trp123 and Pro130, thus allowing His1652 of PLpro to also engage in an interaction. As previously noted by Basters et al. for USP18, this region of ISG15 is conserved across different ISG15 species and was proposed to be a unique hydrophobic patch to ISG15

differing from the well-known hydrophobic patch of Ub, which is commonly utilized for recognition by a vast number of Ub-binding proteins [74, 141]. Our observations and those of others collectively suggest that this hydrophobic region of ISG15 may be a canonical recognition region that is utilized in general by deISGylases to specifically recognize ISG15.

Finally, the water-mediated contacts involving the side-chain imidazole of His1652 are also conserved between ISG15 and Ub bound MERS-CoV PLpro structures (Figure 4.8.E, region #3). Although the interactions appear more extensive in the PLpro-ISG15 complex, His1652 of MERS-CoV PLpro is observed to bridge with a water molecule back to the main-chain carboxyl of a glycine residue, Gly47 of Ub or Gly128 of ISG15, in both structures. A similar water-mediated network was also observed in the SARS-CoV PLpro-Ub aldehyde structure suggesting that this is a canonical region where intervening solvent molecules bridge indirect interactions between the substrate and the protease [77].

Figure 4.8 Superposition of X-ray crystal structures of MERS-CoV PLpro (gray, PDB code 6BI8) bound to ISG15 (blue) and MERS-CoV PLpro (cyan, PDB code 4RF1) bound to Ub (orange)

(A) Overall view of the shared Ub and C-ISG15 binding pocket with key PLpro S1 subsite interactions designated 1-3. (B) $2F_o$-F_c electron density maps are contoured to 1 σ and reveal strong data for interacting residues of the substrate-binding loop, adjacent loop between helix α6-α7, and catalytic triad of PLpro as well as the RLRGG motif of ISG15. (C) Hydrogen bonding network interactions between PLpro and RLRGG motif of ISG15 (represented as black dashed lines). AYE; allylamine. PGE; triethylene glycol. (D) Overlay of MERS-CoV PLpro-ISG15 and –Ub active site interactions reveals a difference in conformation for the Arg155 side-chain of ISG15 (equivalent to Arg74 of Ub) due to a bound PGE molecule. Black and cyan dashes represent hydrogen bonds in the PLpro-ISG15 and –Ub complex, respectively. (E) Magnified views of key interactions (1-3) in both structures. Bonding residues of PLpro to ISG15 and Ub are shown as black or cyan dashed lines, respectively. Region 1; Interactions with PLpro α7 show that Arg1649 orientations are shared while Thr1653 only interacts with Ub. Arg1649 easily accommodates Trp123 of ISG15 while the charged Arg42 of Ub may not be favored. Region 2; Residues of PLpro α7 and β12 form a larger hydrophobic pocket with Trp123 and Pro130 of ISG15. Hydrophobic contacts between Val1691 are shared at equivalent position Thr125 of ISG15. Region 3; His1652 imidazole and peptide backbone residues of PLpro coordinate solvent molecules in a similar fashion to a conserved glycine residue. Less water (cyan spheres) molecules are observed in the Ub-bound structure while in the ISG15-bound structure more waters (red spheres) and a glycerol (GOL, pink sticks) molecule was observed.


A
E
1
2
3
E132
E51
P130
Q49
T1653
W123
R42
R1649
R153
R72
α7
T125
I44
W123
R42
P130
Q49
α7
V1691
β12
H1652
K1686
C1689
E127
A46
GOL
G128
G47
K129
K48
H1652
B
Substrate
binding
loop
Active
Site
C
T1755
PGE
L154
R153
D1646
R1649
R155
G156
D1645
G1758
P1644
AYE
H1759
C1592
D1774
D
T1755
PGE
R155
R74
P1644

4.3.4 Structure-guided design of MERS-CoV PLpro mutants

Structure-guided engineering of SARS-CoV PLpro has previously been shown to be a feasible strategy for selectively decoupling SARS PLpro DUB and/or deISG activities while maintaining the ability of the enzyme to efficiently process the polyprotein [77, 83]. We therefore sought to use our MERS-CoV PLpro bound ISG15 structure and the previously determined Ub-bound structures to design a set of MERS-CoV PLpro mutants that are either DUB or deISG deficient, or both DUB and deISG deficient. The ultimate goal is to design mutants that will serve as tools to delineate the function of DUB versus deISG activities in virus-infected cells and for potentially generating attenuated virus strains for the purpose of creating a live-attenuated MERS-CoV vaccine. We focused our mutagenesis on the following MERS-CoV PLpro residues; Arg1649, His1652, Thr1653, and Val1691, which all reside in the MERS-CoV PLpro SUb1 subsite. Mutants were designed based on two approaches: (i) alanine-scanning mutagenesis to completely remove an interaction, and (ii) insertion of a bulky or charged residue to create steric hindrance, i.e unfavorable Van der Waals overlap, or electrostatic repulsion to disrupt contacts. Additionally, we sought to engineer MERS-CoV PLpro to be more SARS-like, as Lys1550Phe and Arg1649Glu were mutated based on the equivalent residue found in SARS-CoV PLpro to decipher the residues that contribute to SARS-CoV PLpro enhanced turnover rate compared to MERS-CoV PLpro for different substrates [94].

In order to streamline the purification process for the mutants, expression was done at a smaller scale (500 mL cell culture) and gravity flow columns were utilized for purification, which can be run in series. A reduced yield (1-5 mg) of PLpro was obtained using this method but was sufficient for kinetic assays. The wild-type was also purified with this procedure and comparable specific activities values were obtained between protocols (~300-400 μM/min/mg). SDS-PAGE

analysis of the wild-type and mutants MERS PLpro enzymes that were used for kinetics experiments, are provided in Figure 4.9.

Next, we performed an initial kinetic assessment of the kinetic activity of 13 PLpro mutants at a single substrate concentration by testing their ability to hydrolyze three commercially available fluorogenic substrates commonly used to evaluate PLpro peptide, DUB and deISG activities. Results from the initial assessment are summarized in Figure 4.10 as the percent relative activity to the wild-type enzyme. Mutants that did not exhibit a significant alteration in kinetic activity towards any of the three substrates include Lys1550Ala, Lys1550Phe, His1652Ala, and

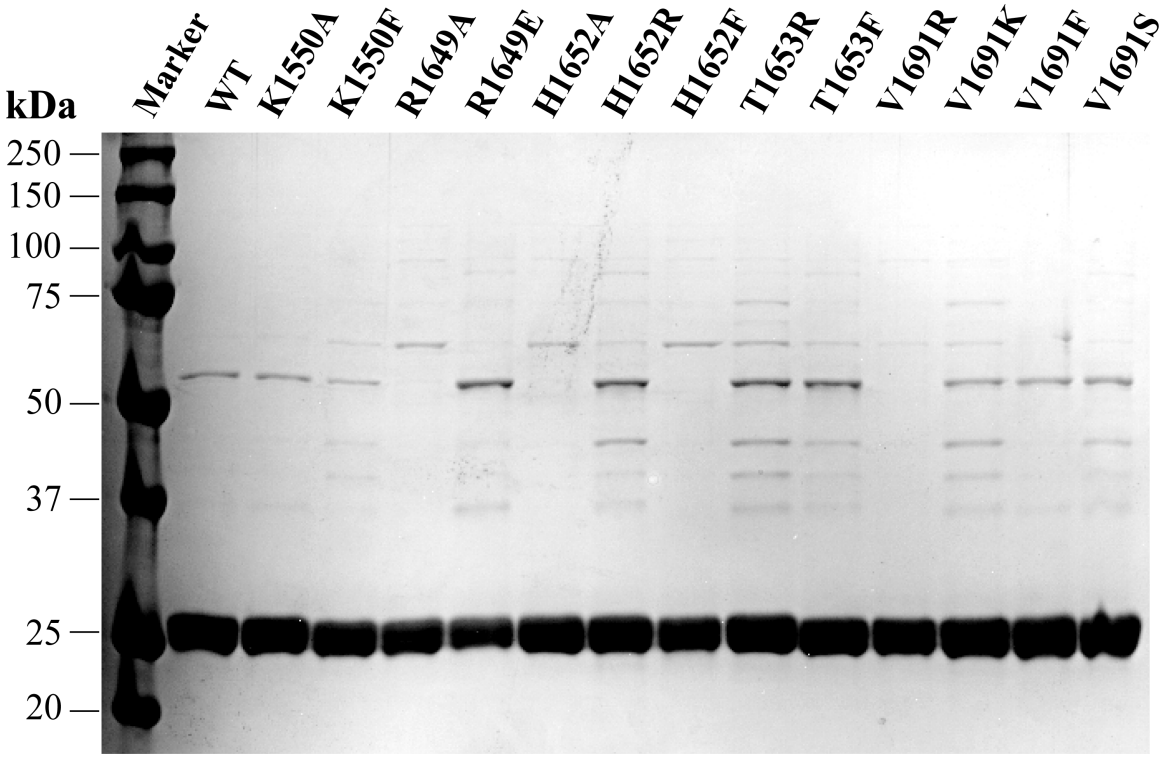


Figure 4.9 12.5% SDS-PAGE analysis of MERS-CoV PLpro wild-type and mutants used for kinetic studies

The proteins (loaded at 4 μg) were estimated to be >90% purity by densitometry.

Thr1653Phe. Mutants that did show reduced peptide hydrolysis and severely impaired hydrolysis for both DUB/deISG activities were His1652Arg, Val1691Arg, and Val1691Lys. Val1691Ser displayed ~50% reduction in both DUB/deISG activities, but the peptide hydrolysis of this mutant

was similar to the wild-type. Surprisingly, Arg1649Ala and Arg1649Glu exhibited enhanced hydrolysis towards both the Ub-AMC and the peptide substrates while the activity towards ISG15-AMC was not significantly altered. Compared to these mutants, Val1691Phe exhibited moderately enhanced Ub-AMC activity. Finally, His1652Phe and Thr1653Arg showed ~30-50% Ub-AMC and the peptide activities but maintained ISG15-AMC activity.

To confirm observed trends and to more fully characterize the kinetic properties for the Arg1649Ala, Arg1649Glu, His1652Phe, Thr1653Arg, Val1691Phe, and Val1601Ser mutants, the kinetic parameters k_{cat} and K_m were determined for each of these mutants with the substrates Ub-AMC and ISG15-AMC and the results are summarized in Table 4.5 along with the parameters previously determined for other mutants [133]. The catalytic efficiency (k_{cat}/K_m) of each mutant and the selectivity ratio (k_{cat}/K_m for ISG15-AMC divided by k_{cat}/K_m for Ub-AMC, and vice versa) for each mutant was also calculated and is compared to the wild-type. From these data, it is observed that mutating the ridge helix residue Lys1550 to either an alanine or the equivalent phenylalanine residue in SARS-CoV PLpro, had no significant effect on the functionality of MERS-CoV PLpro (Figure 4.10). This result supports our prediction that the potential hydrogen bond observed after data refinement between Lys1550 of PLpro and Gln34 of ISG15 is unlikely and not important for ISG15 recognition. This observation stands in stark contrast to SARS-CoV PLpro that has two ubiquitin recognition sites (SUb1 and SUb2) for ISG15 and K48-linked diUb chains [77, 81]. Based on the PLpro-ISG15 recognition model proposed by Ratia et al., SARS-CoV PLpro may utilize its ridge helix to directly engage with the N-terminal domain of ISG15 [77]. Our results indicate that MERS-CoV PLpro does not possess an equivalent SUb2 subsite.

Figure 4.10 Initial and complete kinetic characterization of MERS-CoV wild-type and mutants

(A) Initial assessment of catalysis against three ubiquitin-based fluorogenic substrates at a single substrate concentration: 50 μM RLRGG-AMC (green), 1 μM Ub-AMC (orange), and 1 μM ISG15-AMC (ISG15). PLpro mutants are plotted as a function of relative activity of the wild-type (%) based on the calculated turnover number. Error bars represent standard deviations calculated from triplicate measurements. The location of each residue side-chain position of PLpro (spheres) used for mutagenesis is shown in the structural representation of PLpro (gray). ISG15 was omitted for clarity. (B-D) Kinetic rate (velocity/[Enzyme]) of the wild-type and six mutants to increasing ISG15-AMC (B) and Ub-AMC (C-D) concentrations. Gray shaded box in panel C is magnified for clarity in panel D. All data were fit to the Michaelis-Menten equation with the exception of Thr1653Arg and Val1691Ser, where data failed to reach saturation and were fit to a linear regression equation. Error bars represent a range from duplicate measurements.

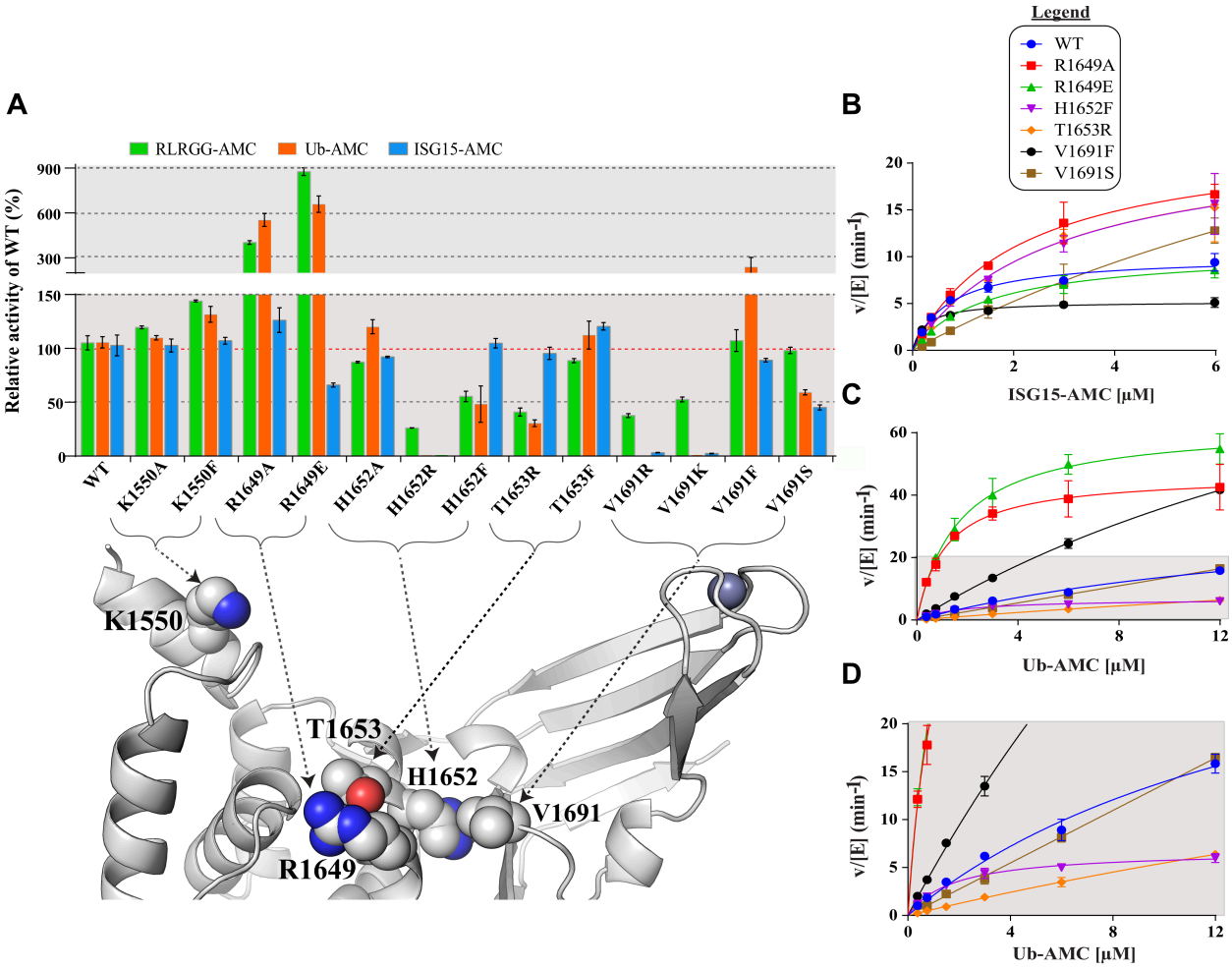

A
RLRGG-AMC
Ub-AMC
ISG15-AMC
Relative activity of WT (%)
900
600
300
150
100
50
0
WT
K1550A
K1550F
R1649A
R1649E
H1652A
H1652R
H1652F
T1653R
T1653F
V1691R
V1691K
V1691F
V1691S
K1550
T1653
H1652
V1691
R1649
B
Legend
WT
R1649A
R1649E
H1652F
T1653R
V1691F
V1691S
v/[E] (min-1)
ISG15-AMC [μM]
C
v/[E] (min-1)
Ub-AMC [μM]
D
v/[E] (min-1)
Ub-AMC [μM]

Table 4.5 Summary of Kinetic parameters of fully characterized PLpro mutants

Mutant	K_m (μM)	k_{cat} (min-1)	Fold activity	k_{cat}/K_m (μM^{-1} min^{-1})	Fold efficiency	Selectivity
ISG15-AMC[a]						
WT	0.73 ± 0.10	10.05 ± 0.44	1.0	13.81 ± 2.04	1.0	6.2
R1649A	2.16 ± 0.33	22.84 ± 1.51	2.3	10.60 ± 1.76	0.8	0.3
R1649E	1.50 ± 0.21	10.69 ± 0.57	1.1	7.13 ± 1.06	0.5	0.2
H1652F	2.80 ± 0.69	22.65 ± 2.63	2.3	8.08 ± 2.20	0.6	2.1
T1653R	2.82 ± 0.74	22.73 ± 2.80	2.3	8.07 ± 2.33	0.6	12.5
K1657E[b]	N.S.[c]	N.S.	N.S.	0.74 ± 0.17	0.1	0.2
V1691F	0.25 ± 0.04	5.20 ± 0.17	0.52	20.80 ± 3.13	1.5	4.1
V1691D[b]	N.S.	N.S.	N.S.	0.23 ± 0.05	0.017	0.1
V1691S	N.S.	N.S.	N.S.	2.23 ± 0.10	0.2	1.6
Ub-AMC[a]						
WT	16.58 ± 4.20	37.07 ± 6.21	1.0	2.24 ± 0.68	1.0	0.2
R1649A	1.13 ± 0.20	46.51 ± 2.36	1.3	41.31 ± 7.58	18.5	3.9
R1649E	1.65 ± 0.21	62.62 ± 2.57	1.7	37.93 ± 5.08	17.0	5.3
H1652F	1.77 ± 0.21	6.76 ± 0.26	0.2	3.81 ± 0.47	1.7	0.5
T1653R	N.S.	N.S.	N.S.	0.55 ± 0.01	0.2	0.1
K1657E[b]	3.17 ± 0.37	12.22 ± 0.43	0.3	3.85 ± 0.12	1.7	5.2
V1691F	26.17 ± 3.01	132.30 ± 11.29	3.6	5.06 ± 0.72	2.3	0.2
V1691D[b]	10.04 ± 2.33	29.44 ± 2.40	0.8	2.93 ± 0.25	1.3	12.7
V1691S	N.S.	N.S.	N.S.	1.37 ± 0.02	0.6	0.6

[a]Steady-state values were determined from duplicate measurements, reported as a mean ± standard error.
[b]Kinetic parameters from [133].
[c]NS, enzyme not saturated. k_{app} values reported as a slope ± best-fit error from a linear regression were used to approximate k_{cat}/K_m.

The data in Figure 4.10.A also suggest that the residue in position 1691 of the PLpro SUb1 subsite is very sensitive to charge, polarity, and the size of the amino acid introduced into this position. Adding a negatively charged aspartate residue at this position was shown to selectively disrupt the recognition of ISG15-AMC by MERS PLpro while the ability to recognize and hydrolyze Ub-AMC was maintained [133]. The aspartate is likely able to make a salt bridge with Arg42 of Ub while in ISG15, the hydrophobic pocket cannot accommodate this charged residue. In contrast, substituting a positively charged arginine or lysine at this position severely impairs both activities of PLpro. For example, Val1691Arg results in a 32-fold reduction in ISG15-AMC hydrolysis and nearly a 2000-fold reduction in Ub-AMC hydrolysis. Bailey-Elkin et al. previously showed that Val1691Arg completely abolishes the ability of PLpro to antagonize IFN, as it no longer suppressed the IFN-β promoter in cells and mimicked the catalytically dead enzyme [83]. The loss of IFN antagonism was attributed to loss of DUB activity but deISG activity was not determined. Based on our results, the loss of IFN antagonism could be due to either the loss of either DUB or deISG activity or both.

In general, inserting a negatively or positively charged residue into the hydrophobic pocket of ISG15 is found not to be favored. In the case of binding Ub, only a positive charge is not permitted, as this would introduce charge-charge repulsion with Arg42 of Ub. To delineate if the observed effects of an insertion of an arginine at position 1691 was simply due to a bulk or steric overlap effect, Val1691 was also mutated to a phenylalanine. Surprisingly, Val1691Phe showed a 13-fold increase in the k_{cat} value for Ub-AMC while other activities were similar to the wild-type. Structurally, Phe1691 would allow for an additional π-cation interaction with Arg42 of Ub. On the other hand, engineering a more conservative mutation i.e. a small polar serine at the Val1691 position only reduced the ISG15-AMC efficiency (6.2-fold based on the k_{app} value) while Ub-AMC efficiency was less affected. The Val1691Ser mutant did not reach saturation for either Ub-

AMC or ISG15-AMC; the loss in efficiency was primarily due to an increase in K_m for ISG15 (Fig. 4.10.B and Table 4.5). The selectivity of Val1691Ser mutant towards ISG15-AMC decreased by 3.8-fold compared to the wild-type and is the only mutant that has similar catalysis for both substrates. This more conservative mutation also restored RLRGG-AMC activity back to that of the wild-type. We suggest that position 1691 in PLpro is critical for substrate recognition and based on the amino acid substitution at position 1691, it is possible to selectively modulate PLpro deISG (Val1691Asp, Val1691Ser) or DUB (Val1691Phe) activities or disrupt both DUB/deISG activities simultaneously (Val1691Arg, Val1691Lys).

His1652Ala, which is structurally adjacent to that of Val1691, did not significantly alter the enzymatic profile of MERS-CoV PLpro in our initial screening (Figure 4.10.A). We originally thought that the hydrophobic contacts and water-mediated interactions made by His1652 were not significant, but instead of performing alanine scanning, we sought to disrupt MERS-CoV PLpro activity by inserting a charged and bulky residue at this position as done with Val1691. Excitingly, mutation of the His1652 to an arginine drastically impaired both PLpro DUB/deISG activities by ~300-400-fold, a similar trend to the Val1691Arg and Val1691Lys mutants. In fact, to observe hydrolysis of these substrates, a higher concentration of enzyme was required and nearly approached the concentration of substrate in the assay. His1652Arg also exhibited a 4-fold reduction in RLRGG-AMC activity. In general, placing a bulky and charged residue in this region of PLpro does affect the active site functionality to a certain degree. We also analyzed if the water-mediated contacts involving His1652 were important by mutating this residue to a phenylalanine, which is incapable of hydrogen bonding. His1652Phe showed only a 1.7-fold change in catalytic efficiency for both ISG15-AMC and Ub-AMC. Interestingly, His1652Phe had a lower K_m value than wild-type for Ub-AMC (1.8 μM versus 16.6 μM) suggesting a 9-fold increase in binding affinity for this substrate, but there was also a 5.5-fold decrease in k_{cat} (Figure 4.10.D and Table

4.5). Overall, the net change was an increase in the catalytic efficiency for Ub-AMC about 2-fold over that for the wild-type enzyme. We observed a similar trend for Val1691Phe catalysis towards ISG15-AMC, but it was not as pronounced (Figure 4.10.B and Table 4.5). Overall, PLpro catalytic activity is more sensitive to having a positively charged amino acid at position 1652 rather than a smaller or bulky hydrophobic residue.

To further evaluate the DUB/deISG deficiencies of the His1652Arg, Val1691Arg, and Val1691Lys mutants, their ability to react and form a covalent modification with Ub-PA and ISG15-PA were tested. First, the optimal molar ratio required for complete conversion to the covalent adduct was determined using the MERS PLpro wild-type. Within one hour, MERS PLpro-wild-type was completely modified by Ub-PA and ISG15-PA at a 1:5 ratio and 1:20 ratio, respectively (Figure 4.11). These molar ratios were then used to analyze the reactivity of the mutants at different time points. While MERS PLpro wild-type completely reacted with both Ub-PA and ISG15-PA within a short 2 min incubation, all three mutants showed impaired reactivity at the earlier time points. Only approximately 50% PLpro-ISG15 complex was observed after the 5 min incubation. For the His1652Arg and Val1691Lys mutants, impaired reactivity at 5 min were also observed with Ub-PA. For Val1691Arg, however, even after a 30 min incubation only about 50% of PLpro-Ub complex was formed while other mutants were able to convert to complex after 30 min. This single mutation elicited a similar defect to Ub-PA as $USP18^{IBB\text{-}1\text{-}USP7}$ to ISG15-PA, which contained four mutations in the ISG15-binding box 1 (IBB-1) [74]. From these data, all three mutants are impaired for Ub/ISG15-PA reactivity. Consistent with our kinetic data, the

Val1691Arg mutation may target DUB activity more so than deISG activity and showed a striking defect in Ub-PA reactivity.

With the goal of creating a PLpro mutant that lacks DUB activity but retains deISG activity, we focused on residue Thr1653 as we observed from the X-ray structure that this residue only engages with the Ub substrate and not ISG15. Previously, the Thr1653Arg mutant showed reduced DUB activity but was not tested for deISG activity [83]. In addition, Thr1653Arg significantly inhibits PLpro-mediated IFN-antagonism in cells, but not quite to the extent of a Val1691Arg mutant [83]. From the kinetic studies presented in Figure 4.10.D, the Thr1653Arg mutant cannot

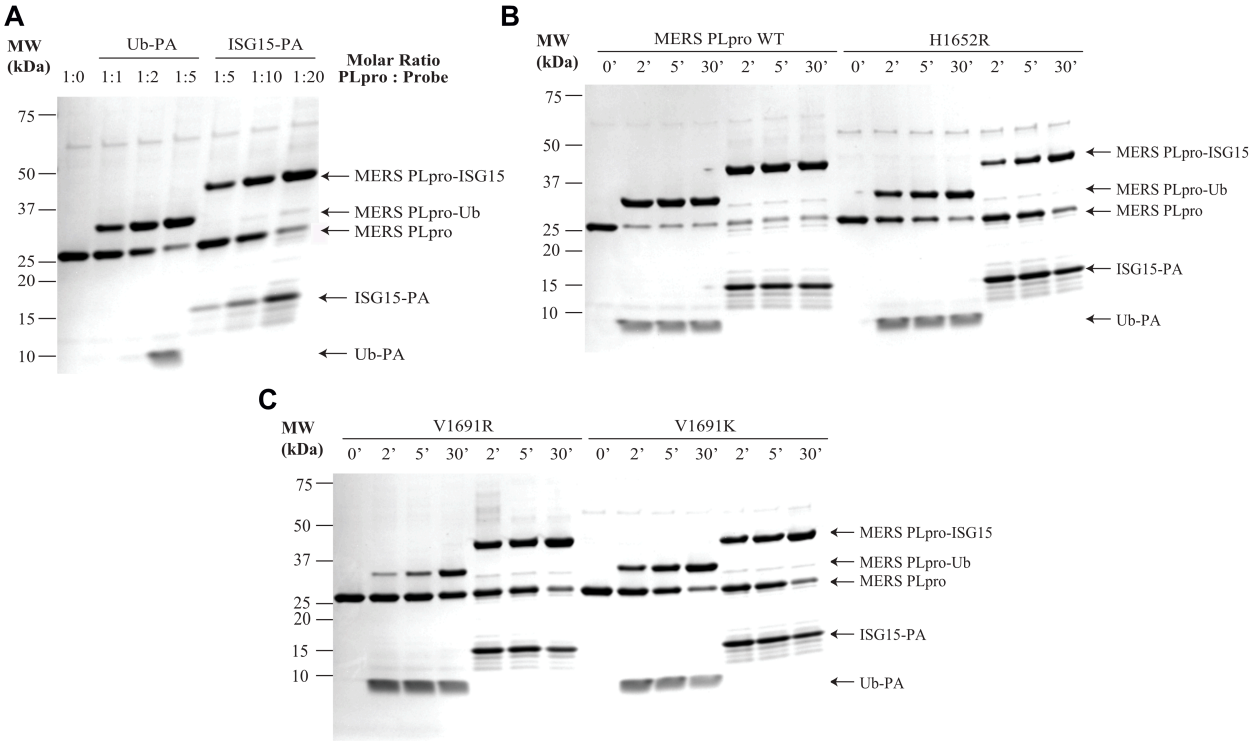


Figure 4.11 Defects in Ub-PA and ISG15-PA probe reactivity by DUB/deISG deficient mutants. (A) Coomassie-blue stained SDS-PAGE gel of the MERS PLpro wild-type reactivity towards Ub-PA and ISG15-PA at different molar ratios of PLpro to probe after an hour incubation. Ratio 1:0 serves as the negative control, which only contains 5 μM PLpro. (B-C) Reactivity of MERS PLpro wild-type and mutants toward Ub-PA (ratio 1:5) and ISG15-PA (ratio 1:20) at different time points (2, 5, and 30 min) from three independent experiments visualized by coomassie-blue stained SDS-PAGE gel. PLpro without addition of the PA probe served as the negative control (0').

be saturated with the Ub-AMC substrate over the concentration ranges tested. Thr1653Arg exhibits a 4.1-fold reduction in Ub-AMC catalytic efficiency compared to the wild-type and only a small

reduction in ISG15-AMC activity (Table 4.5). These results suggest that observed decrease in DUB activity in cells for Thr1653Arg is likely what caused the suppression of IFN-β promoter rather than a loss in deISG activity.

To delineate if the reduction in DUB catalytic activity was due to substituting a bulkier residue rather than a charged residue, we mutated Thr1653 to a phenylalanine. The kinetic data in Table 4.5 show that there is no significant effect on catalysis supporting the idea that the positive charge at this position is detrimental to DUB activity. Insertion of arginine at 1653 places it in close proximity to Arg42 of Ub and to a guanidinium stacking interaction between Arg1649 of PLpro and Arg72 of Ub. The net effect is a charge repulsion that decreases substrate binding and catalysis.

To probe the importance of Arg1649 of PLpro and its involvement in DUB and deISG activities, this residue was mutated to an alanine. Arg1649Ala exhibited an 18.5-fold and 3.9-fold enhancement in Ub-AMC and peptide efficiencies while its ISG15-AMC efficiency was not affected (Table 4.5). This is a surprising observation due to the fact that arginine likely interacts with the P5 arginine in the RLRGG motif of all three substrates. It is possible that removing the stacked guanidinium group interaction alleviates any electrostatic repulsion with Arg42 in Ub so that Ub is more easily accommodated. Conversely, since there is no extra charge at the equivalent Arg42 position in ISG15, the charged guanidinium groups are more properly balanced by negatively charged residues, such as Asp1645 and interacting waters. Therefore, the local environment of Arg1649 appears to differ when binding to each of these substrates.

Since SARS-CoV PLpro has a glutamate at the equivalent Arg1649 position, we thought that engineering MERS to be more SARS-like would also enhance its activity. Indeed, Arg1649Glu exhibited a similar trend as Arg1649Ala, 17-fold and 8.4-fold enhancement in Ub-AMC and peptide activity, while its ISG15-AMC activity was less effected. Both Arg1649Ala and Arg1649Glu exhibited nearly a 15-fold enhancement in K_m for Ub (Table 4.5 and Figure 4.10.C). These newly identified hyperactive DUB mutants have altered selectivity compared to the wild-type and favor Ub rather than ISG15 by 4 to 5-fold. In fact, these mutants have over 5-fold higher efficiency of hydrolysis towards Ub-AMC compared to SARS-CoV PLpro wild-type and similar efficiency to a hyperactive DUB mutant Arg167Glu previously characterized in SARS-CoV PLpro, which would also introduced an additional counter ion in this region near Arg42 of Ub [119]. These results provide insight into how to change MERS-CoV PLpro selectivity from ISG15 to Ub without affecting its deISG activity.

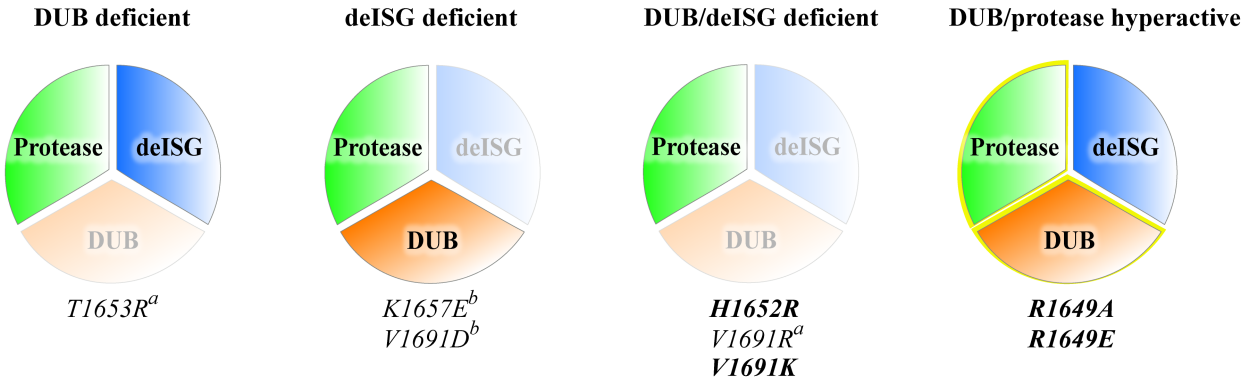


Figure 4.12 Functional tools of different MERS-CoV PLpro mutants

A pie-chart is used to represent the multifunctional activities of PLpro. All mutants investigated to date are shown characterized to be selectively DUB or deISG deficient, both DUB and deISG deficient, or are both DUB and protease hyperactive mutations. New sites of mutation investigated in this study are bolded. [a]PLpro mutants first investigated in [83]. [b]Mutants previously characterized from [46].

Altogether, from our data and previous studies, MERS-CoV PLpro mutants that are selectively DUB deficient, selectively deISG deficient, and both DUB/deISG deficient have been identified that fit each category (Figure 4.12). These mutants can be used as functional tools to

probe for DUB and deISG mechanism in MERS-CoV replication and pathogenesis. It would be interesting to investigate whether complete knockdown of both DUB/deISG activities of PLpro is necessary to efficiently attenuate the MERS virus. We also identified PLpro mutants that exhibited DUB and protease hyperactivity with potential to enhance MERS-CoV pathogenesis. Since these mutations do not exist in nature, it is possible that the DUB hyperactive virus may be nonviable or cannot be recovered using the reverse genetics system. Would these hyperactive viruses reach a steady state in order to exist in the population? This question still remains to be investigated.

4.4 Summary

The MERS-CoV PLpro-ISG15 crystal structure and PLpro mutants provide critical structural and functional data for the reverse genetics system to delineate the importance of DUB versus deISG activity in MERS-CoV pathogenesis and replication. MERS-CoV PLpro only actively engages with the C-terminal domain of ISG15 utilizing one SUb1 subsite, in contrast to SARS-CoV PLpro that may engage with the N-terminus of ISG15 using a second SUb2 subsite. MERS-CoV PLpro recognizes ISG15 by a hydrophobic surface distinct from that of Ub hydrophobic patch and insertion of a charge into this region of ISG15 reduces ISG15 specificity. Selectively disrupting or enhancing the DUB activity of PLpro can be accomplished by inserting a positive charge at position 1653 or removing the charge of Arg1649, respectively. Thr1653Arg will be useful mutation for probing DUB mechanism in MERS pathogenesis. Furthermore, mutation of His1652 or Val1691 to a positive charged residue completely impairs both DUB/deISG activities. DUB/deISG deficient mutants, His1652Arg, Val1691Lys, and Val1691Arg characterized in this study are great candidates for attenuating the MERS virus and may be useful for live-attenuated vaccine design.

CHAPTER 5. EVALUATING UBIQUITIN SPECIFICITY OF ALPHACORONAVIRUSES FIPV AND PEDV PAPAIN-LIKE PROTEASE 2 USING STRUCTURE-GUIDED ENGINEERING

The papain-like protease 2 (PLP2) domain is a conserved enzyme encoded by all coronavirus (CoV) subfamilies. From previous studies, it was determined that PLP2 generally acts as a protease to process the viral polyprotein for RNA replication and as a deubiquitinating (DUB) and deISGylating (deISG) enzyme to suppress innate immune responses for virus pathogenesis. In this study, we investigated the multiple activities of PLP2 *in vitro* from alpha-CoV feline infectious peritonitis virus (FIPV) and porcine epidemic diarrhea virus (PEDV), which cause fatal disease in household cats and suckling piglets, respectively. From kinetics and biochemical studies, we showed that FIPV and PEDV PLP2 are highly efficient DUBs rather than deISG enzymes that favor ubiquitin (Ub) over the interferon-stimulated gene 15 (ISG15) substrate. Additionally, the PEDV PLP2 structure was determined to 1.95 Å resolution and revealed a unique Zn-finger in contrast to other coronaviral PLP2 structures. By superimposing our structure with the MERS-CoV protease-Ub bound complex, residues involved in Ub recognition and catalysis were identified, and several mutants of FIPV and PEDV PLP2 were tested for DUB and protease activity. From our results, we identify a 'hot-spot' in the Ub-binding site that can be used to selectively disrupt DUB activity in both PLP2s with minimal loss in protease activity. Overall, this study provides candidate sites that may be used to attenuate alpha-CoV FIPV and PEDV for vaccine design.

5.1 Introduction

Coronaviruses (CoVs) are enveloped positive-sense single-stranded RNA viruses that cause gastrointestinal and respiratory diseases in humans, livestock, and companion animals. The

reoccurring outbreaks of porcine epidemic diarrhea virus (PEDV) continue to devastate swineherds across the globe since its first isolation in 1978 in Belgium. Since then different strains of PEDV have been detected across Europe and Asia, until PEDV 're-emerged' in 2013/2014 across farms in the United States, Canada and Mexico [19, 142, 143]. Within 1 year, the devastating PEDV epidemic caused 8 million deaths in piglets with economic losses estimated over 1 billion U.S. dollars. PEDV causes acute diarrohea, vomiting, and dehydration in newborn piglets with a morality rate of essentially 100%. Several PEDV vaccines have been developed, especially in Asia and Korea; however, current literature questions the efficacies and safety of these vaccines in the field [19, 144, 145]. With the recent outbreak of another fatal disease in pigs, swine acute diarrhoea syndrome (SADS) in 2016 [18], it stresses the importance of understanding virus-host interactions that facilitate viral pathogenesis and replication, as this may permit the design of effective therapeutics against these devastating viral pathogens.

The papain-like protease (PLP) domain, encoded in the RNA genome of CoVs, either as a single protease termed PLpro or PLP2, is multifunctional enzyme with therapeutic potential suitable for protein engineering studies. In general from studies over the last decade, PLP has been known to possess three activities: (i) viral polyprotein activity, (ii) deubiquitinating (DUB) activity, and (iii) deISGylating (deISG) activity, largely summarized in this review on SARS-CoV PLpro [68]. PLP cleaves the amino-terminal segments of the CoV polyprotein and matures the non-structural proteins, together with the 3-chymotrypsin-like protease (3CLpro), to support RNA replication. In addition, PLP removes ubiquitin (Ub) and interferon-stimulated gene 15 (ISG15) from target proteins, catalytic processes known as DUB and deISG activities, respectively. The DUB/deISG activities of PLP are one mechanism CoV employ to evade host innate immune responses, as PLP has been shown to antagonize IFN-β production, the first line defense of the host, to promote viral pathogenesis. Therefore, removing DUB/deISG activities of PLP but

maintaining the viral polyprotein processing activity is a potential strategy to activate the host innate immune response and possibly attenuate CoV pathogenesis.

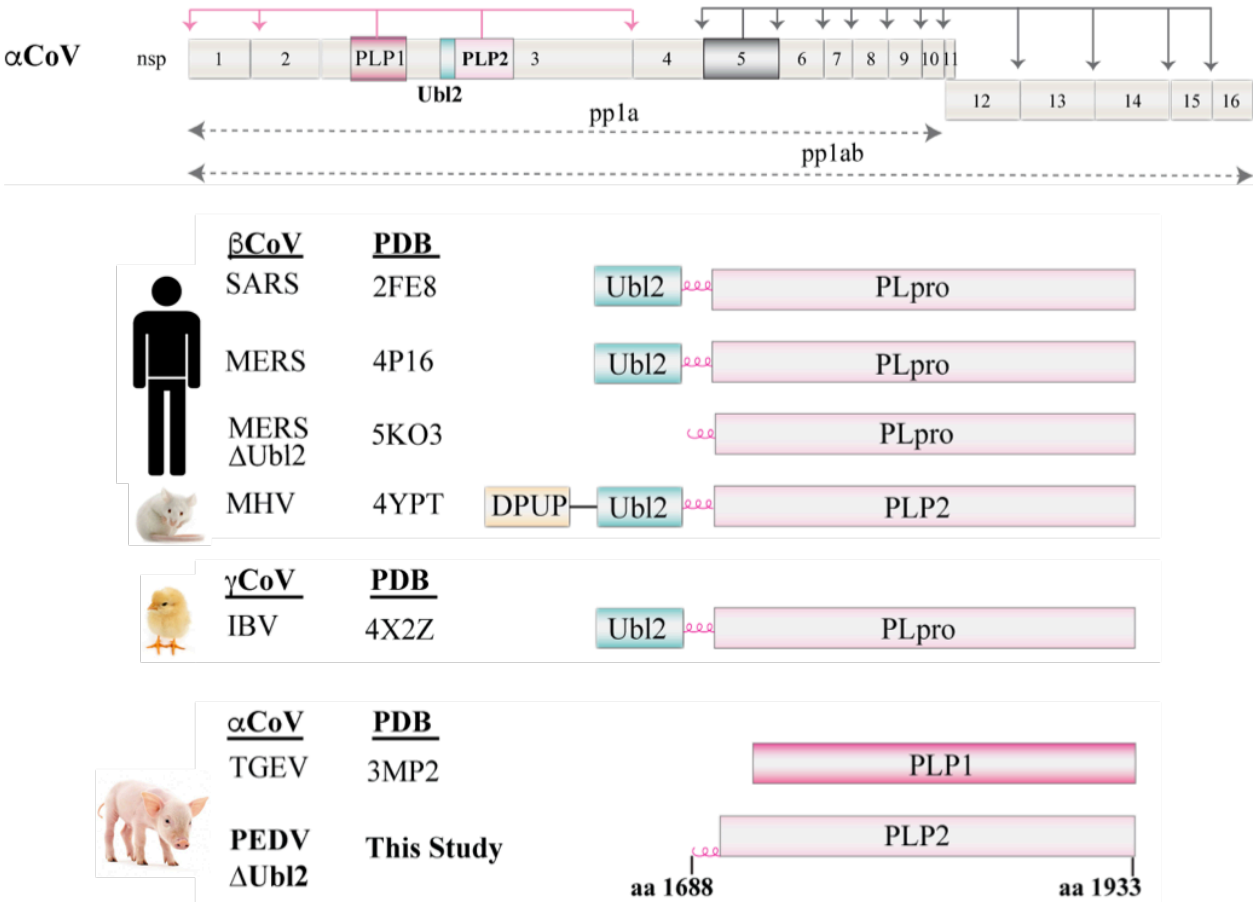


Figure 5.1 αCoV polyprotein organization and the summary of PLP crystal structures solved to date including PEDV PLP2 from this study

Top image – The αCoV viral polyprotein 1a and 1ab with sixteen nonstructural protein labeled. PLP1 and PLP2 are colored in dark and light pink in nsp3, respectively, and 3CLpro in nsp5, is colored in gray. The cleavage sites of these proteases are indicated by arrows and colored accordingly. The Ubl2 domain of nsp3 at the N-terminus of PLP2 is colored in green. Bottom image – Summary of the current PLP X-ray structures from different CoV genera. The PDB codes of the X-ray structures that were determined first for SARS-CoV, MERS-CoV, MHV, IBV and TGEV PLPs are reported. The design of PEDV PLP2 of nsp3 determined in this study is shown. The host (*i.e.* human, mouse, chicken, pig) each CoV is known to infect is shown at the left.

CoVs of the *Nidovirales* order are grouped into four genera: *alphacoronaviruses* (alpha-CoV), *betacoronaviruses* (beta-CoV), *gammacoronaviruses* (gamma-CoV), and *deltacoronavirus* (delta-CoV) (Figure 5.1). While the activities of PLPs have been investigated previously, *in vitro*

characterization has been primarily focused on PLPs from the *betacoronavirus* genus of CoVs. Therefore, there are less structural and functional information that fully characterizes different genera of CoV PLPs. The first X-ray structure of a PLP was solved in 2006 of beta-CoV SARS-CoV PLpro, which resembled the structure of deubiquitinating (DUBs) enzymes of the ubiquitin-specific protease (USP) family [63]. Following this study, other beta-CoVs have been solved: MERS-CoV PLpro [82] in 2014 and MHV PLP2 [54] in 2015. These beta-CoV PLPs adopt a similar catalytic core and tertiary fold to the closely related SARS-CoV PLpro counterpart with the thumb-palm-fingers architecture. Beyond that, only one structure is available for gamma-CoV infectious bronchitis virus (IBV) PLpro [69], which was solved in 2015, and to our knowledge, no αCoV PLP2 has been solved to date. The only αCoV PLP that has been solved is transmissible gastroenteritis virus (TGEV) PLP1 in 2010, which can cleave Ub-chains in biochemical assays, similar to PLP2 paralogs from other CoVs, and recently was reported to have IFN antagonism activity [52].

Here, we focus our efforts on αCoV PLP2s from feline infectious peritonitis virus (FIPV) and porcine epidemic diarrhea virus (PEDV). FIPV is a lethal mutant virus of the avirulent and frequent feline enteric CoV strain that is common to cats [24]. Mutations present in the RNA genome of FIPV enable the virus to efficiently replicate in macrophages, a key tropism that leads to the development of the fatal and incurable FIP disease [24, 146]. FIP affects approximately 5% of cats infected with feline CoV in multi-cat households [23]. There have been exciting reports that have focused on the development of inhibitors targeting FIPV 3CLpro [147-149], but no studies to date have yet characterized the function of FIPV PLP2 in viral polyprotein processing, deISG or DUB activities, or even IFN antagonism activity. From the genomic analysis of FIPV WSU-79/1146 strain, the amino acid positions of PLP1 and PLP2 of the FIPV polyprotein 1ab have been predicted [150]. The phylogenetic analysis supports that the non-structural proteins of

FIPV are closely related to TGEV and to a lesser extended, αCoV human NL63, while MHV and IBV are more distinct homologs [150].

Studies on human CoVs PLPs, SARS-CoV [75, 76, 79, 84] and NL63 [79], as well as MHV PLP2 [78, 86] have shown that these coronaviral proteases inhibit the induction of IFN-Is and counteract the host innate immune system, but the precise mechanism of PLP-mediated antagonism are still under investigation [49]. So far, CoV PLPs have been reported to disrupt IFN regulatory factor 3 (IRF3) and NFκB signaling pathways [76] by directly interacting and/or deubiquitinating key regulatory components, such as scaffolding protein STING (simulator of interferon genes), required for downstream signaling of IFN-β. For example, expression of NL63 PLP2-TM [84] and SARS-CoV PLP2-TM [88] (PLP2 with an intact transmembrane domain of nsp3) reduced ubiquitination of key upstream signaling molecules of the IFN-β gene, such as the pattern recognition receptor RIG-I (retinoic acid inducible gene I), STING, IRF3, and TBK-1 serine/threonine kinase. One mechanism proposed by Chen et al. is that SARS-CoV PLpro directly interacts with STING-TRAF3-TBK1 signaling complex and thereby disrupts the phosphorylation and dimerization of IRF3 for translocation into the nucleus to produce IFN-β [88]. Frieman et al. showed that NL63 PLP2, but not MHV PLP2 was able to block the IRF3 and NFκB signaling pathways, suggesting that the mechanism of IFN antagonism may differ across species of CoVs [76]. Other reports, in contrast, provide evidence to MHV PLP2 targeting TBK-1 [78] and deubiquitinating IRF3 [86] to antagonize IFN production. Collectively studies have supported that disrupting the DUB activity of PLPs could potentially reduce the ability of PLP to act as an IFN antagonist.

Although PEDV PLP2 has not been characterized *in vitro* like the previously mentioned FIPV PLP2, Xing et al. demonstrated that PEDV PLP2 can act as an IFN antagonist as well as a viral DUB in cells [70]. Specifically, PEDV PLP2 was found to interact and deubiquitinate RIG-I

and STING in cells and inhibit IFN-β expression mediated by RIG-I and STING similar to HCoV-NL63 PLP2 [70]. In contrast to what has been reported with HCoV-NL63 PLP2 [79], the enzymatic activity of PEDV PLP2 is essential for IFN antagonism activity, as mutations of PEDV PLP2 catalytic residues (C1729A, H1888A, and D1901A) no longer inhibited IFN-β expression to the extent of the WT. Since these residues are also necessary for PEDV PLP2 to deubiquitinate RIG-I and STING, we hypothesize that selectively disrupting the DUB activity would similarly remove host cell antagonism of PLP2.

In this study, we investigate the structure and function of FIPV and PEDV PLP2 *in vitro*. Our results show that αCoV FIPV and PEDV PLP2 prefer Ub to ISG15 and are more efficient DUBs rather than deISG enzymes. We reveal the first structure of αCoV PEDV PLP2 and generate a model of the Ub-bound structures to predict how both FIPV and PEDV PLP2 interact with Ub. Our study unveils a potential 'hot spot' to selectively disrupt the DUB activity of PLPs across clades of CoVs for live-attenuated vaccine design.

5.2 Material and Methods

5.2.1 Expression and Purification of FIPV PLP2 with the Ubl2 domain workflow 1

The FIPV PLP2 gene coding for residues 1488-1811 of the FIPV polyprotein 1ab (pp1ab), which includes the N-terminal ubiquitin-like 2 (Ubl2) domain followed by the catalytic core of PLP2, was synthesized and cloned into pET11a expression vector by BioBasic, Inc. The resulting FIPV PLP2 construct containing a N-terminal His8-tag was transformed into *E. coli.* BL21 (DE3) cells and expressed by autoinduction. A 100 mL LB starter culture (in a 1 L Erlenmeyer flask) containing 100 μg/mL carbenicillin was grown for 16 hours at 37 °C at 200 rpm using the ATR Biotech Multitron HT Infors Dual-Stack Incubator-Shaker shaker until the OD600 reached ~4. Cultures were expanded in 500 mL super broth containing 100 μg/mL carb by inoculating with

12.5 mL overnight culture at 25°C for 25 hours at 127 rpm using the New Brunswick innova 44. Four 500 mL cultures were combined and pelleted by centrifugation (2900 x g, 4 °C, 22 minutes), and the pellet weighing 15.5 g stored in the -80 °C. Frozen cells were thawed and resuspended with 80 mL lysis buffer (5 mL lysis buffer per 1 g of cells). Lysis buffer contained buffer A (25 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole, 5 mM BME, and 10% glycerol) plus 0.25 mg/ml lysozyme and 10 μg/ml DNase I. After resuspension, cells were sonicated using a Branson Digital Sonifier (60% amplitude; 10 minutes, 6.6 s pulses, 9.9 s delays) and then clarified by centrifugation (24,700 x g, 4 °C, 50 minutes).

The clarified lysate was filtered with a 0.45 μm syringe filter before loading onto a Ni^{2+} charged 5 mL His-Trap FF column (GE Healthcare) that was equilibrated buffer A at 2.5 mL/min. Weakly bound proteins were then washed with at least 5 CV of 3% buffer B at 3.0 mL/min until the absorbance at 280 nm flat-lined. Buffer B contained 25 mM Tris, pH 8, 500 mM NaCl, 500 mM imidazole, 5 mM BME, and 10% glycerol. Tagged FIPV PLP2 was eluted with a linear gradient (3% to 100%) at a flow-rate of 3 mL/min over 30 mL, collecting 3 mL fractions. A peak at 20% buffer B was observed (~1400 mAU) and the gradient was held briefly for 12 mL. A second linear gradient (20% to 100%) was set at 3 mL/min for 21 mL and once the peak returned to base-line the column was held at 100% for cleaning. Fractions were pooled (30 mg) based on activity (ranging from 1000 – 4200 μM/min/mg) and purity by SDS-PAGE analysis and digested with TEV protease at a 1:12.5 ratio (1 mg TEV: 12.5 mg PLP2) while dialyzing against 1 L buffer C (25 mM Tris, pH 8, 100 mM NaCl, 5 mM BME, and 10% glycerol) at 4 °C for ~64 hours. After dialysis, the sample was loaded onto the same column at 2.5 mL/min collecting untagged PLP2 in the flow-through. The column was then washed with 5 CV of buffer C until flat-lined, and both the wash and flow-through (12 mg) was pooled and concentrated to 2 mL using Amicon Ultrafiltration Centrifugal device (10 kDa MW cutoff).

In order to separate out a higher molecular weight contaminating protein, the sample was loaded onto a HiLoad 26/60 Superdex 75 column (GE Healthcare) equilibrated with buffer C containing 10 mM DTT as the reducing agent. FIPV PLP2 was eluted with 325 mL of buffer C at 1 mL/min flow-rate. A peak at ~150 mL elution volume was pooled based on SDS-PAGE and activity analysis. Specific activities of the S75 fractions ranged from 4700 – 5300 μM/min/mg. The pool was concentrated to 200 μL and 4 mg/mL aliquots were stored for kinetic assays and protein samples were prepared for crystallization.

5.2.2 Expression and Purification of FIPV PLP2 with the Ubl2 domain workflow 2

Due to lower protein yields when workflow 1 was repeated, FIPV PLP2-Ubl2 was purified using a second protocol that omitted the size exclusion step and added a strong anion exchange column. Expression of the protein was exactly as described in 4.2.1, but six 500 mL cultures were combined, and the 30 g pellet was used for purification. A 50 μg/mL carbenicillin final concentration was used in the 500 mL super broth cultures. The pellet was resuspended with 145 mL buffer A, and cells were sonicated, clarified, and the clarified lysate was loaded on to the 5 mL Ni^{2+} His-Trap FF column (GE Healthcare) as above. The same wash procedure was followed but tagged FIPV PLP2 was eluted with linear gradient (3 % to 100% buffer B) across 110 mL at 2.5 mL/min flow-rate without holding and a similar peak was observed at ~1400 mAU. Active fractions (ranging from 1500 – 3000 μM/min/mg) were combined and digested with TEV protease at a 1:10 ratio in 1 L buffer C at 4 °C overnight.

To separate out a higher MW contaminant, the 20 mL sample (25 mg) was buffer exchanged into MonoQ buffer A (25 mM Tris, pH 8, 10 mM NaCl, 10% glycerol, 5 mM DTT) and loaded onto a strong anion exchange MonoQ 10/100 GL (GE healthcare) at 0.8 mL/min flow-rate, which was pre-equilibrated with MonoQ buffer A. The column was washed with 5% MonoQ buffer B at 1.5 mL/min while collecting 3 mL fractions. MonoQ buffer B contained 25 mM Tris,

pH 8, 1 M NaCl, 10% glycerol, 5 mM DTT. Untagged FIPV PLP2 was active in the wash with specific activities of ~4000 μM/min/mg, which were pooled. The pool was concentrated and exchanged into buffer C (7.5 mL sample; 10 mg total protein) and passed over the Ni^{2+} His-Trap FF column (GE Healthcare) again to remove lower molecular weight contaminants. The flow-through was collected along with a 10 mL wash of buffer C, which were pooled and concentrated to 2 mL. Finally, the sample was exchanged into buffer C with 10 mM DTT using a PD-10 desalting column (GE Healthcare) and concentrated to 100 μL for crystallization.

5.2.3 Specific activity assays

Throughout purification FIPV PLP2-Ubl2 was assayed for activity using 250 nM Ub-AMC (LifeSensors) at a final protein concentration of 0.125 μg/mg or 3.2 nM PLP2, following a similar procedure described in 2.2.2. The assay buffer contained 50 mM HEPES, pH 8, 0.1 mg/mL BSA, and 2.5 mM DTT.

5.2.4 Attempts at screening FIPV PLP2-Ubl2 for Crystallization

Fresh FIPV PLP2 purified from workflow 1 and 2 were both screened for initial crystallization conditions using Mosquito®Crystal liquid handling robot (TTP Labtech) with Qiagen, Hampton, and Midwest Center for Structural Genomics (MCSG) screens-1-4 from Anatrace. Protein prepared from workflow 1 were screened at different protein concentrations (10 mg/mL, 5 mg/mL, and 2 mg/mL. Protein concentrations were screened at a 1:1 ratio (200 nL total drop size) for the lower concentrations and a 1:2 ratio of protein to reservoir (300 nL total drop size) was used for the 10 mg/mL concertation in 96-3 well sitting drop vapor diffusion plates (Greiner CrystalQuick crystallization plate). Another protein preparation from workflow 2 was screened using only MCSG-2 using different protein concentrations in the subwells. The first subwell contained 150 nL 10 mg/mL FIPV Ubl2-PLP2 at a 1:1 ratio (300 nL total drop size), the

second contained 100 nL 10 mg/mL FIPV Ubl2-PLP2 at a 1:2 ratio (300 nL total dropsize), and the third contained 200 nL 6 mg/mL FIPV Ubl2-PLP2 at a 2:1 ratio (300 nL total dropsize).

5.2.5 Expression and Purification of FIPV and PEDV PLP2s

The genes encoding for FIPV $PLP2_{1558-1794}$ and PEDV $PLP2_{1688-1933}$, were synthesized and cloned into pET11a expression vector by SynbioTech, Inc. with a His_8-tag at the N-terminus. Each construct was expressed in *E. coli* BL21(DE3) by autoinduction using a similar protocol described in [103] at a 1-L scale. For PEDV PLP2, a 500 mL scale autoinduction was also performed and purified with the 1-L scale culture side-by-side. Pellets weighed approximately 7-8 g (5 g for 500-mL culture). All the buffer components for purification are previously described in [103]. On the day of purification, the pellets were resuspended in ~50 mL buffer A with dissolved flakes of DNase I and lysozyme, and the cells were lysed for a 8 minute protocol (5 minute protocol for 5 g pellet) using a Branson Digital Sonifier (65% amplitude with 6.6 s pulses and 9.9 s intervals). The lysed cells were clarified by centrifugation (26,203 x g, 4°C, 25 min).

The clarified lysate was loaded dropwise on a Bio-Rad Econo-Pac® gravity column packed with a 3 mL HisPur™ Ni-NTA resin (Thermo Fisher Scientific), which was equilibrated with buffer A. Unbound proteins were washed from the column with 25 mL buffer A and then 25 mL 10% buffer B. His-PLP2 was then eluted with 25 mL of 25% buffer B and then 25 mL of 50% buffer B. All washes and elutions were tested for absorbance at 280 nm using a BioTek Take3™ Multi-Volume Plate. The 25% and 50% buffer B samples (50 mL total volume) were pooled based on the presence of the his-PLP2 band at the appropriate molecular weight by SDS-PAGE analysis (28.5 kDa for his-FIPV PLP2 and 29.5 kDa for his-PEDV PLP2). Approximately 30-40 mg of his-tagged protein was obtained from the 7-8 g pellets when both 25% and 50% buffer B samples were pooled and approximately 20 mg were obtained from the 5 g PEDV PLP2 pellet. TEV protease at approximately a 1:10 ratio was added to each pool for cleavage of the N-terminal his_8-tag on PLP2,

and the sample was transferred to dialysis tubing (10 kDa MW cutoff) before it was submerged in 1 L buffer A for 16 hours at 4°C.

The dialyzed sample was again loaded over the same gravity column equilibrated with buffer A, this time collecting the flow-through sample with untagged PLP2. For PEDV PLP2, both the untagged PLP2 samples were combined from the two purifications for size exclusion. The untagged PLP2 samples were concentrated to 1 mL using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MW cutoff) and loaded onto a HiLoad 26/60 Superdex 75 column (GE Healthcare) equilibrated with S75 buffer. PLP2 was eluted at 1 mL/min flow rate collecting 5 mL fractions. Active FIPV PLP2 was pooled and then concentrated to ~40 mg/mL while active PEDV PLP2 was pooled and then concentrated to 24 mg/mL before protein was diluted to varying concentrations for the crystallization screening and 1-2 mg/mL samples for kinetic experiments. The S75 final pool of FIPV PLP2 is shown in Figure 5.3.B and yielded ~24 mg pure protein from 1 L cell culture, and the S75 final pool of PEDV PLP2 **(**Figure 5.5.C) yielded ~22 mg pure protein from 1.5 L cell culture. Purified protein was aliquoted in 50-100 µL aliquots and stored at -80°C until further use.

Throughout purification, activity was assayed using 200 nM Ub-AMC with either 6 nM FIPV PLP2 and 5 nM PEDV PLP2 at a 30 µL format in 96-well half-area black microplate from Corning. Fluorescence was monitored using a BioTEK Synergy H1 multimode at ambient temperature in assay buffer containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, 5 mM DTT. It is important to note that both enzymes were very active and lowering the enzyme concentration to 1-3 nM may be beneficial for future specific activity assays.

5.2.6 Steady-state kinetics of FIPV and PEDV PLP2

The kinetic parameters were determined for FIPV PLP2 and PEDV PLP2 hydrolysis of Ub-AMC (Boston Biochem), human ISG15-AMC (Boston Biochem/R&D Systems), and Z-RLRGG-AMC (Bachem) using the same plate-type and assay volume described in [103].

Reactions were performed in assay buffer containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, 5 mM DTT, and fluorescence intensities were monitored by the release of the AMC fluorophore using a BioTEK Synergy H1 multimode microplate reader at 25°C. The Ub-AMC assay was initiated by the addition 0.390 nM PLP2 final concentration across serial dilutions of Ub-AMC between 0.017 – 17.6 μM. The ISG15-AMC assay were initiated with either 0.313 μM FIPV PLP2 or 0.156 μM PEDV PLP2 across serial dilutions of ISG15-AMC between 0.6 – 19.2 μM. The ISG15-AMC assay was performed with 0.5 μM PLP2 across 1.2 – 19.2 μM ISG15-AMC concentrations for FIPV Ubl2-PLP2 enzyme. The RLRGG-AMC was initiated with 0.75 μM PLP2 across serial dilutions of RLRGG-AMC between 1.56 – 50 μM. Due to low activity exhibited by FIPV Ubl2-PLP2, RLRGG-AMC activity was only tested at 50 μM substrate using 0.5 μM PLP2. If saturation was met, these data was fit to classic Michaelis-Menten equation in SigmaPlot (version 12; Systat Software, Inc.). In nonsaturing conditions, data was fit to a linear regression equation in GraphPAD Prism6 for the determination of the apparent $k_{cat}/K_{m.}$. The kinetic parameters are reported in Table 5.3 and also compared with other α- and βCoV PLPs in Table 5.4.

5.2.7 Expression and purification of free feline and porcine ISG15

Genes pET11a-His8-TEVp-feline-ISG15$_{1-157}$-Cys78Ser and pET11a-His8-TEVp-porcine-ISG15$_{1-156}$ were synthesized from BioBasic, Inc and SynbioTech, Inc., respectively (UniProt protein entries B3IX46 and B2ZDZ2). To aid in protein stability, the Cys78Ser mutation was engineered into the hinge region of feline ISG15 (fISG15). Porcine ISG15 (pISG15) does not contain the Cys as in other ISG15 species. Each construct pET11a-His8-TEVp-feline-ISG15$_{1-157}$-Cys78Ser and pET11a-His8-TEVp-porcine-ISG15$_{1-156}$ was transformed into *E. coli* BL21 (DE3) cells for protein expression, and a 50 mL LB starter culture (in a 250 mL flask) containing 100 μg/mL carbenicillin was inoculated with a colony and grown for 16 hours at 37 °C at 200 rpm

using the ATR Biotech Multitron HT Infors Dual-Stack Incubator, and 10 mL was then transferred to 1 L autoinduction media supplemented with 50 μg/mL carbenicillin and grown for 24 hours at 25 ºC shaking at 141 rpm using the New Brunswick innova 44. Two 1 L cultures were combined and harvested by centrifugation (3011 x g, 20 min, 4 ºC). Pellets weighing approximately 12-15 g were obtained and froze in the -80 °C until the day of purification. Pellets were resuspended with ~50-75 mL lysis buffer (5 mL per g of pellet), which contained buffer A (25 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole, 5 mM BME, and 10% glycerol) supplemented with 0.25 mg/ml lysozyme and a pinch of DNase I, and then sonicated using a Branson Digital Sonifier at 60% amplitude for 10 minutes with 5 s pulses and 9 s delays.

The clarified lysate was obtained by removing cell debris by centrifugation (27,200 x g, 4 °C, 45 minutes), which was subsequently loaded onto a charged Ni^{2+} 5 ml His-Trap FF column (GE Healthcare) equilibrated with buffer A at 2.5 mL/min flow-rate. After washing unbound proteins with 5-10 column volumes (30-40 mL) buffer A, ISG15 was eluted with a 33.3 min linear gradient (0% to 50%) of buffer B (25 mM Tris, pH 8, 500 mM NaCl, 500 mM imidazole, 5 mM BME, and 10% glycerol) at a flow rate of 2.5 ml/min, and 5 mL fractions were collected. The eluted fractions (25-30 mL total volume) that displayed a UV signal at 280 nm and contained a 20 kDa band as analyzed by SDS-PAGE were pooled. The His_8-tag was then removed by addition of tobacco etch virus (TEV) protease at a 1:20 ratio after dialyzing against 1-2 L buffer C (25 mM Tris, pH 8, 100 mM NaCl, 5 mM BME, and 10% glycerol) for 4 °C overnight.

After dialysis, the sample was passed over the same His-trap column equilibrated with buffer C, and the flow-through was collected. For pISG15, untagged protein was eluted in the wash with 20 mL of 3% buffer B. The untagged free ISG15 sample was concentrated using an Amicon Ultrafiltration Centrifugal device (10 kDa MW cutoff) to ~2 mL and loaded onto a HiLoad 26/60 Superdex 75 column (GE Healthcare) equilibrated with buffer C containing 10 mM DTT (2 mM

DTT) in place of the BME. ISG15 was eluted with 365 mL of buffer C at 1 mL/min flow-rate, and the elution profile is shown in Figure 5.7. Fractions corresponding to the peak at ~195 mL were collected. Approximately 14 mg pISG15 was purified from a 2 L total cell culture, and a similar yield (10 mg) was obtained for fISG15 after SEC. Feline ISG15 was concentrated to 28 mg/mL and stored in 30 µL aliquots while porcine ISG15 was concentrated to 14 mg/mL and stored in 100 µL aliquots.

Throughout purification fISG15 samples were assayed for activity using 200 nM Ub-Rho with 5.0 µL of fractions or pools in a 30 µL format in 96-well half-area black microplate from Corning, and fluorescence was monitored using a BioTEK Synergy H1 multimode microplate reader at 25 °C in assay buffer containing 50 mM HEPES, pH 8, 0.1 mg/mL BSA, and 2.5 mM DTT. pISG15 samples were assayed using 200 nM Ub-AMC for activity with 100 µM pISG15 in a 30 µL format in 96-well half-area black microplate from Corning. Fluorescence was monitored using a CLARIOstar BMG Labtech microplate reader at 25 °C in assay buffer containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, and 5 mM DTT.

5.2.8 Inhibition assays with free feline, porcine ISG15, and compound 3e

Inhibition assays of FIPV and PEDV PLP2 were performed in triplicate at a 30 µL scale in assay conditions containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, and 5 mM DTT in a 96-well half area black plate (Corning). Fluorescence was measured using BioTEK Synergy H1 multimode microplate reader at 25 °C. Base-line fluorescence of 200 nM Ub-AMC was monitored along with serial dilutions of feline or porcine ISG15 from 1.6 – 100 µM before initiating the reaction with 3.5 nM final PLP2. Compound 3e was screened at 100 µM and inhibition of SARS PLpro (final enzyme concentration of 64 nM) was determined as a positive control. The % inhibition calculation was done as described in 2.2.7 using equation 2.5. The average of the positive control (reactions without inhibitor) was used in all calculations rather than within the

replicate measurement. The IC_{50} value for each inhibitor was determined if applicable based on the Michaelis-Menten equation 2.7. Assuming competitive inhibition, the approximate inhibitor binding affinity or K_i was calculated using equation 5.1:

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}} \tag{5.1}$$

5.2.9 Synthesis of Ub-PA and feline and porcine ISG15-PA

The gene encoding for feline and porcine ISG15 were synthesized cloned into pTYB2 plasmid by BioBasic. Ub-PA, fISG15-PA, and pISG15 propargylamine (ISG15-PA) were synthesized using the chitin intein-fusion protocol described by Wilkinson et al. [134]. Plasmid of either pTYB2-fISG15$_{1–154}$Cys78Ser or pTYB2-pISG15$_{1–153}$ was expressed in *E. Coli* BL21(DE3) cells in 1 L autoinduction media for 24 hours at 25 ºC before cells were harvested by centrifugation (3011 x g, 20 min, 4 ºC). The pTYB2-Ub$_{1-75}$ plasmid was expressed in *E. Coli* BL21(DE3) + RIPL cells. Pellets were frozen in -80 ºC and then thawed and resuspended in 50 mL chitin-column buffer (50 mM MES pH 6.0, 350 mM sodium acetate), supplemented with 50 µg/mL DNase and 1 mM $MgCl_2$. The resuspended cells were lysed using a Branson Digital Sonifer by sonicating at a 60% amplitude for 15 minutes with 5 s pulses and 9 s delays and then centrifuged to clarify (28,960 x g, 30 minutes, 4 ºC).

The clarified lysate (50-60 mL) was combined with 20 mL chitin resin (New England Biolabs), which was equilibrated with chitin-column buffer using vacuum filtration and allowed to incubate with gentle stirring for at least 2 hours at 4 ºC. Unbound proteins were washed using vacuum filtration with ten volumes of 50 mL column buffer. The resin was resuspended in 50 mL cold reaction buffer (50 mM MES pH 6.0, 350 mM sodium acetate, 120 mM sodium 2-mercaptoethanesulfonate (MESNa)), transferred to a beaker, and allowed to incubate at 4 ºC overnight with gentle stirring.

The next day, newly formed Ub/ISG15-MESNa was separated from the resin using vacuum filtration. The resin was then resuspended and washed with two 50 mL portions of column buffer to recover any residual Ub/ISG15-MESNa. The elution and washes were combined and concentrated to approximately 20 mL using an Amicon Ultra-15 Centrifugal Filter Unit (either 3 kDa or 10 kDa MW cutoff for Ub-MESNa and ISG15-MESNa, respectively), and propargylamine HCl (Sigma-Aldrich) was added to a final concentration of 250 mM. The pH was adjusted to pH 10 by adding 2-3 mL of 2 M NaOH dropwise. To generate Ub/ISG15-PA, the reaction was allowed to precede overnight at room temperature, and the final reaction products were centrifuged (28,960 x g, 10 min, 4 ºC) to remove any precipitation. The resulting supernatant was concentrated to ~2 mL and buffer exchanged into storage buffer using a PD-10 desalting column (GE Healthcare). PA probes were aliquoted and flash frozen with liquid N_2 before storage in -80 ºC. fISG15-PA was stored in 50 mM Tris, pH 7.5, 1 mM DTT while pISG15-PA was stored in S75 storage buffer described in [103]. Approximately ~600 μM stocks of f- and pISG15-PA were used in probe reactivity assays.

5.2.10 Ub- and ISG15-PA probe reactivity assays

Probe reactivity assays were set-up at a 10 μL scale with pure proteins and PA probes. The reactivity of FIPV and PEDV PLP2 WT was evaluated with both Ub-PA and ISG15-PA probes at varying molar ratios. Ub-PA and either f- or pISG15-PA were diluted in 10 mM Tris, 100 mM NaCl, 5% glycerol containing fresh 10 mM DTT at final concentrations ranging from 5-25 μM. Each reaction was initiated with 5 μM PLP2 and allowed to react for 1 hour at room temperature before quenching with 5X loading buffer (250 mM Tris-HCl pH 6.8, 50% glycerol, 10% SDS, 0.02% bromophenyl blue, and fresh 400 mM DTT). The reactivity of PLP2 WT and mutants were evaluated at single molar ratios of 5 μM PLpro and either 25 μM Ub-PA or ISG15-PA (1:5 molar ratio), which were quenched at different time points (2, 5, 30 min) with loading buffer. The entire

reaction was loaded onto 4-20% Mini-PROTEAN® TGX™ precast SDS-PAGE gels from BioRad, and proteins were visualized with Coomassie.

5.2.11 Ubiquitin-chain cleavage assays

Different Ub_2 chains (3.9 μM Ub_2 final), including K6, K11, K29, K33, K48, and K63, and linear Ub_2, were incubated with either 100 nM FIPV or PEDV PLP2 in assay buffer 50 mM HEPES, 0.1 mg/mL BSA, 5 mM DTT for 2 hours and analyzed by Coomassie stain (no boiling) by loading 0.5 μg of Ub_2 chain on a 4-12% Bis-Tris NuPAGE gel. Reactions were done at a 7.5 μL scale. The negative control was also loaded, which only contained Ub_2. For assays with mutants, 3.9 μM K48-Ub_2 were incubated with two different enzyme concentrations, 0.1 μM and 2 μM. Time-dependent cleavage assays were performed by incubating K48- and K63-Ub_4 (2 μM Ub_4 final) with 20 nM FIPV or PEDV PLP2 in assay buffer and was allowed to react for 5, 10, 30, 60, 120 min before quenched with 4X NuPAGE® lithium dodecyl sulfate, pH 8.4 (LDS) sample buffer. Time point 0 min only contained Ub_4 substrate and not enzyme. Cleavage products were boiled for 1 min at 95°C and then visualized with Coomassie stain by loading 0.5 μg of Ub_4 chain on a 4-12% Bis-Tris NuPAGE® gel. All assays with FIPV PLP2-Ubl2 were done with 100 nM enzyme in assay buffer 50 mM HEPES, pH 8, 0.1 mg/mL BSA, 100 mM NaCl, 5 mM DTT; Ub_2 chains and products were ran on a on a 4-12% Bis-Tris NuPAGE gel, and Ub_4 chains and products were ran on a 4-20% Mini-PROTEAN® TGX™ precast SDS-PAGE gels from BioRad. The same procedure above was used except the latter samples were quenched with 2.5 μL of 5X loading buffer (200 mM Tris-HCl pH 6.8, 20% glycerol, 10% SDS, 0.02% bromophenyl blue, and fresh 100 mM DTT) and boiled for 3 min at 95°C.

5.2.12 Crystallization of PEDV PLP2

The freshly purified PEDV PLP2 was used to screen for initial crystallization conditions at three protein concentrations: 5 mg/mL, 10 mg/mL, 15 mg/mL. The Mosquito®Crystal liquid

handling robot (TTP Labtech) dispensed 100 nL of purified PEDV PLP2 to 100 nL of reservoir solution in 96-3 well sitting drop vapor diffusion plates (Greiner CrystalQuick crystallization plate). The 200 nL drops were stored and imaged daily in the Rigaku Minstrel® HT plate hotel at 20°C. FIPV PLP2 protein was also screened, but regrettably no diffractable quality crystals were obtained. After five days, an initial crystal hit at 15 mg/mL PEDV PLP2 was observed from the Anions suite by Qiagen, which contained 0.1 M HEPES, pH 7.5 and 1.2 M K/Na tartrate reservoir. The crystals resembled a 'sea urchin' or fuzzy microcrystal. Optimization was done at 4°C in a 24-well sitting drop format with drops containing 2 µL 15 mg/mL PEDV PLP2 and 2 µL reservoir with varying amounts of K/Na tartrate from the initial hit. Thin needles clusters were obtained after 11 days in reservoir containing 0.1 M HEPES, pH 7.5 and 1.3 M K/Na tartrate. After 20 days, two separate crystal rods, measuring roughly ~200 µm long by 15 µm thick, were observed in the same well adjacent to the thin needle clusters (Figure 5.5.D). These crystals rods were harvested with a 0.1-0.2 mm nylon loop, briefly soaked in cryoprotectant, which contained reservoir solution supplemented with 20% glycerol, and then plunged in liquid nitrogen. Crystals were stored in SPINE pucks before shipment to the Advanced Photon source (APS) at Argonne National Laboratory (ANL). Two data-sets were collected from both rod-shaped crystals at the automated MAR-21-ID-F beam-line in the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) sector.

The structure of PLP1 from alpha-CoV transmissible gastroenteritis virus (TGEV), PDB code 3MP2 from [51], was used as a search model for molecular replacement [136]. The PHENIX AutoBuild wizard was used for initial model building and structure refinement [151], and the structure was finished using manual building and Phenix.refine [110]. After using PHENIX AutoBuild wizard, the structure refined with a R_{work} of 21.85% and R_{free} of 26.74%. The final refined structural model statistics are reported in Table 5.5.

5.2.13 Expression and Purification of FIPV and PEDV PLP2s mutants

A series of site-directed mutants were generated using the QuickChange™ site-directed mutagenesis system [137] from Stratagene (La Jolla, CA). pET11a-his8-TEVp-FIPV-PLP2$_{1558\text{-}1794}$ and pET11a-his8-TEVp- PEDV-PLP2$_{1688\text{-}1933}$ WT plasmids were used as a template and primers were designed (Table 5.1) and synthesized using Integrated DNA Technologies (IDT). These were done following the same PCR cycling procedure as described in 4.2.5. Correct mutations in the FIPV PLP2 and PEDV PLP2 genes were confirmed by Sanger-based sequencing through the Purdue Genomics Core Sequencing Facility.

Expression and purification of mutant enzymes were performed similar for the WT as described above. The only differences were that the expression volume was reduced to a 500-mL scale and the size-exclusion step was eliminated. From a 500-mL culture yield, 10-20 mg of PLP2 was recovered with the exception of E97R^{FIPV}, which yield was only 2 mg, and slightly less pure compared to the rest of the mutants (Figure 5.17). Proteins were flash frozen in S75 buffer as done with the WT and stored at a protein concentration ranging from 2-6 mg/mL PLP2.

Table 5.1 Primers used for site-directed FIPV and PEDV PLP2 ubiquitin binding mutants

Mutation	Forward Primer (5' – 3')	Reverse Primer (5' – 3')
FIPV PLP2		
E97R	GACGCACGTCTGACCCTGCATAAACTGGTCGATCTG	GTTTATGCAGGGTCAGACGTGCGTCGCCCGGTTGACCTTTG
L98R	GCAGAACGTACCCTGCATAAACTGGTCGATC	GTTTATGCAGGGTACGTTCTGCGTCGCCCGGTTG
H101R	CTGACCCTGCGTAAACTGGTCGATCTG	GACCAGTTTACGCAGGGTCAGTTCTGCGTC
K102A	CTGCATGCACTGGTCGATCTGATGAGCAGC	GACCAGTGCATGCAGGGTCAGTTCTGCGTC
K102E	CTGCATGAACTGGTCGATCTGATGAGCAGC	GACCAGTTCATGCAGGGTCAGTTCTGCGTC
A136K	GTTGTTAAAGCACCGCTGCTGGTTTGCGGTAC	GGTGCTTTAACAACCGGACCGGTAAAGGTTTC
A136R	GTTGTTCGTGCACCGCTGCTGGTTTG	GTGCACGAACAACCGGACCGGTAAAGGTTTC
PEDV PLP2		
E97R	GATAGCCGTAACGCCCTGAACATGCTGAGC	CGTTACGGCTATCGCTCGGCTGACCTTTATC
N98R	CGAACGCGCCCTGAACATGCTGAGCAAATAC	CATGTTCAGGGCGCGTTCGCTATCGCTC
N101R	CCTGCGCATGCTGAGCAAATACATCGTCC	CTCAGCATGCGCAGGGCGTTTTCGCTATC
M102R	CCTGAACCGTCTGAGCAAATACATCGTCC	GCTCAGACGGTTCAGGGCGTTTTCGCTATC

5.2.14 Activity screen of FIPV and PEDV PLP2 mutants

The catalytic turnover rates were determined for FIPV and PEDV PLP2 WT and mutants at one single substrate concentration of 50 μM RLRGG-AMC and 1 μM Ub-AMC. Ub-AMC and RLRGG-AMC assays were run at a 30 μL and 100 μL scale, respectively in either a half area and full area 96-well black plate (Corning) in assay buffer containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, and 5 mM DTT. Fluorescence was measured using BioTEK Synergy H1 multimode microplate reader at 25 °C for all assays except for the PEDV PLP2 peptide assays where fluorescence was monitored at 4.2 mm focal height using a CLARIOstar BMG Labtech microplate reader at 25 °C with a signal gain of 676. The extinction / emission wavelengths and associated bandwidths were used based on the fluorophore AMC (7-amino-4-methylcoumarin), which were 350-15 nm and 440-20 nm, respectively.

A fixed PLP2 concentration of 0.75 μM was used for the hydrolysis of the peptide substrate. The Ub-AMC assay was initiated with a lower final enzyme concentration of 0.39 nM for FIPV PLP2 WT, L98R^{FIPV}, K102A^{FIPV}, and K102E^{FIPV} as well as PEDV PLP2 WT and M102R^{PEDV}. Due to low substrate turnover, a higher concentration of 0.05 μM was used for E97R$^{FIPV/PEDV}$, N98R^{PEDV}, N101R^{PEDV}, A136K^{FIPV}, and A136R^{FIPV} and the highest of 0.156 μM was used for H101R^{FIPV}. The % relative turnover k_{cat} (min^{-1}) of each PLP2 was calculated and compared to the WT for analysis.

5.2.15 Temperature-dependent inactivation experiments

A series of 50 μL reactions at 7.5 μM PLP2 were prepared and allow to incubate in a water bath set to 30 °C and 37 °C for 5, 15, 30, and 60 minutes in assay buffer containing 50 mM HEPES, pH 7.5, 0.1 mg/mL BSA, and 5 mM DTT. The turnover number of FIPV PLP2 WT, H101R^{FIPV}, PEDV PLP2 WT, and N101R^{PEDV} following the incubation was assayed in duplicate with 50 μM RLRGG-AMC using a CLARIOstar BMG Labtech microplate reader at 25 °C at a final

concentration of 0.75 μM PLP2. The reaction at 0 minutes was incubated at 25 °C on the lab bench and used to normalize against the other time points. Data analysis was done by plotting the percent remaining activity as a function of incubation time.

5.3 Results/Discussion

5.3.1 Purification of FIPV PLP2 with the Ubl2 domain and failed crystallization attempts

In an effort to characterize αCoV FIPV PLP2 activity and structure, we first expressed and purified the FIPV PLP2 construct with the N-terminal ubiquitin-like 2 (Ubl2) domain, represented amino acid sequence 1488-1811 of the FIPV pp1ab from FIPV strain WSU-79/1146. This is the identical construct that was used in cellular studies by our collaborators in Dr. Susan Baker's lab at Loyola University. Initially, FIPV Ubl2-PLP2 was purified using a 3-step column procedure, which included a Ni His-trap column 1, Ni His-trap column 2 (after incubation with TEV protease), followed by SEC. The last SEC step was necessary to separate a 70 kDa contaminating protein (Figure 5.2.A) that was observed during the first affinity column, which is represented in Figure 5.2.D, lane 1-2. FIPV Ubl2-PLP2 was purified with a purification fold of 21, percent recovery of 14%, and the final specific activity value achieved was ~4900 μM/min/mg (Table 5.1). SDS-PAGE analysis of the final sample, shown in Figure 5.2.B, indicated >90% purity, which was an ideal sample for crystallization.

The S75 pool sample was screened for crystallization conditions, and several hits from the MCSG-2 suite was observed, which contained 10 mg/mL FIPV Ubl2-PLP2 at a 1:2 protein to reservoir ratio. Approximately 40 micron crystals grew after 24 hours at room temperature in the following condition: 0.2 M calcium acetate, 0.1 M sodium acetate, pH 4.5, and 30% (w/v) PEG-400 (Figure 5.2.C). Regrettably, all the protein from workflow 1 was used for optimization attempts that were unsuccessful and only produced precipitated protein at a 24-well format.

Table 5.2 Purification activity summary of FIPV Ubl2-PLP2$_{1488-1811}$ for workflow 1 & 2

Sample	Total Protein (mg)	Total Units[1]	Specific Activity (μM/min/mg)	Purification Fold	Yield (%)
Workflow 1[2]					
Lysate	1264	14.5	229	1.0	100
Ni His-trap Pool 2	11.8	0.76	1281	5.6	5.2
Superdex-75 Pool	8.5	2.06	4927	21.5	14.2
Workflow 2[3]					
Lysate	1726	12.66	147	1.0	100
Ni His-trap Pool 1	27.2	1.17	862	5.9	9.2
MonoQ Pool	10.3	2.14	4103	28.0	16.9
Ni His-trap Pool 2	3.5	0.42	2433	16.6	3.4

[1]Unit (μmol/min): μmol of products produced per minute from assay conditions – 0.250 μM Ub-AMC, 3.5 nM PLP2 in assay buffer containing 50 mM HEPES, pH 8, 0.1 mg/mL, 2.5 mM DTT
[2]Workflow 1: The amount of protein was generated from 2 L of cell culture.
[3]Workflow 2: Protein produced from 3 L of cell culture.

After six attempts to repeat workflow 1, only ~5 mg of protein was generated following the Ni His-trap 2 step, which was not enough sample to run SEC for crystal optimization. To troubleshoot this problem, the amount of 500 mL expression cultures was increased from 2 L to 3 L total cultures, and to separate out the contaminant a smaller 8 mL MonoQ anion exchange column was run instead of the 350 mL SEC (Figure 5.2.D, lane 4). At pH 8, FIPV Ubl2-PLP2 eluted during the wash step while the higher MW contaminant eluted at almost 1 M NaCl. To remove lower MW contaminants, such as TEVp, a final Ni His-trap 2 column was performed. The final FIPV Ubl2-PLP2 pool was purified by 17-fold, but only 3.4% was ultimately recovered. While the final pools were not analyzed in the same assay plate for direct comparison, the final specific activity of workflow 2 was 2-fold less than the workflow 1 suggesting the previous

workflow was a better purification procedure. A loss in specific activity was also observed between the MonoQ pool and the final Ni His-trap pool 2 for workflow 2 (Table 5.2).

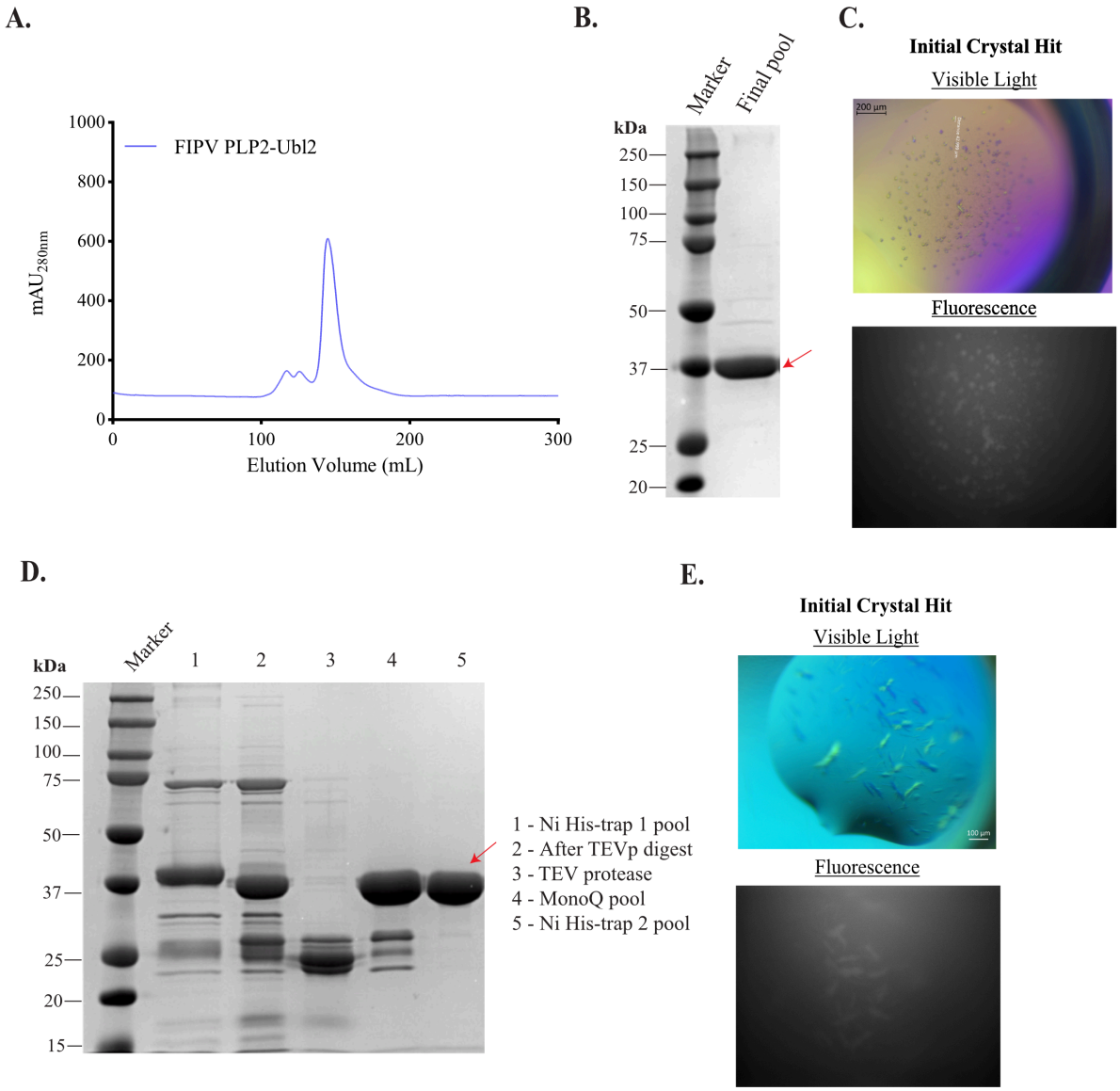


Figure 5.2 Purification and initial crystal hits of FIPV PLP2-Ubl2 using two workflows

(A) SEC chromatogram showing the elution profile of FIPV PLP2-Ubl2. (B) 12.5% SDS-PAGE gel of final pool from workflow 1. Protein was loaded at 5 µg (C) Top image – FIPV PLP2-Ubl2 crystals (~40-50 µm) under polarized light. Bottom image – Fluorescence of protein crystals after UV exposure. (D) 12.5% SDS-PAGE purification summary gel using workflow 2 with MonoQ column. Protein was loaded at 10 µg. (E) Visible light and fluorescent images of preliminary crystal hit from workflow 2 showing a thin needle-like morphology.

Due to the low yield, the final pool generated from workflow 2 was screened for crystallization with only the MCSG-2 screen, where most of the other initial crystal hits were

previously observed. The best initial hit was grown in condition containing 1.6 M ammonium sulfate, 0.1 M MES, pH 6.5, and 10% (v/v) dioxane using a 2:1 ratio of 6 mg/mL FIPV Ubl2-PLP2 to reservoir (Figure 5.2.E). These crystals were very thin and resulted in large needle clusters. Unfortunately, when optimizing this condition in a 24-well format, only clear drops were observed. Another crystal hit from the condition 0.1 M Tris, pH 8.5, 3 M NaCl was also sought for optimization. Although successful, the optimization resulted in only small microcrystals and finely clustered needles that resembled spherulites. One explanation as to why there were problems in reproducing the FIPV Ubl2-PLP2 crystals were inconsistencies in the drop volumes dispensed by the Mosquito® robot within the software during this time. It was therefore difficult to transition from a small scale to a large-scale drop size when the crystal hits were produced from unknown protein to reservoir ratios.

5.3.2 Purification of FIPV PLP2 without the Ubl2 domain and attempts at crystallization

To improve protein yield for crystallization, we redesigned the FIPV PLP2 construct by mimicking our MERS-CoV PLpro construct choosing residues of the FIPV polyprotein 1ab that correspond to only the PLP catalytic core [103]. FIPV PLP2 corresponds to amino acid sequences 1558-1794 of the FIPV polyprotein 1ab (pp1ab) from FIPV strain WSU-79/1146. As predicted, removing the Ubl2 domain significantly improved the purification of FIPV PLP2. Approximately 21 mg of protein was obtained for kinetics and crystallization from a 1 L culture. The higher MW band present throughout the FIPV Ubl2-PLP2 purification was not observed, and only 1 single peak was eluted from the S75 column (Figure 5.3.A). Therefore, a cleaner protein sample was obtained after the two Ni-NTA affinity columns even before SEC.

The final FIPV PLP2 S75 sample was active with a specific activity of ~18,000 μM/min/mg assayed with 200 nM of Ub-AMC (Figure 5.3.B). Another sample expressed in 500 mL instead of 1 L culture and purified with only two Ni-NTA gravity columns also had a

comparable specific activity of ~17,000; therefore, FIPV PLP2 can be purified at 1 L or 500 mL scale. This is in contrast to FIPV Ubl2-PLP2 construct that must be purified in a 500 mL volume, as attempts at a 1 L scale resulted in overexpression of protein in inclusion bodies. Similar turnover rates (~14 min^{-1}) were obtained between FIPV PLP2 with and without the Ubl2 domain assayed at 200 nM Ub-AMC. Overall, FIPV PLP2 without the Ubl2 domain is a robust and active enzyme amiable for kinetic assays and crystallization screening.

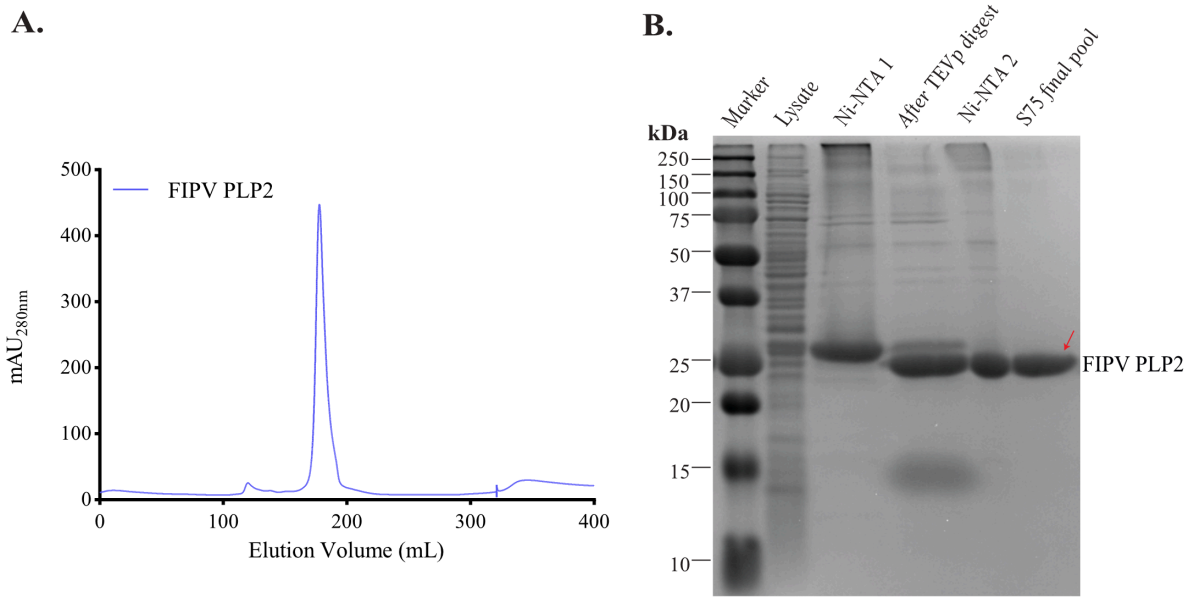


Figure 5.3 Purification of FIPV PLP2 without the Ubl2 domain

(A) SEC chromatogram showing elution profile of FIPV PLP2. (B) 12.5% SDS-PAGE summary gel of purification and final FIPV PLP2 protein used for kinetic experiments and screened for crystallization.

FIPV PLP2 without the Ubl2 domain was surveyed for crystallization conditions using the same protocol as in 2.2.3 with the MERS PLpro core construct. While many crystals were fluorescent in different conditions, crystal hits were probably not diffraction quality crystals since they did not diffract X-rays at APS, or they were salt crystals with protein aggregate adhered to the surface. Nevertheless, further optimization and screening is necessary to obtain a suitable crystal of FIPV PLP2 for data collection.

5.3.3 Ubl2 domain does not impact FIPV PLP2 catalysis and substrate specificity

In our previous study, we show that N-terminal ubiquitin-like 2 (Ubl2) domain directly adjacent to the catalytic domain of MERS PLpro is not required substrate specificity and recognition, and the core PLpro is sufficient for its catalytic function [103]. Therefore, we tested if the Ubl2 domain of FIPV PLP2 influences the hydrolysis against the commercial AMC fluorophores. The kinetic curves for the Ubl2-PLP2 construct alongside the PLP2 core construct are shown in Figure 5.4 and kinetic parameters extracted from these curves are reported in Table 5.3. FIPV Ubl2-PLP2 catalyzes the hydrolysis of Ub-AMC and RLRGG-AMC at a similar efficiency compared to the PLP2 catalytic core but exhibits a 5-fold lower efficiency towards the hISG15-AMC substrate. Although there is a difference in ISG15-AMC catalysis, the same trend holds true that both enzymes prefer Ub-AMC over hISG15-AMC. Therefore, we conclude that the Ubl2 domain does not appear to have a significant effect on catalysis, which is consistent with previous work in βCoVs SARS- and MERS-CoV PLpro [103, 105].

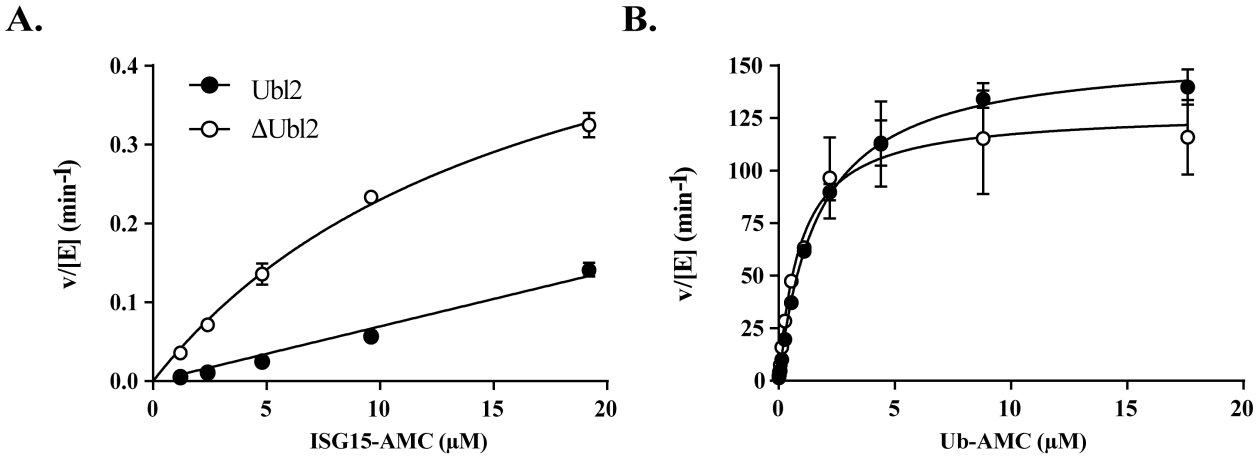


Figure 5.4 Kinetic curves for FIPV PLP2 turnover of ISG15-AMC (A) and Ub-AMC (B) with (dark circles) and without (open circles) the Ubl2 domain

5.3.4 Purification and crystallization of PEDV PLP2

Initial characterization of αCoV PEDV PLP2 was done by a former student Dr. Yafang Chen using the PEDV PLP2$_{1630\text{-}1922}$ construct that contained the Ubl2 domain. However, there

were significant problems with protein yield (<1 mg) and expression as described in [102]. Therefore, it was necessary to redesign this construct for crystallization, which was again done by removing the Ubl2 domain. The PEDV PLP2 core contained sequences 1688-1933 of the PEDV pp1ab from PEDV strain USA/Indiana34/2013, which also contained 11 additional residues at the C-terminus compared to the former construct.

A summary gel of PEDV PLP2 at different stages of purification is shown in Figure 5.5.C. Throughout the purification, higher order oligomers of PEDV PLP2 were formed on the SDS-PAGE gel and were maintained even after boiling samples. After the TEVp digest, there was a shift in MW from his_8-tag cleavage, not only with the PEDV PLP2 monomer, but with the higher order bands as well, further supporting the possibility that PEDV PLP2 oligomerizes (Figure 5.5.C). Fractions 35-37 from the S75 column showed bands at ~54 kDa and ~108 kDa (indicated by red arrows), which would correspond to the PEDV PLP2 dimer and tetramer size, respectively (Figure 5.5.B). These fractions were loaded at 10 μg while fractions 34 and 38, loaded at a lower concentration of <5 μg, did not show distinct higher order bands. Therefore, the formation of oligomers appeared to be concentration-dependent, as a single PEDV PLP2 monomer band was observed when a lower amount of protein (5 μM) was loaded on the gel (Figure 5.9.B), and only 1 single peak was observed during SEC (Figure 5.5.A).

From specific activity analysis, the Ni-NTA pool 1 expressed in a 1 L culture exhibited a higher activity of 17,600 μM/min/mg compared to the Ni-NTA pool 1 from the 500 mL culture (8000 μM/min/mg). After running the Ni-NTA column 2 on each pool, the flow-throughs were combined to increase protein yield for SEC, and the resulting Ni-NTA pool 2 exhibited a specific activity of 17,000 μM/min/mg. After SEC, fraction 35-37 with a specific activity of 13,700 μM/min/mg were pooled for crystallization (~24 mg), and fractions 34 and 38 with a similar specific activity of 11,400 μM/min/mg were pooled for kinetic experiments (~4 mg). The turnover

rates of PEDV PLP2 at 200 nM Ub-AMC was determined to be ~11 min^{-1}, which was comparable to turnover rate for FIPV PLP2 at the same concentration of Ub-AMC (~14 min^{-1}). PEDV Ubl2-PEDV construct, however, was ~3-fold less active. Overall, the core PEDV PLP2 is a robust enzyme for structural and functional studies.

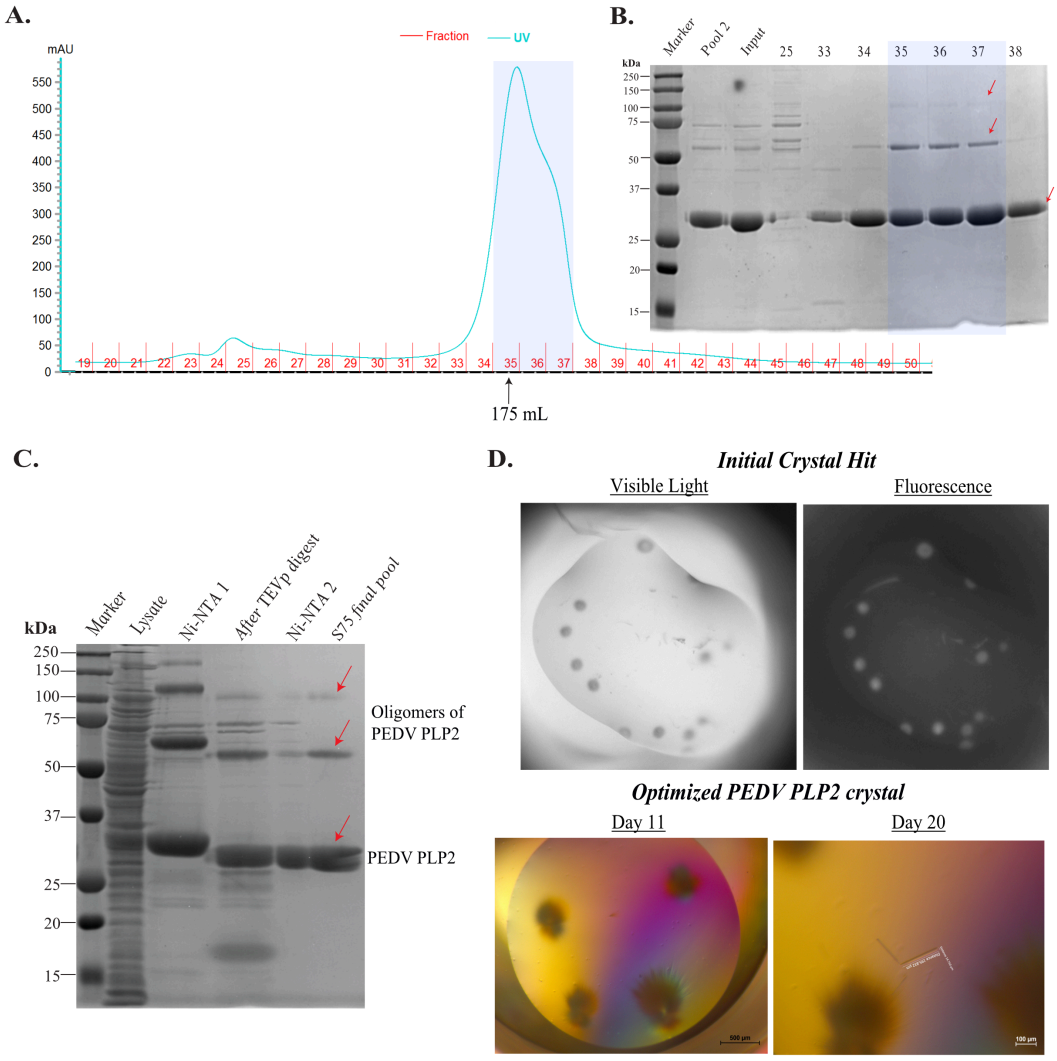


Figure 5.5 PEDV PLP2 purification and crystals

(A) SEC chromatogram of PEDV PLP2. The mAU at 280 nm is plotted as a function of the fraction number. The elution volume where there was a peak in absorbance is indicated. (B) 12.5% SDS-PAGE gel of SEC fractions. Highlighted fractions were pooled for crystallization and loaded at 10 µg. (C) Purification summary gel as analyzed by12.5% SDS-PAGE gel. (D) Top images –Visible light and fluorescence of initial 'sea urchin' crystal hits from 200 nL drops. Bottom images – Visible light images of the condition optimized in 24-well format. PEDV PLP2 rods were 200 µm in length by 15 µm wide.

After screening the S75 final pool for crystallization conditions, several initial crystal hits of PEDV PLP2 were observed. Many crystals resembled 'sea urchin' or fuzzy microcrystals that grow from a single nucleation point. Crystals also grew in common conditions, such as in MCSG-1 condition containing 0.1 M Tris, pH 8.5, 1.8 M magnesium sulfate at all three protein concentrations, in anions suite containing 0.1 M HEPES, pH 7.5, 1.2 M K/Na tartrate at 15 mg/mL PEDV PLP2 (Figure 5.5.D, top images), and finally in the MCSG-3 at 15 mg/mL PEDV PLP2 containing 0.2 M lithium sulfate, 0.1 M Tris, pH 7, and 1 M K/Na tartrate. After setting up an optimization tray of the Anions suite condition at 4°C, thicker protein rods were formed in one single subwell and only varied slightly from the initial crystal condition (0.1 M HEPES, pH 7.5, 1.3 M K/Na tartrate). These exact crystals shown in Figure 5.5.D (bottom images) were harvested for data collection.

5.3.5 FIPV and PEDV PLP2 substrate specificity for Ub and ISG15

To relate αCoV FIPV and PEDV PLP2 catalysis to other βCoV PLPs, we first tested pure PLP2 enzymes for protease, DUB, and deISG activity, respectively, using commercial fluorogenic substrates, RLRGG-AMC, Ub-AMC, and human ISG15-AMC (hISG15-AMC). The kinetic parameters of FIPV and PEDV PLP2 hydrolysis of each Ub-based substrate are reported in Table 5.3 and kinetic curves are shown in Figure 5.6. We also compare the kinetic profiles of these αCoV PLP2s to βCoVs PLPs that have been characterized to date in Table 5.4.

In general, both FIPV and PEDV PLP2 are more efficient DUBs and hydrolyze Ub-AMC at a higher efficiency compared to hISG15-AMC. FIPV PLP2 is nearly ~3800-fold more efficient at cleaving Ub-AMC compared to hISG15-AMC, and PEDV PLP2 is ~1900-fold more efficient towards Ub-AMC over hISG15-AMC. While FIPV PLP2 did approach saturation for ISG15-AMC, the catalytic efficiency was similar to PEDV (within 2-fold). According to apparent k_{cat}/K_m, PEDV PLP2 catalyzes the hydrolysis of hISG15-AMC with a similar efficiency to the small peptide

RLRGG-AMC substrate. PEDV PLP2 is 24-fold more efficient than FIPV PLP2 at cleaving the peptide substrate and even cleaves the peptide similar to (within 3-fold) to SARS-CoV PLpro [94], the most efficient protease against RLRGG-AMC. In contrast, RLRGG-AMC is a poor substrate for FIPV PLP2, as FIPV PLP2 prefers Ub-AMC ~73,000-fold and even prefers hISG15-AMC ~3800-fold to RLRGG-AMC. This suggests that FIPV PLP2 requires addition interactions beyond the S1-S5 subsites for possibly aligning the catalytic triad for optimal catalysis.

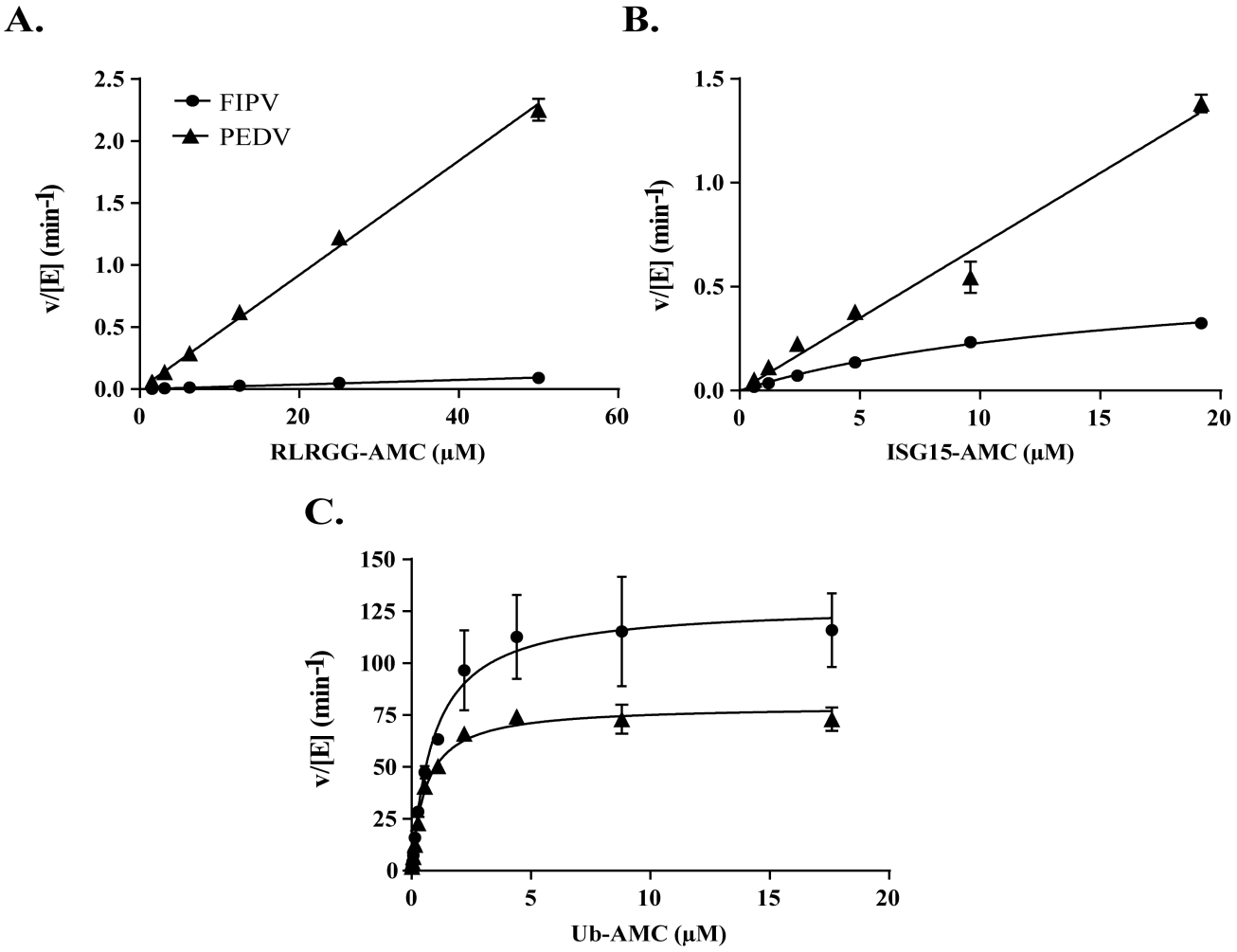


Figure 5.6 FIPV and PEDV PLP2 prefer Ub-AMC over the peptide and ISG15-AMC fluorophore Kinetic curves of FIPV (black circles) and PEDV (black triangle) PLP2 turnover of increasing RLRGG-AMC (A), human ISG15-AMC (B), and Ub-AMC (C) concentrations.

The Ub-AMC substrate preference exhibited by FIPV PLP2 (Ub-AMC >>> hISG15-AMC > RLRGG-AMC) and PEDV PLP2 (Ub-AMC >>> hISG15-AMC ≅ RLRGG-AMC) is in agreement to kinetic data observed previously for MHV PLP2 [54] (Table 5.4). The Ub-AMC preference is in stark contrast to the human CoVs SARS and MERS PLpro that favor hISG15-AMC over Ub-AMC [94]. Human CoV NL63 PLP2 is also able to hydrolyze hISG15-AMC with much higher efficiency, ~1000-fold compared to FIPV and PEDV PLP2 and 25-fold to MHV PLP2, even though this protease prefers Ub-AMC over hISG15 by 15-fold. In general, αCoV PLP2s, prefer the Ub-AMC substrate, and MHV PLP2 is the only βCoV that also follows this trend.

Since the ISG15-AMC substrate used in our assay is the human ISG15 species, the low deISG activity exhibited by animal CoVs, FIPV, PEDV, and MHV PLP2 could simply be due to the differences between human ISG15 compared to ISG15 from their natural host i.e. cats, pigs, and mouse, respectively. The sequence identity of ISG15 among mammals ranges from ~35-80%, depending on the sample size [119]. Human ISG15 is more closely related to porcine ISG15 (77% sequence identity) while feline and mouse ISG15 differ from hISG15 by 69% and 64%, respectively. MHV PLP2 has been shown to cleave immature pro-ISG15 substrates more selectively compared to SARS-CoV and MERS-CoV PLpro, which recognize a diverse panel of pro-ISG15s from human, sheep, fish, mouse, camel, and bat species [119]. Therefore, FIPV and PEDV PLP2 may be similar to MHV PLP2 and only recognize their specific ISG15 species.

Table 5.3 Summary of kinetic parameters and inhibition for FIPV and PEDV PLP2 with ubiquitin-based fluorogenic substrates and inhibitors

Enzyme Kinetic Parameter	**Substrate** RLRGG-AMC[b]	ISG15-AMC[a]	Ub-AMC[a]
FIPV Ubl2-PLP2$_{1488-1801}$			
k_{cat}/K_m (μM^{-1} min^{-1})	0.00098[d]	0.0069 ± 0.0002[b]	90.9 ± 3.9
k_{cat} (min^{-1})	N.S.	N.S.	156 ± 1.7
K_m (μM)	N.S.	N.S.	1.7 ± 0.1
IC_{50} (μM) free Ub			32.8 ± 10.6 μM
IC_{50} (μM) free fISG15			N.I.
FIPV PLP2$_{1588-1794}$			
k_{cat}/K_m (μM^{-1} min^{-1})	0.0019 ± 0.00003	0.0367 ± 0.0048	138.32 ± 24.92
k_{cat} (min^{-1})	N.S.	0.615 ± 0.040	128.19 ± 5.92
K_m (μM)	N.S.	16.77 ± 1.90	0.927 ± 0.161
IC_{50} (μM) free Ub			3.2 ± 0.56
IC_{50} (μM) free fISG15			N.I.
PEDV PLP2$_{1688-1933}$			
k_{cat}/K_m (μM^{-1} min^{-1})	0.0460 ± 0.0005	0.0698 ± 0.0022[b]	131.19 ± 11.78
k_{cat} (min^{-1})	N.S.	N.S.	79.71 ± 1.70
K_m (μM)	N.S.	N.S.	0.608 ± 0.053
IC_{50} (μM) free Ub			25.6 ± 7.96
IC_{50} (μM) free pISG15			N.I.

[a]Steady-state values reported as a mean ± standard deviation, were determined from a minimum of duplicate measurements.
[b]Value of k_{app} with nonsaturating substrate approximates k_{cat}/K_m.
[c] N.S., enzyme not saturated.
[d]Value calculated from only one triplicate point at 50 μM RLRGG-AMC.
[e]N.I., no inhibition.

Table 5.4 Summary of kinetic parameters for different animal and human α- and βCoV PLPs using three FRET ubiquitin-based substrates

	Genus	Enzyme Kinetic Parameter	Substrate RLRGG-AMC[b]	hISG15-AMC[a]	Ub-AMC[a]
Animal CoVs	α	**FIPV PLP2**			
		k_{cat}/K_m (μM^{-1} min^{-1})	0.0019 ± 0.00003	0.0367 ± 0.0048	138.32 ± 24.92
		k_{cat} (min^{-1})	N.S.	0.615 ± 0.040	128.19 ± 5.92
		K_m (μM)	N.S.	16.77 ± 1.90	0.927 ± 0.161
	α	**PEDV PLP2**			
		k_{cat}/K_m (μM^{-1} min^{-1})	0.0460 ± 0.0005	0.0698 ± 0.0022[b]	131.19 ± 11.78
		k_{cat} (min^{-1})	N.S.	N.S.	79.71 ± 1.70
		K_m (μM)	N.S.	N.S.	0.608 ± 0.053
	β	**MHV DPUP-Ubl2-PLP2**[d]			
		k_{cat}/K_m (μM^{-1} min^{-1})	0.0016	2.30 ± 0.1[b]	38.3 ± 6.3
		k_{cat} (min^{-1})	N.S.	N.S.	49.8 ± 2.9
		K_m (μM)	N.S.	N.S.	1.3 ± 0.2
Human CoVs	β	**SARS Ubl2-PLpro**[e]			
		k_{cat}/K_m (μM^{-1} min^{-1})	0.141 ± 0.002	50.7 ± 9.0	7.22 ± 0.56[b]
		k_{cat} (min^{-1})	N.S.	602 ± 49	N.S.
		K_m (μM)	N.S.	11.9 ± 1.9	N.S.
	β	**NL63 Ubl2-PLP2**[f]			
		k_{cat}/K_m (μM^{-1} min^{-1})	0.193 ± 0.001	59 ± 5	880 ± 90
		k_{cat} (min^{-1})	N.S.	3.2 ± 0.1	8.8 ± 0.2
		K_m (μM)	N.S.	0.054 ± 0.004	0.010 ± 0.001
	β	**MERS PLpro**[g]			
		k_{cat}/K_m (μM^{-1} min^{-1})	0.003	15.1 ± 0.8	4.7 ± 0.6
		k_{cat} (min^{-1})	N.S.	19.0 ± 0.3	21.4 ± 0.8
		K_m (μM)	N.S.	1.3 ± 0.1	4.6 ± 0.6

[a]Steady-state values reported as a mean ± standard deviation, from a minimum of duplicate measurements.
[b]Value of k_{app} with nonsaturating substrate approximates k_{cat}/K_m.
[c] N.S., enzyme not saturated. [d]The kinetic parameters for MHV DPUP-Ubl2-PLP2 are reported in [54].
[e]The kinetic parameters for SARS Ubl2-PLP2 are from [119]. [f]The kinetic parameters for NL63 Ubl2-PLP2 are from by [48]. [g]The kinetic parameters for MERS PLpro are reported in [103].

5.3.6 Purifying free ISG15 substrates for inhibition studies

To assess the binding of FIPV and PEDV PLP2 to their natural ISG15 species, free feline and porcine ISG15 were expressed and purified using two Ni his-trap columns from a 2 L total culture. Surprising, throughout purification, we observed activity for ISG15 towards Ub conjugated fluorophores, suggesting that a contaminating DUB in *E. coli* BL21 (DE3) cells may have been co-purified with the substrate. *E. coli* proteins have been shown to have DUB activity *in vitro* [152], and there are prokaryotic homologs of Ub that have the similar β-grasp fold to

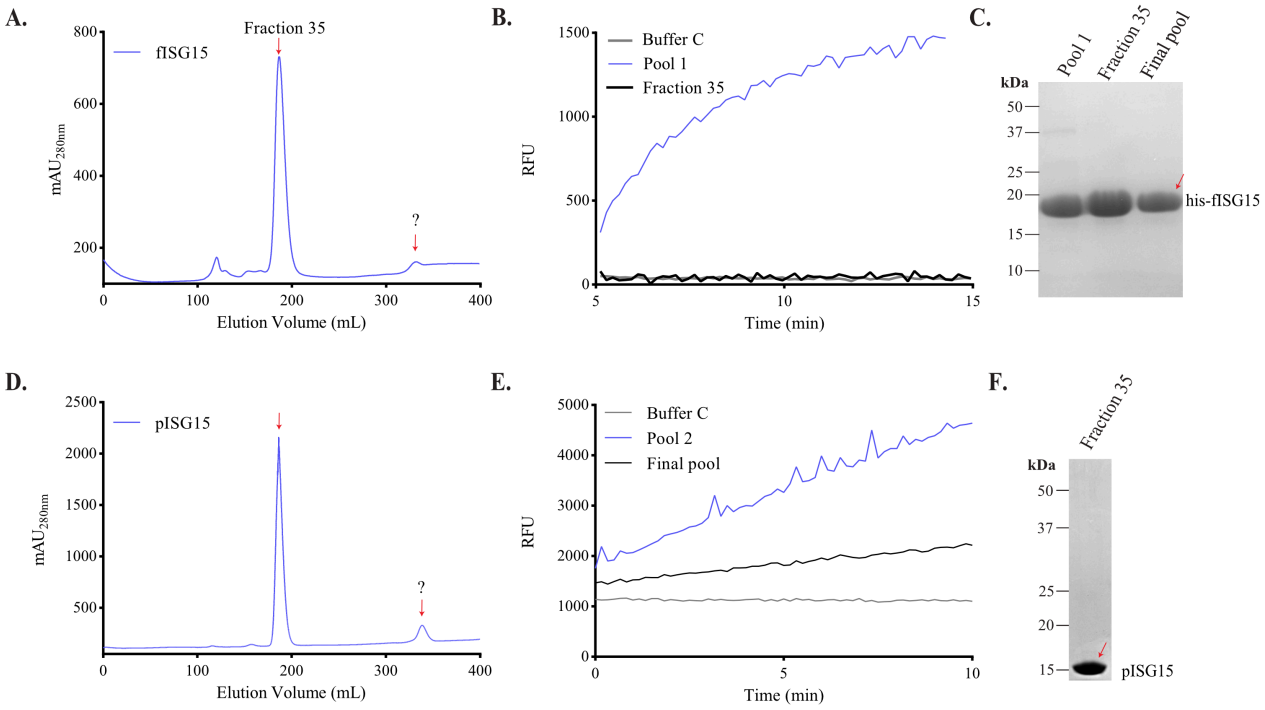


Figure 5.7 Free ISG15 substrates purify with contaminant that shows activity towards FRET ubiquitin substrates

(A) SEC chromatogram of his-fISG15 elution profile. A question mark is indicated at a small peak that was also observed when running Ni his-trap pool 2 untagged fISG15 sample. (B) Fluorescence time course curve assaying activity of Ni his-trap 1 pool versus fraction 35, which had a peak mAU at 280 nm from panel A. Pool 1 and fraction 35 were assayed at a final concentration of 383 µM and 10 µM fISG15, respectively. (C) 12.5% SDS-PAGE analysis of Ni his-trap pool 1, fraction 35, and final pool of his-fISG15 loaded at 10 µg. (D) SEC chromatogram of pISG15 elution profile. A question mark is indicated at the small protein peak. (E) Fluorescence time course curve assaying activity of Ni his-trap 2 pool untagged pISG15 sample at final concentration of100 µM pISG15. (F) 12.5% SDS-PAGE analysis of purity of fraction 35.

ISG15 [153, 154]. Therefore, it is possible that one of these *E. coli* proteins had modest affinity to these ISG15 substrates.

Using SEC, the contaminant was removed from the ISG15 sample, and the final ISG15 pool used for inhibition studies was determined to have no confounding DUB activity (Figure 5.7). It is possible that this contaminant was a very small protein, indicated by a question mark in the SEC chromatogram, or maybe the *E. coli* protein was not abundant enough to visualize by SDS-PAGE but expressed enough to be detected with a fluorescence-based assay. Nevertheless, free fISG15 and pISG15 substrates were purified >90% based on densiometric analysis. In addition, it was important to confirm the substrate exhibited no DUB activity. Therefore, the fluorescence base-line at 100 μM ISG15 was tested using 200 nM Ub-AMC, before assessing its ability to inhibit FIPV and PEDV PLP2.

5.3.7 Free mono-Ub inhibits FIPV and PEDV PLP2 but not free ISG15

To compare the binding of FIPV and PEDV PLP2 to their natural ISG15 species versus Ub, the ability of free mono-Ub and ISG15 to inhibit these proteases were evaluated. As expected, we obtained an ideal inhibition curve when we tested free mono-Ub against FIPV and PEDV PLP2 (Figure 5.8). Assuming competitive inhibition, K_i of Ub was calculated to be 2.6 μM for FIPV PLP2 and 19.3 μM for PEDV PLP2, respectively. Ub is a low micromolar inhibitor for FIPV PLP2 (IC_{50} = 3.2 μM). Based on calculated K_i, FIPV PLP2 binds 7-fold tighter to Ub than PEDV PLP2. Typically, the binding affinity (assuming $K_m \cong K_s$) of Ub-AMC as a substrate is higher compared to free-Ub, which is the product-form inhibitor of the enzymatic reaction. In the case of FIPV PLP2, both the substrate and the product bind with low micromolar affinity while PEDV PLP2 prefers the Ub-AMC substrate over free mono-Ub inhibitor by 32-fold.

In contrast to Ub, when we tested fISG15 and pISG15 for FIPV and PEDV PLP2 inhibition, no inhibition was observed up to 100 μM, suggesting that these enzymes do not bind and have

weak or no affinity to their natural ISG15 species *in vitro*. These data are in consistent to our kinetic assays with hISG15-AMC that support that these enzymes may be poor deISG enzymes in general. Are FIPV and PEDV PLP2 deISGylating enzymes in nature? Our collaborators in Dr. Susan Baker's lab have shown that FIPV PLP2 does remove mouse ISGylated proteins in HEK293T cells (personal communication). Since our *in vitro* system does not take into account the target protein, there is still a possibility that the target protein may be necessary and/or regulate deISG activity of these enzymes. Taken together, FIPV and PEDV PLP2 seem to prefer Ub, as a substrate or product-based inhibitor over ISG15.

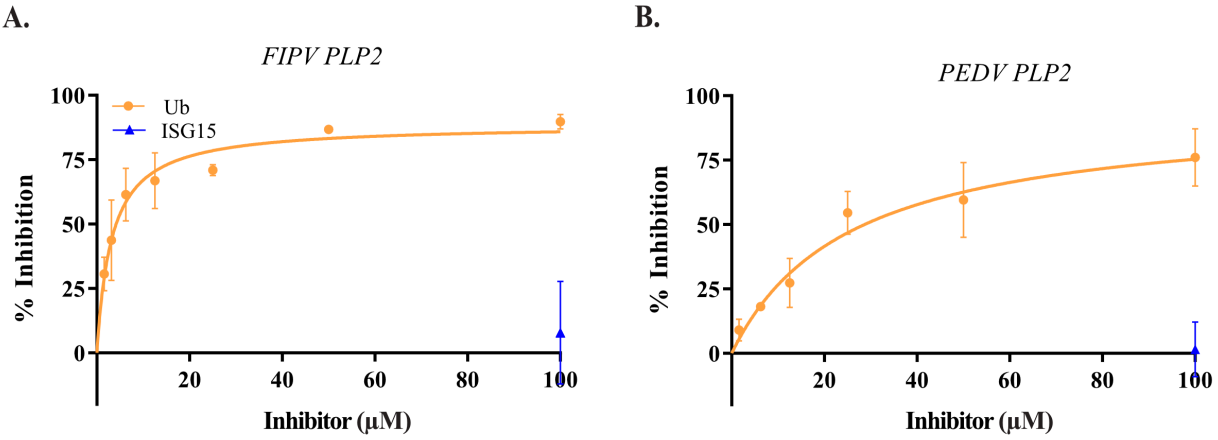


Figure 5.8 Ub but not ISG15 inhibits FIPV and PEDV PLP2

% Inhibition curves for FIPV (A) and PEDV PLP2 (B) across inhibitor concentrations. Serial dilutions of free mono-Ub (orange circles) were tested against both enzymes. One single point at 100 µM feline and porcine ISG15 (blue triangles) were tested against FIPV and PEDV PLP2 inhibition, respectively. Activity of PLP2s was measured using 200 nM Ub-AMC.

5.3.8 FIPV and PEDV PLP2 reactivity with Ub- and ISG15-PA probes

To further assess the host-specific deISG activity of FIPV and PEDV PLP2, we then tested the ability of PLP2 to react and form a covalent adduct with different molar equivalents of PLP2 to propargylamine (PA) probes, either feline ISG15-PA (fISG15-PA) or porcine ISG15-PA (pISG15-PA) versus Ub-PA, using a gel-shift reactivity assay (Figure 5.9.A,B). Each PA probe

contains a modified C-terminal propargylamine (PA) warhead group that is designed to readily react with the catalytic cysteines of PLPs and traps the enzyme in the 'F' acyl-enzyme intermediate-like state.

Surprisingly, FIPV PLP2 only converted ~50% of complex even up to excess molar equivalents of the PA probe (Figure 5.9.A). We observed roughly 50% of unreacted FIPV PLP2 on the SDS-PAGE gel, even up to 1:20 molar ratio of fISG15-PA. Despite FIPV PLP2 robust Ub-AMC activity, the same amount of unreacted FIPV PLP2 was also observed with Ub-PA. Our positive control, MERS PLpro, on the other hand, was able to reach full conversion to MERS PLpro-fISG15 complex at a 1:20 ratio and PEDV PLP2 converted with Ub-PA at a 1:5 ratio. These results indicate that probe integrity was intact. Ubiquitin carboxyl-terminal hydrolase isozyme L3 (UCHL3) exhibited low to high reactivity with different Ub-PA derivatives that present varying length R groups at the C-terminus [155]. Therefore, DUBs could preferentially react with certain PA probes, and FIPV PLP2 may have poor reactivity simply due to the nature of the probe itself.

Another possibly is that during the reaction FIPV PLP2 becomes inactive i.e. is in an unproductive conformation for catalysis against the PA probe or the catalytic cysteine becomes oxidized. To test this hypothesis, we assayed the activity of each PA reacted sample and calculated the remaining activity relative to PLP2 without addition of the probe (Figure 5.9.C,D). Although FIPV PLP2 does not react fully with the PA probes on the SDS-PAGE gel, unreacted PLP2 appears to be inactive, as no activity was detected using Ub-AMC at increasing concentrations of either Ub-PA or fISG15-PA probes (Figure 5.9.C). PEDV PLP2, on the other hand, roughly showed ~60% remaining activity at a 1:1 ratio of PLP2 to PA probe, roughly ~15-18% activity at 1:2 ratio of PLP2 to PA probe, and no activity once completely reacted with the PA probe at a 1:5 ratio. The amount of remaining activity corresponded quite nicely to the amount of unreacted PEDV PLP2

visualized on the SDS-PAGE gel. We conclude that, in the case of FIPV PLP2, full reactivity is not reached because the remaining, unreacted FIPV PLP2 is inactive.

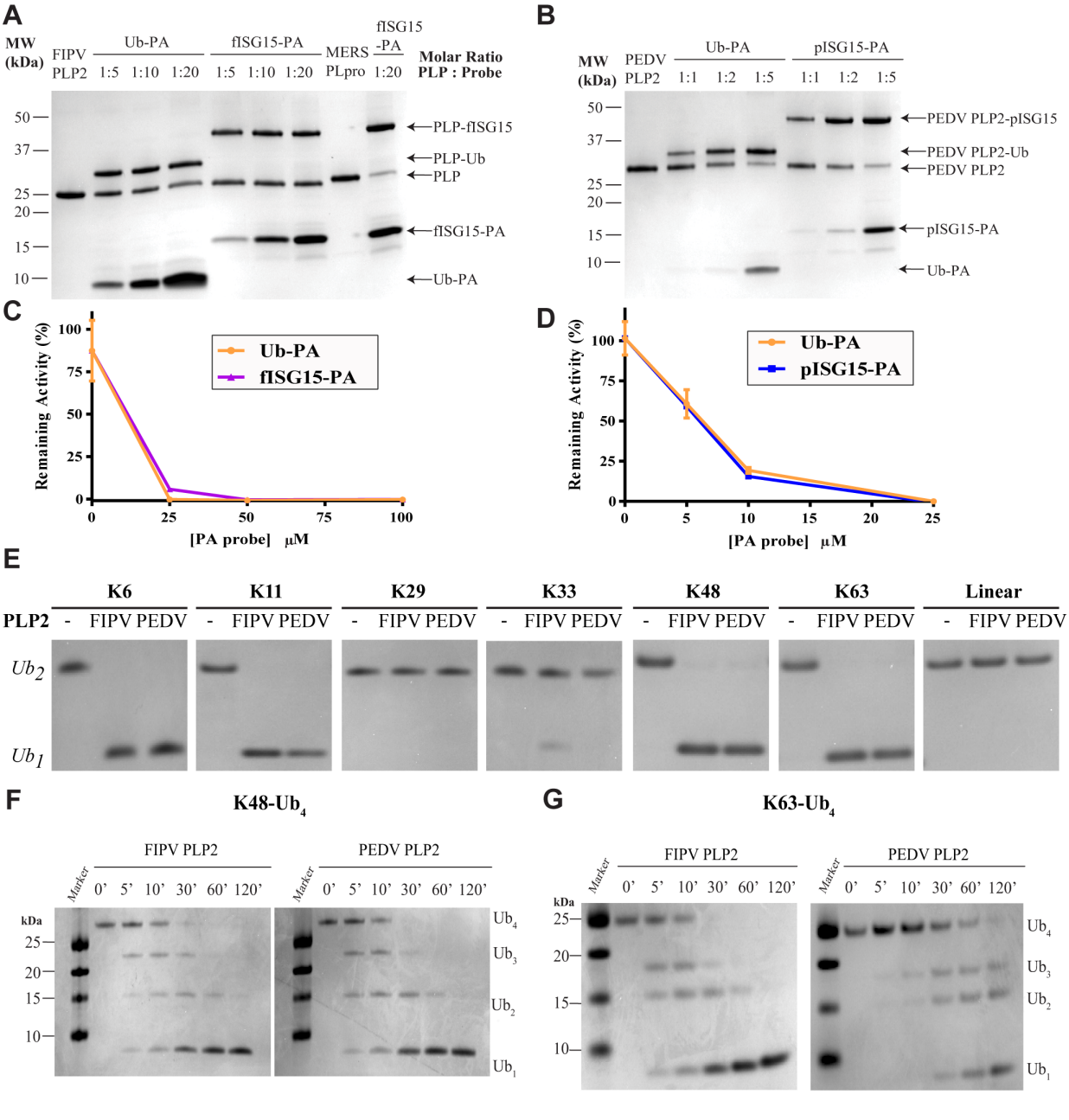
In contrast to FIPV PLP2, PEDV PLP2 was able to reach full conversion of complex within 1 hour at a 1:5 molar ratio with both Ub-PA and pISG15-PA probes (Figure 5.9.B). Although PEDV PLP2 reached full conversion with both modifiers, we noticed that PEDV PLP2 converted with Ub-PA at a faster rate, within 2 min, while conversion with pISG15-PA was delayed and did not occur until a prolonged incubation. Therefore, these data suggest that PEDV PLP2 preferentially reacts with Ub- to ISG15-PA further supporting its preference for Ub.

5.3.9 FIPV and PEDV PLP2 Ub specificity and cleavage for isopeptide-linked chains

Since FIPV and PEDV PLP2 appear to be efficient DUBs, we further investigated the mechanism of their Ub recognition, cleavage, and specificity of Ub-chains, which display the physiological isopeptide bond. To survey Ub-chain specificity, we first performed an Ub-chain panel assay by incubating either FIPV or PEDV PLP2 with different di-ubiquitin (Ub_2) chains linkages for 2 hours (Figure 5.9.E). Similar to other Ub-panels, FIPV and PEDV PLP2 converted K6- , K11-, K48-, and K63-Ub_2 to mono-Ub within 2 hours while partially reacted Ub_2 chains, included K29-Ub_2. No cleavage by FIPV or PEDV PLP2 was observed for linear Ub_2, a commonly linkage that is not recognized by PLPs. We observed partial cleavage of K33-Ub_2 by FIPV PLP2 while PEDV PLP2 failed to cleave this chain linkage. Overall, FIPV and PEDV PLP2 has similar promiscuity for recognizing different Ub_2 chains compared to MERS-CoV PLpro and MHV PLP2 [54, 103] while SARS PLpro is more selective [105].

Figure 5.9 FIPV and PEDV PLP2 reactivity and substrate specificity of Ub and ISG15

(A, B) SDS-PAGE analysis of FIPV (A) and PEDV PLP2 (B) probe reactivity at different molar ratios of Ub- and ISG15-PA. The natural ISG15 species was used, either feline (fISG15-PA) or porcine (pISG15-PA). (C, D) Analysis of the remaining activity of FIPV (C) and PEDV (D) PLP2 after incubation with different molar ratios of PA probe. Relative turnover of PLP2 was tested towards 0.2 μM Ub-AMC and then normalized to the PLP2 control without addition of the probe. (E) FIPV and PEDV PLP2 cleavage and specificity of different Ub_2 chain-linkages, tested at either 0 ("-") or 100 nM PLP2. (F, G) Time-course of FIPV and PEDV PLP2-mediated cleavage of K48-Ub_4 (D) versus K63-Ub_4 from 5-120 min. Negative control (0') was performed without addition of PLP2 while other reactions were initiated with 20 nM PLP2. All covalent adducts and Ub-chains were visualized by coomassie-stain.


A
MW (kDa)
FIPV PLP2
Ub-PA
1:5 1:10 1:20
fISG15-PA
1:5 1:10 1:20
MERS PLpro
fISG15 -PA
1:20
Molar Ratio PLP : Probe
50
37
25
20
15
10
PLP-fISG15
PLP-Ub
PLP
fISG15-PA
Ub-PA
B
MW (kDa)
PEDV PLP2
Ub-PA
1:1 1:2 1:5
pISG15-PA
1:1 1:2 1:5
50
37
25
20
15
10
PEDV PLP2-pISG15
PEDV PLP2-Ub
PEDV PLP2
pISG15-PA
Ub-PA
C
Remaining Activity (%)
100
75
50
25
0
Ub-PA
fISG15-PA
0 25 50 75 100
[PA probe] μM
D
Remaining Activity (%)
100
75
50
25
0
Ub-PA
pISG15-PA
0 5 10 15 20 25
[PA probe] μM
E
K6 K11 K29 K33 K48 K63 Linear
PLP2 - FIPV PEDV - FIPV PEDV - FIPV PEDV - FIPV PEDV - FIPV PEDV - FIPV PEDV - FIPV PEDV
Ub2
Ub1
F
K48-Ub4
FIPV PLP2
PEDV PLP2
Marker
0' 5' 10' 30' 60' 120'
kDa
25
20
15
10
Ub4
Ub3
Ub2
Ub1
G
K63-Ub4
FIPV PLP2
PEDV PLP2
Marker
0' 5' 10' 30' 60' 120'
kDa
25
20
15
10
Ub4
Ub3
Ub2
Ub1

In a previous study, PEDV PLP2 was shown to have global promiscuity in cleaving ubiquitinated proteins as well as chain specificity for K48- and K63 polyubiquitin chains [70], the most abundant and well-characterized chains in signaling pathways that trigger the innate immune system for host protection against viral infection [66, 156]. To gain insight into the mechanism of FIPV and PEDV PLP2 hydrolysis of K48- and K63-linked Ub chains, we monitored the cleavage of extended K48- and K63-tetraubiquitin (Ub_4) chains by incubating each substrate with PLP2 and quenching the reaction at five time points between 5-120 minutes. Results show that FIPV PLP2 cleaves both substrates with equal efficiency, as all of the Ub_4 was converted to mono-Ub within the time-course study for both K48- and K63-linked chains (Figure 5.9.F,G**)**. In contrast, PEDV PLP2 appears to cleave K48-Ub_4 at a faster rate compared to K63-Ub_4. Even at 2 hours, there is still K63-Ub_3 and K63-Ub_2 species present (Figure 5.9.G) while K48-Ub_4 was completely converted to mono-Ub by PEDV PLP2 within 2 hours (Figure 5.9.F). Therefore, PEDV PLP2 may prefer K48-linked polyubiquitin chains compared to K63-linked chains while FIPV PLP2 may not have a preference. Since there was no distinct build-up of Ub_2 as with SARS PLpro and also with TGEV PLP1 [51], it is likely that both of these enzymes have one single recognition site for Ub-chains, similar to MERS PLpro that interacts with poly-Ub chains and ISG15 with one single ubiquitin recognition subsite (SUb1) [103]. SARS PLpro is the only PLP to our knowledge that is capable of recognizing K48-linked Ub_2 [81] and ISG15 [77] with a second ubiquitin recognition subsite (SUb2).

We also performed the same assays with FIPV Ubl2-PLP2 and no striking differences were observed compared with the core PLP2, which suggests that the Ubl2 domain does not influence Ub-chain specificity nor the processing of extended Ub_4 chains (Figure 5.10). It is important to note that 5-fold more FIPV Ubl2-PLP2 enzyme was used in the Ub_4 time course assay, so the core PLP2 may be slightly more active at cleaving these chains. However, these assays were not run

side-by-side for direct comparison. Similar cleavage patterns exhibited by PLP2 with and without the Ubl2 domain is consistent with studies done on the SARS [105] and MERS-CoV PLpro core versus the Ubl2-PLpro enzyme [103].

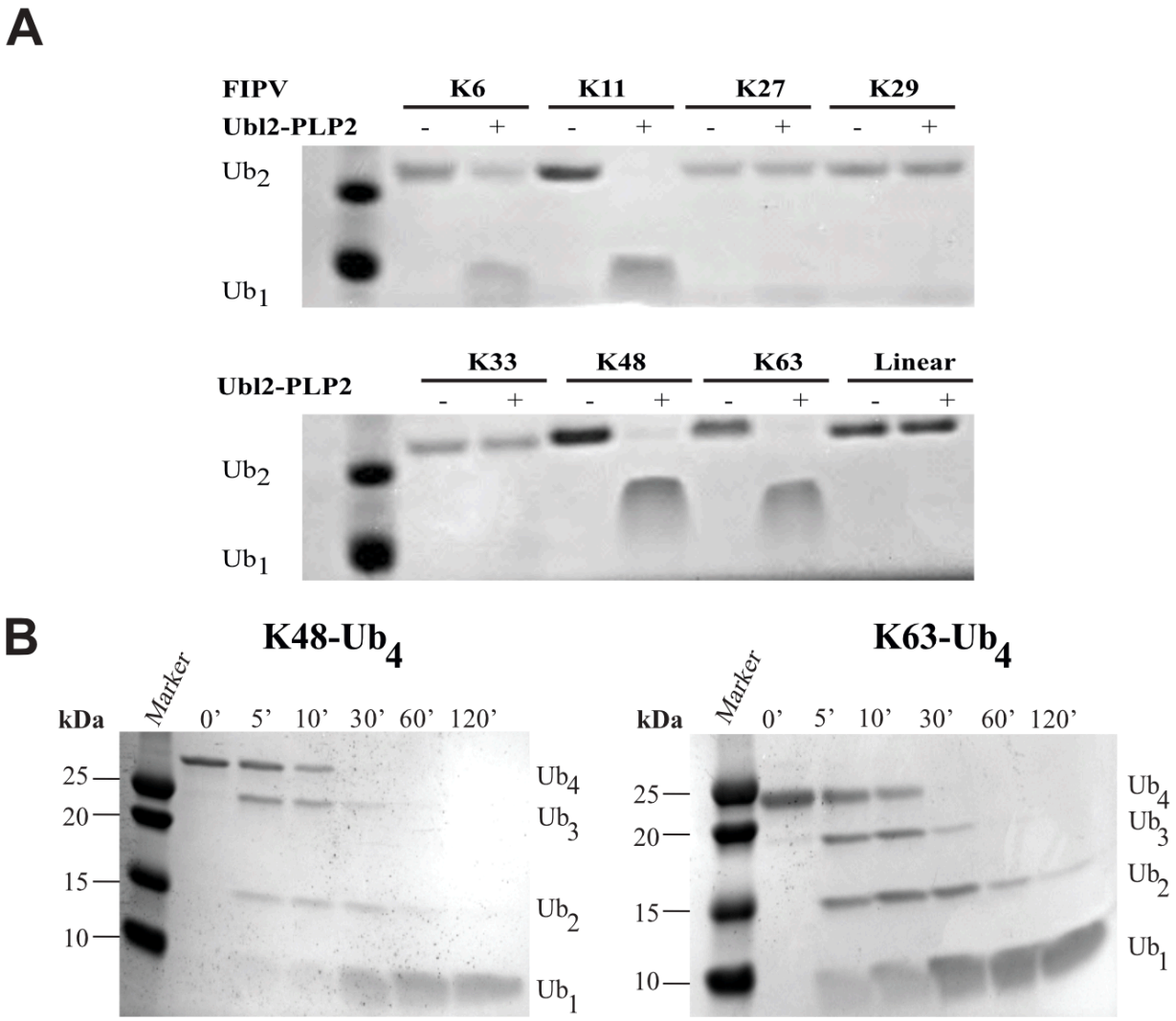


Figure 5.10 Assessing Ub-chain specificity and polyubiquitin processing of FIPV Ubl2-PLP2

(A) A Ub_2 panel was used to survey FIPV Ubl2-PLP2 cleavage specificity of different Ub-chains at a 2 hour time point. (B) Time course of PLP2-dependent cleavage of K48 and K63-linked Ub_4 chains from 5-120 min. The negative (-) symbol and the 0' time point were performed without addition of enzymes. All reactions were initiated with 100 nM FIPV Ubl2-PLP2, and Ub proteins were visualized with coomassie stain.

5.3.10 X-ray crystal structure determination of PEDV PLP2

To shed light on the structure of the αCoV subfamily of PLPs, we attempted to crystallize both FIPV and PEDV PLP2 enzymes. While attempts to crystallize FIPV PLP2 failed to yield diffractable quality crystals, freshly purified PEDV PLP2 crystallized in the C2221 space group at 1.95 Å resolution with a R_{work} of 14.5% and R_{free} of 19.3% (Table 5.5). One unliganded PLP2 molecule was observed in each asymmetric unit while eight PLP2 molecules resided in each unit cell.

To determine the initial phases, the most homologous structure (RMSD of 1.76 Å using secondary-structure matching (SSM) superimpose in coot) to PEDV PLP2 was used for molecular replacement, PLP1 from αCoV transmissible gastroenteritis virus (TGEV), PDB code 3MP2 from ref [51], which yielded one single molecular replacement solution with a top log likelihood gain (LLG) of 31.8 and translation function Z-score (TFZ) of 5.3. We also tried performing molecular replacement with βCoV MERS PLpro (PDB code 5KO3), which similar to our structure does not contain the Ubl2 domain. However, a likely solution was not found. Five molecular replacement solutions were suggested and a negative value LLG of -42.8 indicated that the model was worse than randomly distributed atoms. This suggests that the structural features of our PEDV PLP2 structure is more closely related to αCoV TGEV PLP1 than βCoV MERS-CoV PLpro.

In our PEDV PLP2 structure, the catalytic cysteine of the catalytic triad is singly oxidized (Figure 5.11.B). The sulfenic acid (CSO) adopts two partial conformations. The slightly higher occupancy (occ = 0.60) CSO42 interacts with the main-chain carbonyl of adjacent residue Asn40 while the other (occ= 0.4) forms a productive conformation to the catalytic His201, which is in hydrogen bonding distance to the Asp214 of the catalytic triad. Overall, the catalytic triad of PEDV PLP2 appears to be properly aligned for catalysis, despite the oxidization state of the cysteine.

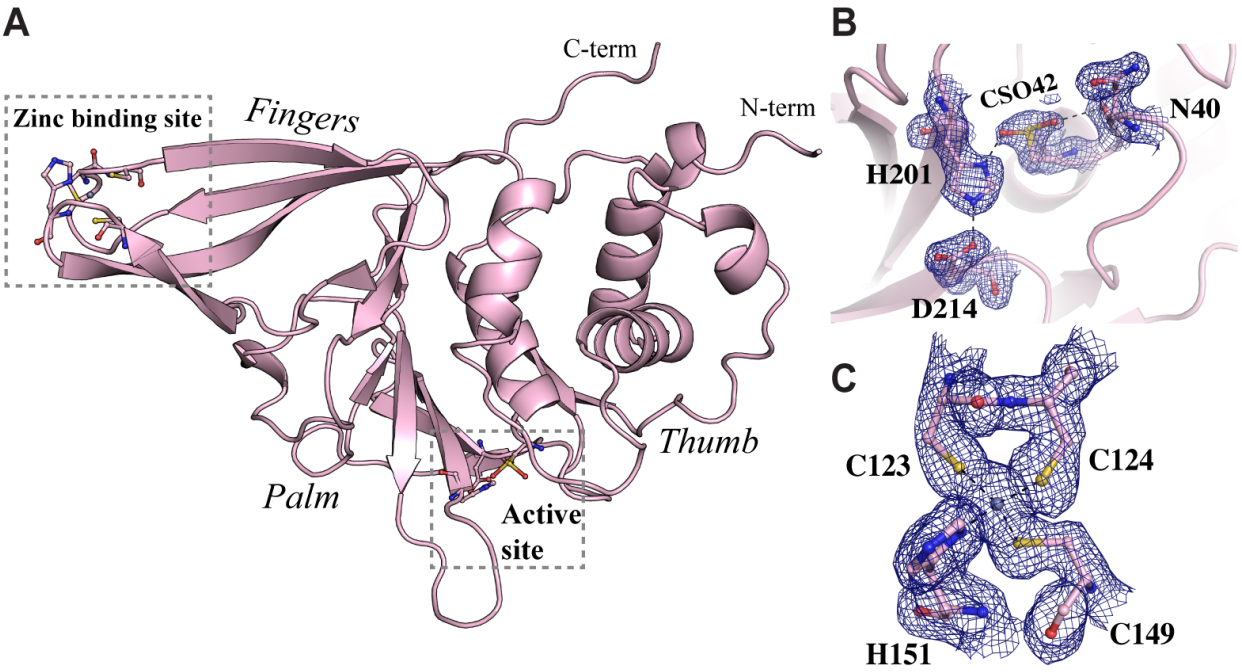


Figure 5.11 Crystal structure of PEDV PLP2

(A) Overall cartoon view of the USP-fold of PEDV PLP2 (PDB code 6NOZ, light pink) with thumb-palm-fingers domains labeled. The active site and zinc-binding site of PEDV PLP2 are indicted with dotted-lined boxes. (B, C) 2Fo-Fc electron density maps (blue mesh, contoured to 1.0 σ) of the zinc-binding motif and active site of PEDV PLP2. (B) The catalytic H201 and D214 of PEDV PLP2 as well as residue N40 interacts with the alternative conformation of the singly oxidized catalytic cysteine (CSO). (B) The structural zinc atom is coordinated by a unique ligand arrangement comprising of C123-C124-H151-C149. Atoms in Panels B-C are colored as follows; nitrogens (dark blue), oxygens (red), sulfur (yellow) carbons (light pink).

Table 5.5 Data Collection and Refinement Statistics of PEDV PLP2

PDB entry	**6NOZ (PEDV PLP2-ΔUbl2)**
Data-collection parameters	
Beamline	MAR225-ID-F
Wavelength (Å)	0.979
Space group	*C2221*
Unit cell dimensions:	
a, b, c (Å)	*49.722, 65.672, 170.038*
α, β, γ (°)	*90, 90, 90*
Resolution (Å)	100-1.95 (1.98-1.95)[a]
Number of reflections observed	311509
Number of unique reflections	20917
R_{merge} (%)[b]	16.4 (58.1)
R_{pim} (%)[c]	6.6 (23.5)
$CC_{1/2}$ (%)	99.0 (91.3)
CC* (%)	99.7 (97.7)
I/σI	17.69 (4.39)
% Completeness	100.0 (100.0)
Redundancy	3.8 (3.6)
Refinement	
Resolution range (Å)	35.9-1.946 (1.995-1.946)
No. of reflections in working set	20248
No. of reflections in test set	1945
R_{work} (%)[d]	14.47 (16.66)
R_{free} (%)[e]	19.31 (22.70)
Wilson B factor (Å^2)	13.45
Average B factor (Å^2)	16.6
RMSD from ideal geometry	
Bond length (Å)	0.007
Bond angle (deg)	1.03
Ramachandran plot	
Most favored (%)	99.2
Allowed (%)	0.80
Disallowed (%)	0

[a]Values in parentheses are for the last (highest resolution) shell.

[b]$R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of a given reflection, and $\langle I(hkl)\rangle$ is the mean intensity of symmetry-related reflections.

[c]$R_{pim} = \sum_{hkl}\sqrt{(\frac{1}{n}-1)}\sum_i|I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i(hkl)$, where n is the multiplicity for multiplicity-weighted R_{merge}.

[d]$R_{work} = \sum_{hkl}\big||F_{obs}| - |F_{calc}|\big| \,/\, \sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[e]R_{free} was calculated using 9.6% of the data set chosen at random that were excluded from the refinement.

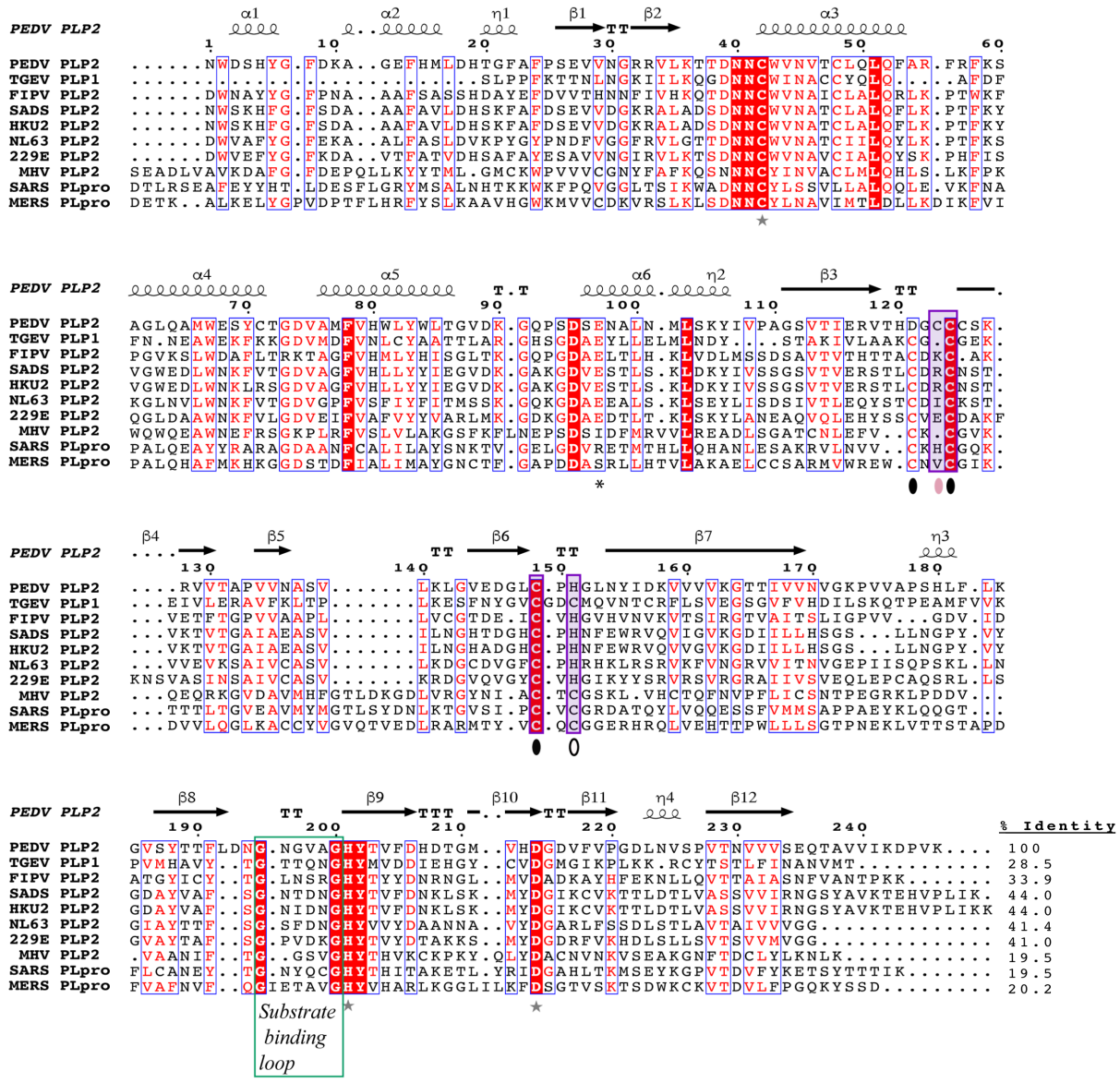
5.3.11 Comparison of PEDV PLP2 to other solved PLP structures.

The overall structure of PEDV PLP2 resembles the global fold and architecture of other viral USPs with a typical right-handed fold containing the thumb, palm, and finger domains. However, within these domains, there are less distinct secondary structure elements and more looping regions in comparison to other PLPs (Figure 5.11.A). For example, the first two α-helices of PEDV PLP2 in the thumb domain are shorter compared to MERS and SARS PLpro. The first α-helix, also known as the ridge helix in SARS and MERS PLpro [77, 103], is not adopted. Instead, a loop is formed before the start of a short α-helical turn in PEDV PLP2.

One unique feature to the PEDV PLP2 structure is that the zinc atom of the structural zinc finger is coordinated by Cys_3-His (Figure 5.11.C) instead of the canonical Cys_4 ligands in other PLP zinc-binding sites that have been solved to date [51, 54, 63, 82]. Although this organization has not been observed yet in the finger domain of other PLP crystal structures, Cys_3-His as well as Cys_2His_2, Cys_3His and Cys_4 are all common ligand cores in structural zinc sites of proteins [157]. Interestingly, when the PEDV PLP2 sequence was aligned to other subfamilies of CoV PLPs, the His151 ligand of PEDV PLP2 appears to be conserved amongst αCoVs, including FIPV, HCoV-NL63, HCoV-229E, SADS, batCoV-HKU2 PLP2 while a cysteine is found at this position in αCoV TGEV PLP1 and βCoVs MHV PLP2, SARS, and MERS PLpro (Figure 5.12). Therefore, it is possible that the Cys_3-His organization is conserved structural feature in αCoVs PLP2s. On the other hand, we noticed that Cys123 ligand of PEDV PLP2 is not conserved in other PLPs. In fact, PEDV PLP2 contains three consecutive Cys in this looping region while other PLPs adopt the equivalent position Cys121 rather than Cys123. Cys124 and Cys149 appear to be conserved ligands of the Zn-finger between different subfamilies of PLPs (Figure 5.12).

Figure 5.12 Sequence alignment of α- and βCoV PLPs

Gray stars; residues of catalytic triad. Purple shaded box; ligands of PEDV PLP2 zinc-binding site. Black oval; conserved Cys ligand of PLP zinc binding sites. Pink oval; unconserved Cys ligand of PEDV PLP2 zinc-binding site. Open circle; conserved His ligand of αCoV PLP zinc binding sites. Green-box; residues of the substrate-binding loop of PLPs. Asterisk; Conserved E97 residue in αCoV PLPs that may be important in catalytic function. Accession numbers of PLPs from replicase pp1ab: PEDV PLP2, AIM47747.1; TGEV PLP1, ABG89333.2; FIPV PLP2, Q98VG9-1; SADS PLP2, AVM41570.1; bat coronavirus HKU2 PLP2, A8JNZ0; HCoV-NL63 PLP2, AFD98833.1; HCoV-299E PLP2, AGW80946.1; MHV PLP2, P0C6X9.1; SARS PLpro, AAP41036.1; MERS PLpro, K4LC41. Multiple sequence alignment was performing using multiple sequence comparison by log-expectation (MUSCLE) and figure was generated using Escript [158]. Secondary structure elements are shown based on the PEDV PLP2 structure. Identical residues are shaded in red with white character. Conserved residues are shown with red character and blue frame represent similarity across groups.


PEDV PLP2
PEDV PLP2
TGEV PLP1
FIPV PLP2
SADS PLP2
HKU2 PLP2
NL63 PLP2
229E PLP2
MHV PLP2
SARS PLpro
MERS PLpro
Substrate binding loop
% Identity
100
28.5
33.9
44.0
44.0
41.4
41.0
19.5
19.5
20.2

Another striking feature of our PEDV PLP2 structure was the C-terminal tail, which is longer than past PLP structures reported to date. This region was manually built in Coot after molecular replacement with TGEV PLP1. In Figure 5.13, we show this looping region of PEDV PLP2 in comparison to MERS-CoV PLpro (RMSD of 2.93 Å using secondary-structure matching (SSM) superimpose in coot). While other PLPs, like MERS PLpro, adopt the last antiparallel β-sheet of the fingers domain, the C-terminus of PEDV PLP2 forms an extended tail that stretches back towards the N-terminus (Figure 5.13). It is tempting to speculate that this long-extended tail could enable flexibility between the different flanking domains of nsp3 to allow for protein-protein interaction. However, it is likely an artifact of crystallization from our construct design. In Figure 5.14, we illustrate crystal contacts made between the C-terminal tail with an adjacent PEDV PLP2 molecule of the neighboring unit cell. Specifically, we observe both hydrophobic and hydrogen-bonding contacts between the C-terminal tail and residues of PLP2 thumb and palm domain near the substrate-binding loop, the highly mobile loop adjacent to the PLP active site (Figure 5.14.C).

The substrate-binding loop of PEDV PLP2 is a region that deviates significantly from MERS PLpro, a nearly ~15 Å swing (Figure 5.13). This is not that surprising, as this loop is

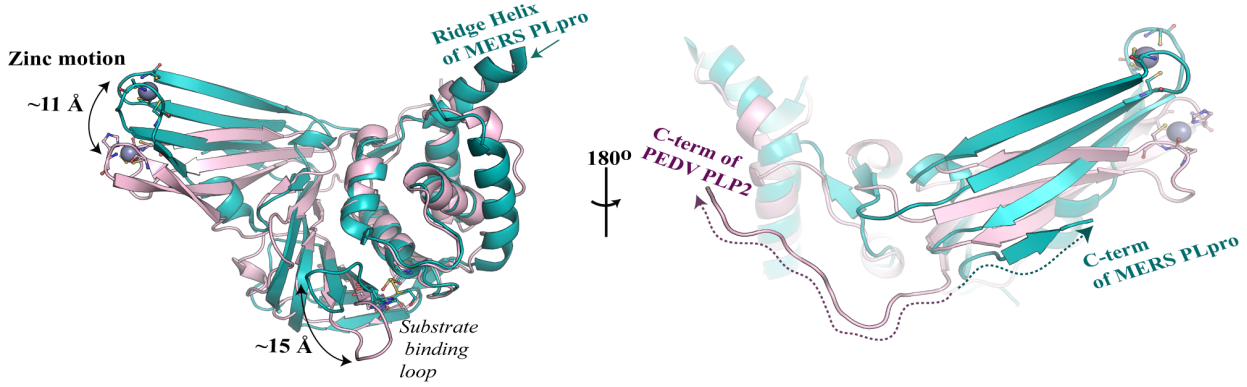


Figure 5.13 Comparison of PEDV PLP2 (PDB code 6NOZ, light pink) to MERS PLpro (PDB code 5K03, teal) crystal structures

Front view, major deviations of zinc-finger and substrate-binding loop are shown. The first α-helix ('Ridge helix') of MERS PLpro is indicated. Back facing view, The C-terminus of each construct, the C-terminal tail of PEDV PLP2; last β-sheet of MERS PLpro.

inherently flexible and must accommodate the RLRGG tail of Ub and ISG15 and is also near crystal contacts in our structure (Figure 5.14.C). PEDV PLP2 contains 6 residues in the substrate-

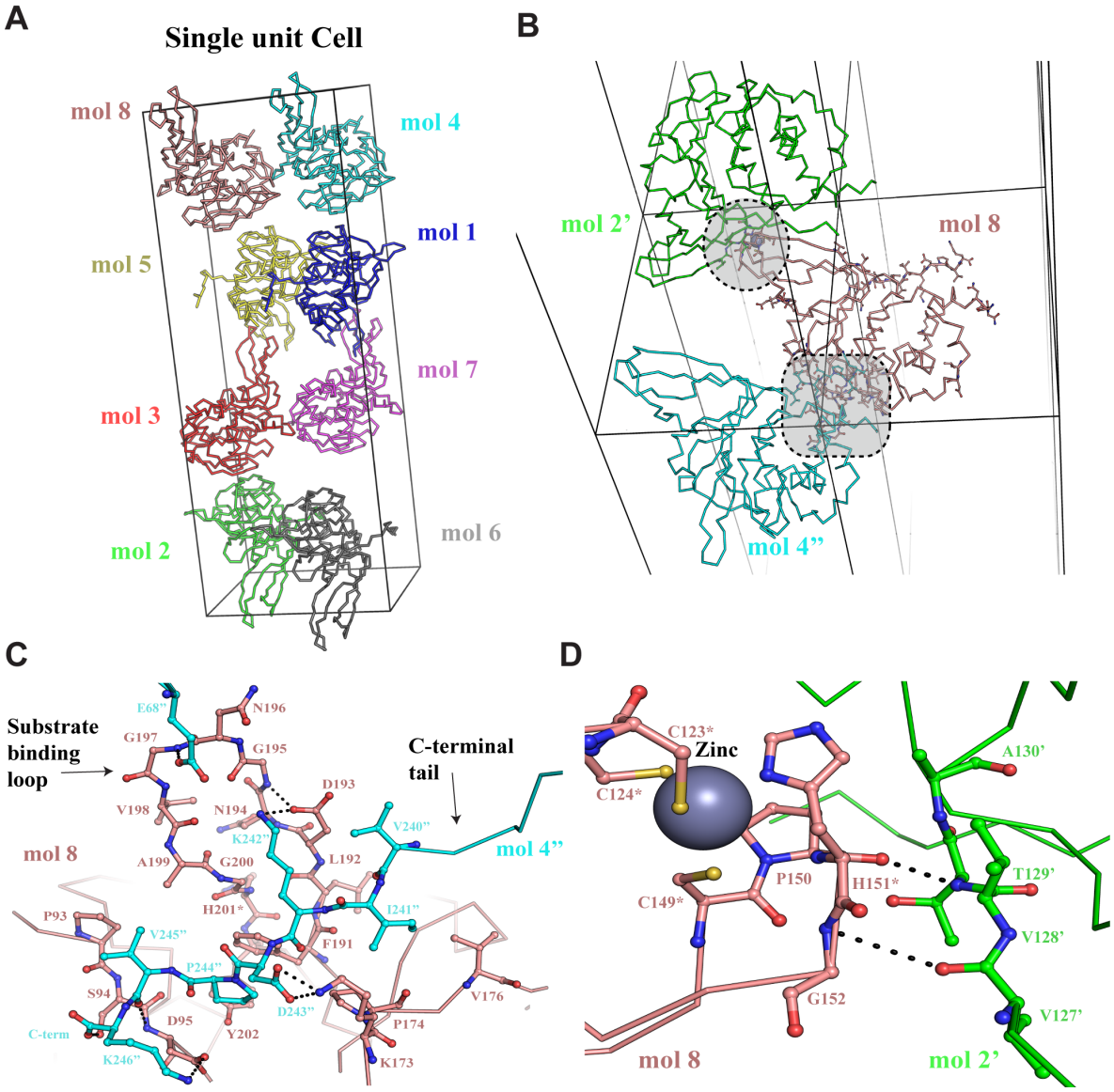


Figure 5.14 Unit cell organization and crystal contact interactions

(A) C2221 space group arrangement of individual PEDV PLP2 molecules within a single unit cell. Each PEDV PLP2 molecule is labeled (1-8) with a different color. (B) Contacts made by three PEDV PLP2 molecules across unit cells. Black lines represent boarders between unit cells. (C) Zoomed in view of crystal contacts between looping residues near the zinc finger motif of mol 8 and residues of the fingers domain in mol 2'. *; zinc-binding ligands (D) Crystal contacts between the C-terminal tail of PEDV PLP2 mol 4'' and residues near the active site and substrate binding loop of mol 8, between the palm and thumb domains. Catalytic H201 is labeled with an asterisk.

binding loop, the same number of residues in most other PLPs (Figure 5.12). MHV PLP2 is lacking one residue while MERS PLpro contains an extra residue in this β-turn (Figure 5.12).

Over the last decade, there has been a significant effort to develop both potent and selective small molecule inhibitors for SARS-CoV PLpro that are metabolically stable and effective at inhibiting viral replication with low cytotoxicity [68, 93, 94]. These competitive inhibitors bind in the substrate-binding loop of SARS-CoV PLpro via an induce-fit mechanism. To determine if these compounds could inhibit either FIPV or PEDV PLP2 in a similar manner, we tested for FIPV and PEDV PLP2 inhibition by compound **3e** (structure shown in Figure 2.9) at 100 μM. Results showed no inhibition for FIPV and PEDV PLP2 by **3e** while SARS PLpro exhibited 96.3 ± 0.3 % inhibition at 100 μM compound **3e**. Compound **3e** has only been reported to inhibit SARS-CoV PLpro and human CoV NL63 PLP2. These enzymes both contain a bulky aromatic residue at position 198 of the substrate-binding loop (numbering in Figure 5.12), where the inhibitor binds and forms a nice hydrophobic clamp. While SARS and NL63 PLP contain a tyrosine and phenylalanine in their substrate-binding loop, PEDV and FIPV PLP2 have a glycine and asparagine at this equivalent position, respectively. In fact, this bulky residue appears to be absent in all other PLPs including MERS-CoV PLpro, which **3e** also failed to inhibit [94]. The nature of the substrate-binding loop in most other PLPs pose a significant challenge when designing a small molecule inhibitor to target PLPs beyond SARS-CoV and human CoV NL63 PLP, as these small and/or flexible residues would not likely form an ideal pocket to enable small molecule inhibitor binding.

In addition to the substrate-binding loop, there was also notable mobility in the Zn-finger motif between PLPs structures, illustrated between PEDV PLP2 and MERS PLpro in Figure 5.13. The dynamic nature of the Zn-finger motif is typical, and the conformation of the structural zinc has been shown to vary between past PLP crystal structures [83, 133]. The Zn-finger motif is also

a point of crystal contact in our PEDV PLP2 structure, as we observed main-chain interactions in this region with neighboring PEDV PLP2 molecules within the unit cell (Figure 5.14). Therefore, crystal packing likely influences the conformation adopted here.

5.3.12 Structure-guided removal of FIPV and PEDV PLP2 DUB activity

Since FIPV and PEDV PLP2 are robust DUBs *in vitro* and DUB activity is implicated in the ability of PLP2 to antagonize and suppress IFN expression [70] to promote viral infection, we focused our efforts towards disrupting Ub recognition by FIPV and PEDV PLP2. The goal is to create PLP2 mutants that maintain their ability to cleave the viral polyprotein sequence for virus replication but can no longer recognize Ub conjugates for the suppression of the innate immune system. PLP2 mutants that are selectively DUB deficient may be useful for reverse genetics to create live-attenuated FIPV and PEDV viruses for vaccine design.

In our recent study described in Chapter 4, we utilized a structure-guided approach to prevent Ub-binding by MERS-CoV PLpro without significantly affecting its ability to cleave the RLRGG-AMC substrate, as a measure of active site functionality. We demonstrated that introducing a positive charge either at position 101 or 136 (previous nomenclature H1652 and V1691 in Chapter 4) selectively eliminated the DUB activity of MERS-CoV PLpro by disrupting its interaction with R42 and hydrophobic patch residue I44 of Ub.

To gain insights into how a positive charge might have altered the Ub-binding site of PLpro and compare that with other PLPs, we analyzed the electrostatic surface potential of PLPs solved to date and labeled residues of the SUb1 subsite investigated in MERS-CoV PLpro with equivalent residue positions in our PEDV PLP2 structure (Figure 5.15). We identified the residues of the SUb1 subsite in FIPV and PEDV PLP2 by superimposing our PEDV PLP2 structure with the structure of MERS PLpro in complex with Ub [83] as a molecular guide (Figure 5.16). A homology

model of FIPV PLP2 was generated with the Swiss-Model program to predict residues found in FIPV PLP2 [159].

Although electrostatic potentials across the surface of PLPs vary, there are some interesting trends that we observed. For instance, positions 101 and/or 136 are in a neutral region of the protease. Specifically, SARS and MERS PLpro as well as FIPV and MHV PLP2 all contain a hydrophobic residue at position 136. From previous studies in our group, V136 and F136 in MERS PLpro (Chapter 4) and MHV PLP2 (nomenclature F1812 in [54]) respectively, both form a key hydrophobic interaction with I44-V70 patch of Ub, and is therefore notably important for DUB activity. Similarly, M136 in SARS (Chapter 3, nomenclature M209 in [119]) also engages with this hydrophobic patch region.

While position 136 is often hydrophobic, residues at position 97, especially for αCoV PLPs, are nested in an acidic pocket. Therefore, it is possible that inserting a basic residue at these positions would likely alter the charge distribution in this region and hence affect Ub binding. In fact, basic residues that are naturally found in these proteases actually hinder DUB activity, as mutagenesis of these arginine residues to an acidic side-chain commonly enhance DUB activity. For example, MERS-CoV PLpro appears to have a unique arginine at position 98 that forms a small basic patch in an overall more neutral SUb1 compared to other PLPs (Figure 5.15). When we mutated R98 residue in MERS-CoV PLpro to an alanine or to the equivalent glutamate in SARS-CoV PLpro, we observed a significant enhancement of DUB and peptide activity by MERS PLpro (Chapter 4). R98 forms a unique guanidium stacking arrangement with R72 of Ub and, unlike E98 in SARS-CoV PLpro, R98 must accommodates the charge of R42 for Ub binding. Therefore, an acidic residue is preferred at position 98 in MERS-CoV PLpro for Ub recognition and catalysis.

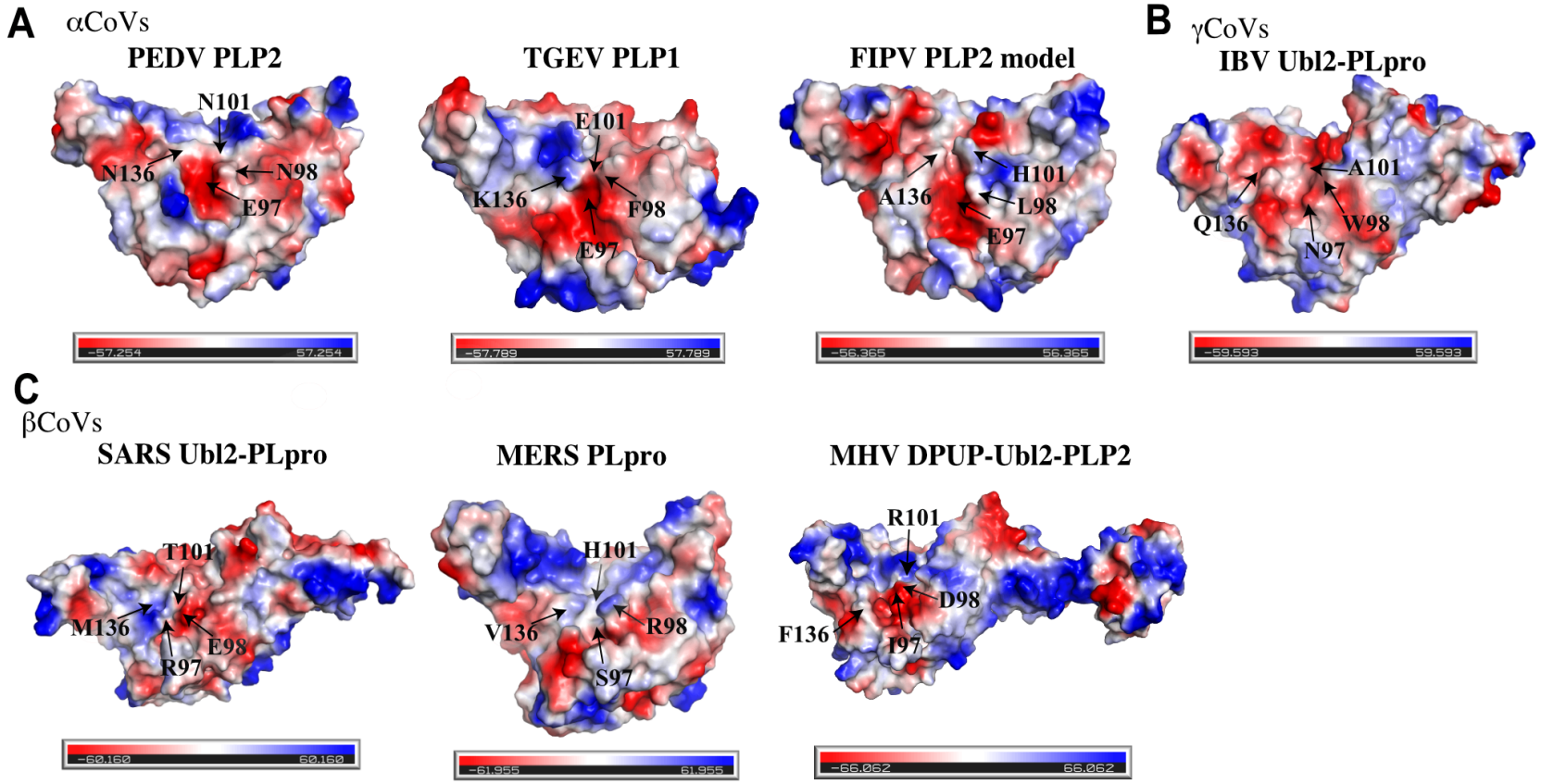
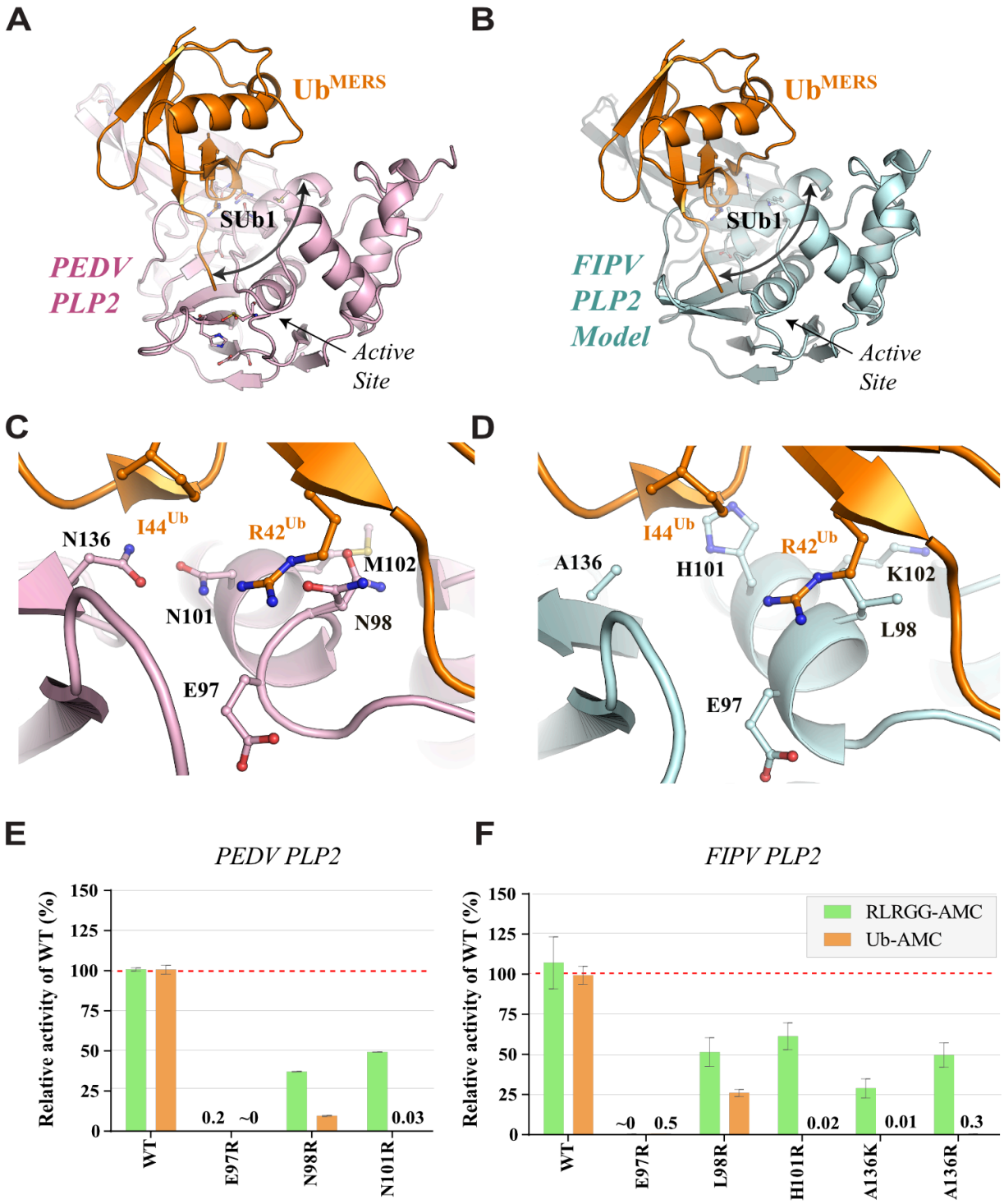


Figure 5.15 Electrostatic surface maps of different subfamily of CoV PLPs

(A) αCoVs PEDV PLP2, TGEV PLP1 (PDB code 3MP2), FIPV PLP2 swiss model, (B) γCoV IBV Ubl2-PLpro (PDB code 4X2Z), (C) βCoV SARS Ubl2-PLpro (PDB code 3E9S), MERS PLpro (PDB code 5KO3), and MHV DPUP-Ubl2-PLP2 (PDB code 4YPT) structures with surface colored based on electrostatic potential. Equivalent residues at positions 97, 98, 101, and 136 are indicated for each PLP. Charge smoothing potential was generated using vacuum electrostatics in Pymol.

Figure 5.16 Structure-based design of DUB deficient mutants in PEDV and FIPV PLP2

(A, B) Overall views of PEDV PLP2 structure (panel A, PDB code 6NOZ, light pink) and FIPV PLP2 model (panel B, light cyan) superimposed with MERS PLpro-Ub bound structure (PDB code 4RF1, orange). MERS PLpro is omitted for clarity. The proposed SUb1 binding pocket of PLP2 is indicated, along with the location of the catalytic active site. (C, D) Magnified view of residues sought for mutation in the PEDV (C) and FIPV PLP2 (D) SUb1 binding pocket, which are near residues R42 and I44 of Ub (PDB code 4RF1, yellow). (E, F) Relative turnover number of PEDV (E) and FIPV (F) PLP2 mutants compared to the WT towards 50 μM RLRGG-AMC (green) and 1 μM Ub-AMC (yellow).


A
Ub^MERS
SUb1
PEDV
PLP2
Active
Site
B
Ub^MERS
SUb1
FIPV
PLP2
Model
Active
Site
C
N136
I44^Ub
R42^Ub
M102
N101
N98
E97
D
I44^Ub
A136
H101
R42^Ub
K102
L98
E97
E
PEDV PLP2
Relative activity of WT (%)
150
125
100
75
50
25
0
0.2
~0
0.03
WT
E97R
N98R
N101R
F
FIPV PLP2
Relative activity of WT (%)
RLRGG-AMC
Ub-AMC
~0
0.5
0.02
0.01
0.3
WT
E97R
L98R
H101R
A136K
A136R

While MERS-CoV PLpro has a unique arginine at position 98 compared to other PLPs, SARS-CoV PLpro appears to have a unique arginine at position 97. R97 in SARS-CoV PLpro interacts with Q39 of Ub and is positioned away from R42 of Ub that forms a salt bridge with E98 of PLpro [77, 81]. When we swapped R97 in SARS-CoV PLpro to a glutamate, it resulted in a hyperactive enzyme with a 10.6-fold increase in catalytic efficiency towards Ub-AMC [119], a similar enhanced trend, as investigated with R98 mutants in MERS-CoV PLpro. Interestingly, E97 is a conserved residue in αCoV PLPs (Figure 5.12). Taken together, an acidic residue at position 97 is preferred in SARS-CoV PLpro for DUB activity and also conserved in the αCoV PLP2s. Therefore, it may be important for Ub recognition by FIPV and PEDV PLP2.

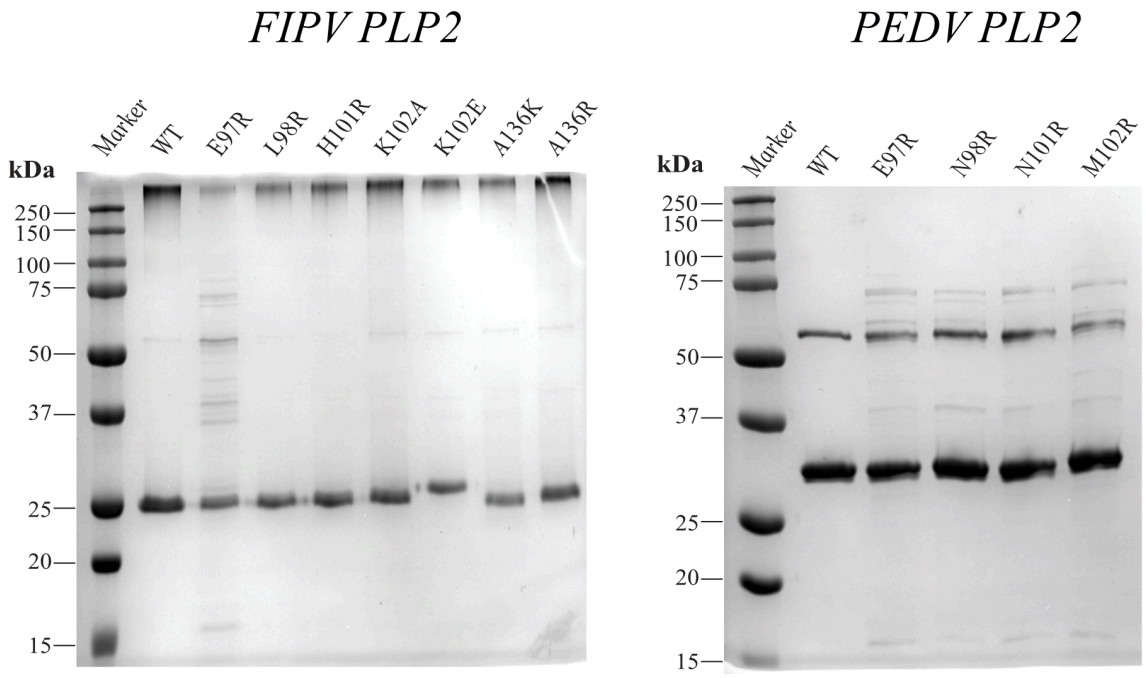


Figure 5.17 12.5% SDS-PAGE analysis of pure FIPV (A) and PEDV PLP2 WT (B) and mutants used for kinetic experiments.

Protein was loaded at 4 µg. Some aggregation was observed at the top of the gel for the FIPV PLP2 mutants. The higher order band observed in PEDV PLP2 WT and mutants corresponds to the PEDV PLP2 dimer (~54 kDa).

Since inserting a basic, positively charged residue seems to be a common strategy to disrupt DUB activity, and vice versa for enhancing DUB activity, in coronaviral PLPs, we designed

several FIPV and PEDV PLP2 mutations at residue positions 97, 98, 101, 102, and 136 by substituting either an arginine or a lysine residue (Table 5.1). FIPV PLP2 has a lysine at position 102, so this residue was changed to either an alanine or a glutamate. These residues, as described above, are predicted to be in close proximity to R42 and hydrophobic patch residue I44 of Ub (Figure 5.16.C and 5.16.D). The FIPV and PEDV PLP2 mutant enzymes purified well, shown in Figure 5.17, using a Ni-NTA gravity flow batch method with 500 mL total cell culture. We tested these mutant enzymes for Ub-AMC and RLRGG-AMC activity at a single substrate concentration and compared the relative turnover rate of hydrolysis to that of the WT enzyme. All DUB deficient mutations that showed a defect in DUB activity (>3-fold different compared to the WT) are shown in Figure 5.16.E for PEDV PLP2 and Figure 5.16.F for FIPV PLP2. Mutants that only showed a marginal effect on DUB activity was $K102A^{FIPV}$, $K102E^{FIPV}$, and $M102R^{PEDV}$ were excluded from this figure for clarity, but are reported in Table 5.6.

Overall, we observed very similar trends between FIPV and PEDV PLP2 mutants. Both E97R mutants in FIPV and PEDV PLP2 showed a drastic decrease in DUB activity, but also a significant loss in RLRGG-AMC activity. FIPV PLP2 E97R showed almost no detectable RLRGG-AMC activity while there was a 200-fold reduction in Ub-AMC activity. PEDV PLP2 E97R showed nearly no detectable Ub-AMC activity while 500-fold less RLRGG-AMC activity was observed compared to the wild-type. As previously noted, E97 appears to be conserved among αCoVs PLP2s as well as TGEV PLP1, but is absent in βCoVs, so it is possible that this residue plays an essential role in the αCoV PLP catalytic mechanism for maintaining functionality of the active site (Figure 5.12 and 5.15). We hypothesize that this residue may contact R72 of Ub, which is the P5 arginine of RLRGG sequence, as removing this salt bridge interaction with E97 of PLP2 may be attributed to the activity loss observed with the RLRGG-AMC peptide.

Table 5.6 Summary of the turnover rate and fold activity of FIPV and PEDV PLP2 mutants towards 50 μM RLRGG-AMC and 1 μM Ub-AMC

	Mutant	**RLRGG-AMC k_{cat} (min^{-1})[a]**	**Fold Activity reduction**	**Ub-AMC k_{cat} (min^{-1})[a]**	**Fold Activity reduction**
FIPV PLP2	WT	0.10 ± 0.2	1.0	39.1 ± 2.2	1.0
	E97R	~0	2000	0.2 ± 0.01	200
	L98R	0.05 ± 0.01	2.1	10.3 ± 0.85	4
	H101R	0.06 ± 0.01	1.7	0.009 ± 0.0001	4400
	K102A	0.31 ± 0.05	0.3	54.0 ± 3.9	0.7
	K102E	0.38 ± 0.07	0.3	65.3 ± 1.5	0.6
	A136K	0.03 ± 0.01	3.7	0.005 ± 0.0004	8400
	A136R	0.05 ± 0.01	2.2	0.132 ± 0.013	300
PEDV PLP2	WT	1.2 ± 0.01	1.0	33.6 ± 1.1	1.0
	E97R	0.002 ± 0.0006	500	~0	>Very High
	N98R	0.44 ± 0.002	2.7	3.2 ± 0.12	10.5
	N101R	0.58 ±0.001	2.0	0.01 ± 0.0002	3000
	M102R	0.61 ± 0.02	2.0	20.1 ± 2.9	1.7

[a]Turnover rates were determined from triplicate measurements, reported as a mean ± standard deviation.

H101R^{FIPV}, N101R^{PEDV}, A136K^{FIPV} and A136R^{FIPV} mutants showed a severe disruption in DUB activity while peptide activity was less effected compared to the E97R mutants. N98R^{PEDV} maintained 9.6% Ub-AMC activity while L98R^{FIPV} maintained 26% Ub-AMC activity. Even though these mutants showed a 10.5-fold and 4-fold loss, respectively, in Ub-AMC activity, their deficiency was not as pronounced as the remaining mutants. Excitingly, H101R^{FIPV} and N101R^{PEDV} mutants shows a drastic decrease in DUB activity, 3000-4000-fold, while the RLRGG-AMC activity was still within 2-fold. A136K^{FIPV} and A136R^{FIPV} showed a similar trend to the H101R^{FIPV} mutant in a loss of Ub-AMC activity, 8400 and 300-fold, respectively. While A136R^{FIPV} maintained ~50% RLRGG-AMC activity (within 2-fold of WT), A136K^{FIPV} was less efficient towards the peptide substrate and exhibited 3.7-fold reduction. However, the deficiency of A136K^{FIPV} DUB activity is much more severe in comparison to its reduction in RLRGG-AMC activity. From these data, inserting a charge in positions 98, 101, and 136 of FIPV PLP2 and positions 98 and 101 of PEDV PLP2 are all strategies that can be used to selectively disrupt DUB activity of PLP2 with the latter positions as the most promising target.

To further examine the DUB deficiency exhibited by H101R^{FIPV} and N101R^{PEDV} mutants, we tested the ability of these enzymes to react with 1:5 molar equivalents of PLP2 to Ub-PA and ISG15-PA probes, as done with the WT, at different time points between 2-30 min. While the FIPV PLP2 WT converted to ~50% PLP2-Ub and PLP2-fISG15 complex within 2 min, the H101R^{FIPV} showed a delay in conversion to complex at the earlier time points (Figure 5.18.A). Interestingly, H101R^{FIPV} showed a stronger defect at the earlier time points with fISG15-PA, suggesting that this mutation also targets the ability of PLP2 to react with fISG15-PA as well as Ub-PA. A similar trend was observed with the N101R^{PEDV} mutant. While PEDV PLP2 WT was able to react with the PA probes within 30 min, N101R^{PEDV} mutant exhibited little to no reactivity up to 5 minutes and only 50% complex formed after the 30 min incubation (Figure 5.18.B).

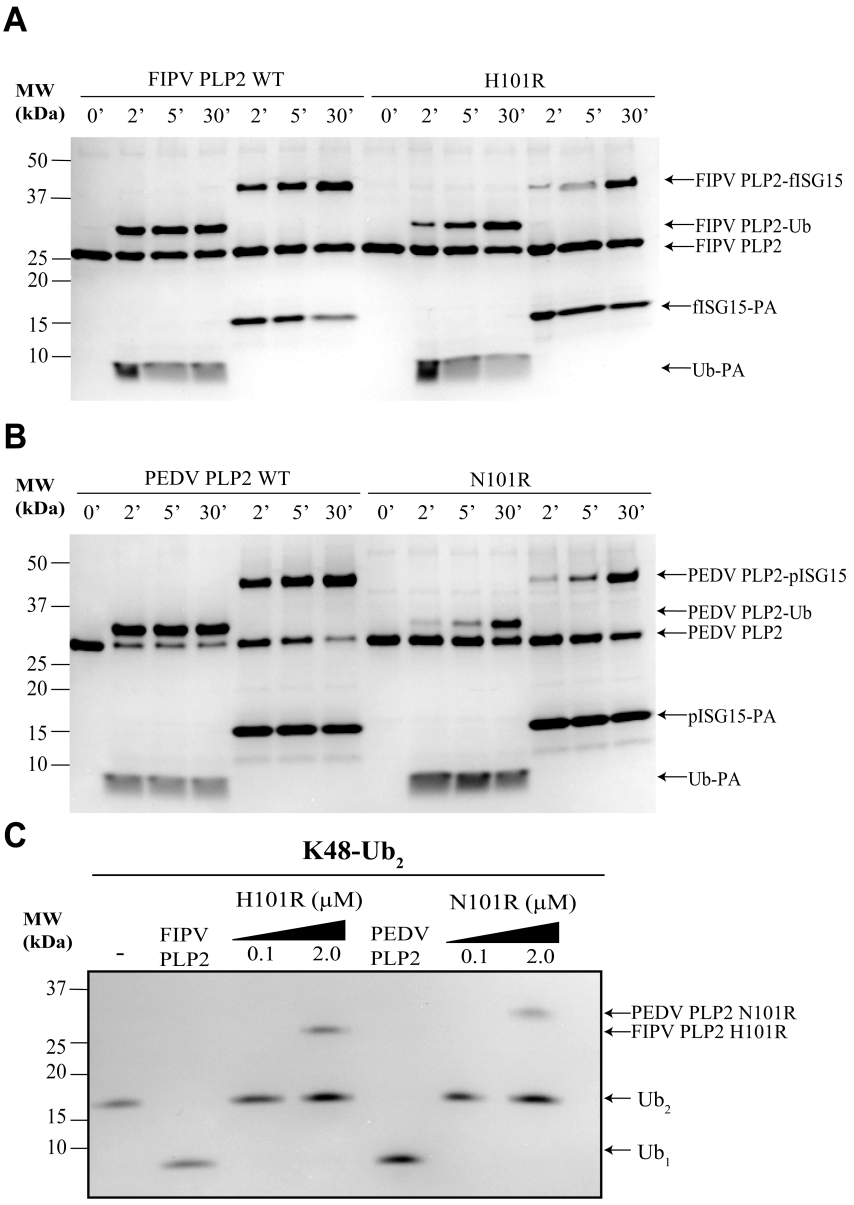


Figure 5.18 H101R^{FIPV} and N101R^{PEDV} deficiency in Ub- and ISG15-PA probe reactivity and K48-Ub_2 cleavage.

(A, B) Time-dependent PA probe reactivity of FIPV (A) and PEDV (B) PLP2 WT versus position 101 mutants. Reactions were performed independently at 2, 5, 30 min with a fixed 5 μM PLP2 to 25 μM Ub-PA or ISG15-PA final concentration. Negative control (0') was performed without addition of PA probe. (C) K48-Ub_2 recognition and cleavage by PLP2 WT and increasing concentration of Arg101 mutants. All covalent protein adducts and Ub_2 chains were visualized with Coomassie stain.

Since both mutants showed a defect for ISG15-PA reactivity, these mutations may target deISG activity as well as DUB activity. As previously mentioned, position 101 in MERS PLpro (referred to as H1652 in Chapter 4) was shown to interact with the conserved Pro and Trp of ISG15 species [74], which is present in both fISG15 and pISG15. These Arg mutants at position 101 could therefore sterically clash with ISG15 at this hydrophobic region. To access if $H101R^{FIPV}$ and $N101R^{PEDV}$ mutants are impaired in their cleavage of Ub-chains linked by isopeptide bonds, we tested their cleavage of K48-Ub_2 chains compared to the WT. While FIPV and PEDV PLP2 cleaved K48-Ub_2 to mono-Ub within 2 hours at 100 nM, no cleavage was observed for the $H101R^{FIPV}$ and $N101R^{PEDV}$, even up to 2 μM PLP2. Therefore, we conclude that $H101R^{FIPV}$ and $N101R^{PEDV}$ mutants are impaired in their recognition and cleavage of mono-Ub as well as Ub-chains.

Past MHV PLP2 mutants have shown temperature-dependent inactivation [54, 101]. While $H101R^{FIPV}$ and $N101R^{PEDV}$ mutants are DUB deficient, it is possible that this mutation may alter protein stability at physiological temperatures of the cell. To test for this, the turnover rates of these mutant enzymes for RLRGG-AMC were assayed compared to the wild-type after incubation from 0-60 minutes at 30°C and 37°C. The results expressed as percent remaining activity, shown in Figure 5.19, indicate that these mutants are not temperature sensitive. We observed a minor decrease in activity for the $N101R^{PEDV}$ mutant (74% remaining activity) after 1 hour at 37°C. Given that these $H101R^{FIPV}$ and $N101R^{PEDV}$ mutants are active at high temperatures and drastically deficient in their DUB activity compared to the wild-type, they are great candidates for attenuating FIPV and PEDV. Further studies using the reverse genetics system are necessary to determine if these FIPV and PEDV mutants effectively attenuate pathogenesis and are still sufficient for virus replication for vaccine design.

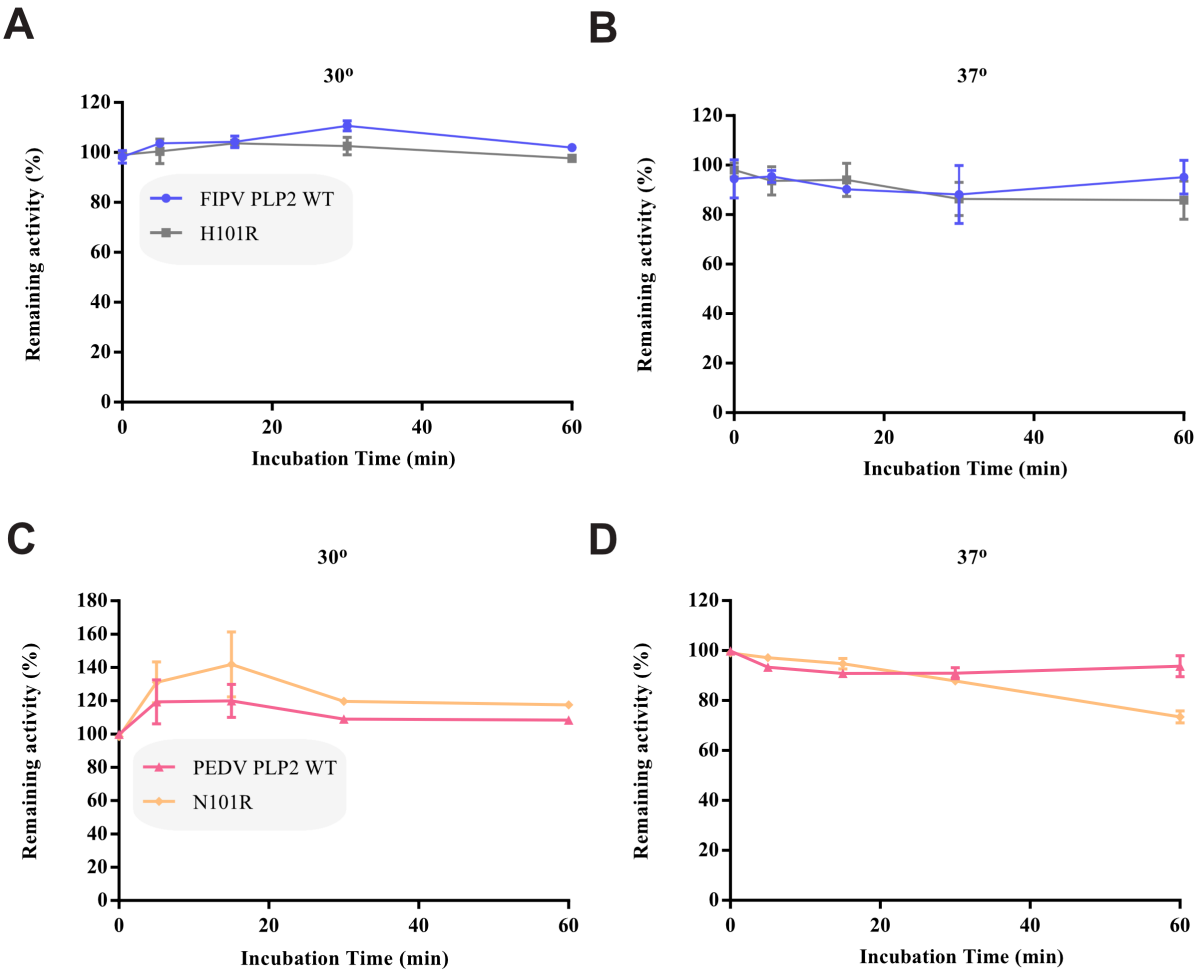


Figure 5.19 H101R^{FIPV} and N101R^{PEDV} maintain their RLRGG-AMC activity at physiological temperatures

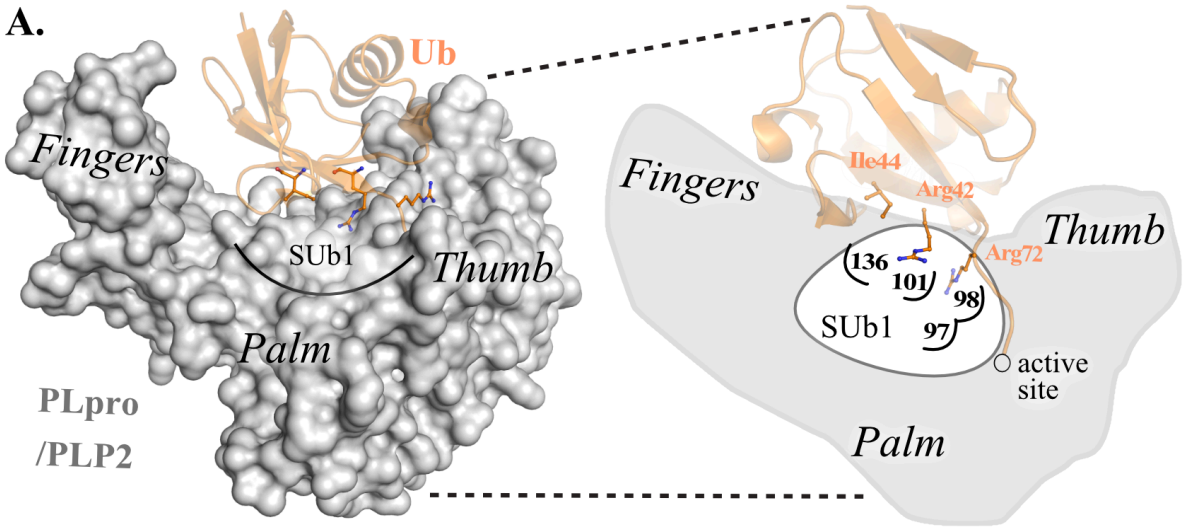
(A, B) Normalized % remaining activity of FIPV PLP2 wild-type (blue circles) and H101R^{FIPV} (gray squares) as a function of incubation time (0-60 minutes) at 30°C (A) and 37°C (B). (C,D) PEDV PLP2 wild-type (pink triangles) and N101R^{PEDV} (orange diamonds) normalized % remaining activity plotted against incubation time (0-60 minutes) at 30°C (C) and 37°C (D). Data were normalized by determining the ratio of the relative turnover number at time = t and time = 0 at 50 μM RLRGG-AMC for each enzyme.

5.4 Summary

In Chapter 5, we provide important structural, biochemical, and kinetic data to better understand αCoV FIPV and PEDV PLP2 substrate specificity and recognition of Ub versus ISG15. Both FIPV and PEDV PLP2 are efficient DUBs *in vitro* and greatly prefer Ub over ISG15. The structure of PEDV PLP2 revealed a unique zinc-binding mode suggesting that αCoV PLP2s utilize different Zn-finger coordinating residues than βCoVs. In addition, the Zn-finger of PEDV PLP2 is likely unique among the αCoVs. Using our structure, we were able to generate a FIPV PLP2 model and Ub-bound models to identify residues of the SUb1 recognition site, which are critical for Ub recognition and catalysis. We show that engineering an arginine at position 101 is one strategy that can be used to disrupt DUB activity in both αCoV PLP2s, and may also reduce deISG activity, without affecting PLP2s' ability to cleave a small peptide substrate, even at physiological temperatures. These residues are near Arg42 as well as the putative hydrophobic patch of Ub, where most PLPs bind Ub. Therefore, we identify a 'hot spot' in the Ub-binding site of α/βCoV PLPs that can be used to selectively remove their DUB activity. Mutations of the PLP2 'hot spot' may be useful in attenuating FIPV and PEDV for vaccine design.

CHAPTER 6. SUMMARY

In this dissertation, we investigate the substrate specificity of PLpro/PLP2 across the alpha- and beta-CoV genus towards host Ub and ISG15. We determined that alpha-CoV FIPV and PEDV PLP2 are efficient DUBs while beta-CoVs, MERS- and SARS-CoV are efficient at cleaving both Ub and ISG15 substrates. Most PLPs possess one single ubiquitin recognition (SUbl) site for Ub-chains, and ISG15 in the case of MERS-CoV, while only SARS-CoV PLpro, and possibly TGEV PLP1, have a second ubiquitin (SUb2) subsite. We showed that it is possible to disrupt Ub-binding by mutating residues of the SUb1 subsite at equivalent positions across the beta- and alpha-CoV



B.

Position	Promote DUB	Disrupt DUB
97/98	Glu	Arg
101	Neutral polar	Arg
136	Hydrophobic	Arg/Lys

Figure 6.1 Molecular fingerprint to remove PLP DUB activity

(A) PLP is shown as a surface representation next to a cartoon drawing with positions of the SUb1 labeled. Ub (orange, PDB code 4RF1) is shown as a cartoon model with Ile44 and Arg42 shown as sticks, which are residues involved in PLP-mediated Ub recognition. The position P5, Arg72, is also shown as sticks. (B) Table summarizing the hot-spots for mutation in panel A that promote or disrupt DUB activity from mutagenesis studies. Residue positions correspond to the PEDV PLP2 (PDB code 6NOZ) crystal structure.

clade of PLPs, revealing hot spots for mutagenesis (Figure 6.1). These hot spots are located at position 97, 98, and 101 in α6 of the thumb domain and position 136 in β5 of the palm (refer to Figure 5.12).

Although the amino acid may vary across different PLPs, in general from the collection of mutagenesis studies on PLPs, inserting a Glu at position 97 and 98 promotes a favorable interaction with Arg42 of Ub for DUB activity while mutating either position to an Arg completely disrupts DUB function (Figure 6.1). In the case of SARS-CoV, Arg97 is already adopted and substituting this residue to a Glu resulted in DUB hyperactivity (nomenclature Arg167 in Chapter 3). The same trend was observed for the Arg98Glu mutant in MERS-CoV PLpro (nomenclature Arg1649 in Chapter 4). The results from the arginine charge flip in SARS- and MERS-CoV supports the fact that Arg is less preferred at position 97 and 98.

In the case of MHV PLP2, an Ile97Arg mutant (nomenclature Ile249Arg in [102]) was able to reduce DUB activity of the enzyme while protease activity was maintained. Another mutant in MHV, Asp98Arg (nomenclature Asp250Arg in [102]) showed slightly reduced protease activity while DUB activity was significantly impaired compared to the wild-type, as this residue forms a direct interaction with Arg42 of Ub in the MHV PLP2-Ub bound structure. At the same time, both Ile97Arg and Asp98Arg also resulted in an enhancement in deISG activity. Since Trp123 of ISG15 resides in the equivalent Arg42 position, it is possible that substituting in an Arg at this region may promote a more favorable interaction with ISG15 for deISG activity.

Although 97 and 98 have shown to be important for DUB/deISG activities, in many cases these positions may be close enough to the active site to effect protease function of PLPs. In SARS, Glu98 makes a salt bridge with the position P5 of Arg72 [37]. MERS-CoV PLpro Arg98 also forms a guanidium stacking interaction with Arg72. However, a Glu at this position is preferred for protease activity. In the case of MHV, Asp98 does not interact with the P5 position Arg72 and

is less important for protease recognition. In the case of alpha-CoV FIPV and PEDV PLP2, and likely alpha-CoV clade of PLPs, Glu97 residue is conserved and also important for protease function, as substitution of this residue to an Arg drastically reduced both protease and DUB activities.

For study of new PLPs, positions 101 and 136 of the SUb1 subsite, rather than position 97-98, may serve as better candidates for mutagenesis to reduce DUB activity while not effecting protease function. Position 101 is often a neutral polar residue in PLPs while position 136 is usually hydrophobic residue. Both these positions have been shown to interact with Ile44 of Ub hydrophobic patch, which is close in proximity to Arg42 of Ub (Figure 6.1). Therefore, mutation of either these positions to a positive charged residue i.e. Arg is one strategy to reduce DUB activity by steric block and charge repulsion. From mutagenesis work, $Val136^{MERS}$ (nomenclature Val1691 in Chapter 4), $Phe136^{MHV}$ (nomenclature Phe290 in [102]), Met^{SARS} (nomenclature Met209 in Chapter 3), and $Ala136^{FIPV}$ (refer see Chapter 5) are all important for Ub-binding. Specially, $Val136Arg/Lys^{MERS}$ and $Ala136Arg/Lys^{FIPV}$ mutants were drastically DUB deficient while protease activity is less effected. A similar trend was observed for $His101Arg^{FIPV,MERS}$ $Asn101Arg^{PEDV}$. In the case of MERS, inserting a positive or negative charge at positions His101 and Val136 can also reduce deISG activity by clashing with the small hydrophobic pocket of ISG15 consisting of Trp123 and Pro130 (see Chapter 4). All in all, position 101 and 136 are hot spots for reducing DUB activity across beta-CoV and alpha-CoV genera.

From these mutagenesis data, we were able to develop a molecular fingerprint of how PLPs recognize Ub. We also provide predictions of how these sites may influence ISG15 specificity and active site functionality. Inserting an Arg at position 101 and 136 appear to be the best strategy for obtaining a DUB deficient PLP enzyme, which would be great candidate sites for suppressing CoV pathogenesis for live attenuated vaccines. This molecular fingerprint will be incredibly useful for

rapidly engineering PLPs from new emerging CoVs. Although subtle differences between PLPs are inevitable, this dissertation provides a guide to remove the DUB function of new PLPs. Future studies are required to test how these DUB deficient or DUB/deISG deficient mutants affect host antagonism in cells as well as viral replication and pathogenesis in an animal model using reverse genetics system.

APPENDIX A. TABLE OF PLP CONSTRUCTS

CoV	Strain	GenBank	Protein	Pp1ab start	Pp1ab end
MERS	*Human betacoronavirus 2c EMC/2012*	AFS88944.1	Ubl2-PLpro	1484	1802
			PLpro	1544	1801
SARS	*SARS coronavirus Urbani*	AAP13442.1	Ubl2-PLpro	1541	1855
PEDV	*WSU-79/1146*	Q98VG9.2	Ubl2-PLP2	1488	1811
			PLP2	1558	1794
FIPV	*USA/Indiana34/2013*	AID56702.1	PLP2	1688	1933

After TEV protease cleavage, for MERS, FIPV, and PEDV, Ser and Asn are left at the N-terminus of the protein construct.

APPENDIX B. CHARACTERIZING THE ENZYMATIC ACTIVITY OF PLP WILD-TYPE AND MUTANTS

As discussed earlier, PLPs are multifunctional enzymes. The native substrates of PLPs include portions of the viral polyprotein with the LXGG motif sequence, Ub-conjugated proteins (i.e. K48-linked IκBα, K63-linked STING, TBK1, TRAF3), and ISG15-conjugated proteins. To characterize PLPs protease, DUB, and deISG activities in an *in vitro* setting, different Ub and Ubl tools were used as substrates or inhibitors, which is described in this section (Figure B.1). These are commonly used in the literature to characterize the enzymatic function of DUB and deISG enzymes.

For kinetic experiments, commercially available FRET-based substrates, Z-RLRGG-AMC, Ub-AMC, and ISG15-AMC, were used to estimate protease, DUB, and deISG activity, respectively, of PLPs (Figure B.1.A,B). Z-RLRGG-AMC is a short P5-P1 sequence of the C-terminal tail of Ub/ISG15, which also contains a N-terminal Z protective group. PLPs cleave the peptide bond of these FRET-based substrates to release the AMC fluorophore. Fluorescence of AMC can be easily monitored over time to determine kinetic parameters (k_{cat}, K_m) of PLPs and the catalytic efficiency k_{cat}/K_m can be calculated. Currently, only the human form of ISG15 is commercially available; therefore, to evaluate the binding of different ISG15 species other experiments are necessary.

Product-based inhibitors or the free forms of Ub/ISG15 were used for competitive inhibition assays to determine IC_{50} and calculate binding affinity (K_i) of PLPs towards free substrates (Figure B.1.A). In these assays, different species of ISG15s can be designed. Furthermore, free product forms of Ub/ISG15 can be used for obtaining crystal structures of noncovalent PLP-Ub and PLP-ISG15 complexes. These complexes are generated by mutating the

catalytic Cys of PLPs to a Ser to promote interaction with the carboxy terminus of Ub/ISG15. These complexes mimic the transient "EQ" state in the catalytic mechanism (Figure 1.9).

The cleavage of PLPs towards Ub-chains can be used to assay catalysis of the native isopeptide bond (Figure B.1,A,C). Di-Ub panels were used to survey substrate specificity of PLPs towards different Ub-chain linkages, including K6-, K11-, K29-, K33-, K48-, K63-, as well as M1 (linear)-linked Ub_2. K48- and K63-linked tetra-Ub was used to monitor PLPs cleavage towards extended poly-Ub chains for insight into the mechanism of Ub-chain processing. Although these Ub-chains display the physiological isopeptide bond, they are missing the target protein, which could potentially alter PLPs specificity and recognition for Ub-chains. In general, PLPs specificity for target proteins has been is not well characterized in the literature. In cell-based assays, PLPs remove different ubiquitinated and ISGylated proteins in the cell. Therefore, it is unclear if PLPs exhibit specificity and selectivity for certain targets in the innate immune response.

Lastly, Ub and Ubl-probes that contain the propargylamine (PA) warhead at the C-terminus were used react with catalytic Cys of PLPs (Figure B.1.A,D). Ubl-probes are easily purified using the chitin-intein system, a general method in the literature described by Wilkinson et al. in 2005 [134] (Appendix D). These PA probes can be used for gel-based assays to determine reactivity of PLPs for different Ubl proteins and compare the wild-type reactivity to selected PLP mutants. Additionally, PA probes can be utilized for generating covalent PLP-Ub and PLP-ISG15 complexes, which mimic the acyl-enzyme intermediate or "F" state in the catalytic mechanism (Figure 1.9). MERS-CoV PLpro-ISG15 and PEDV PLP2-Ub complexes were successfully crystallized by reacting PLPs with their respective modifier (Chapter 4 and Appendix D). These Ubl-probes are customizable for different enzymes. For example, the Ubl protein can be changed based on the context (i.e. different species of ISG15) and the probe itself can be altered if enzymes exhibit low reactivity. Ekkebus et al. describes the influence of the probe length on ubiquitin

carboxyl-terminal hydrolase isozyme L3 (UCHL3) ability to react with Ub probes [155]. The PA probe was chosen in our studies because PA is inexpensive and preferentially react with many DUBs based on the literature. We used mono-Ub-PA and focused on PLPs that have one ubiquitin recognition subsite (SUbl). Recently, nonhydrolyzable di-Ub probes have been developed to assay for SUb1-SUb1' specificity (i.e. the in-between di-Ub probe) as well as SUb2-SUb1 (i.e. S2 pocket di-Ub probe [160]. The latter was used to obtain SARS-CoV PLpro bound with K48-linked di-Ub [81]. Rather than an isopeptide bond, the S2 pocket probe has a triazole linkage that is protease-resistant.

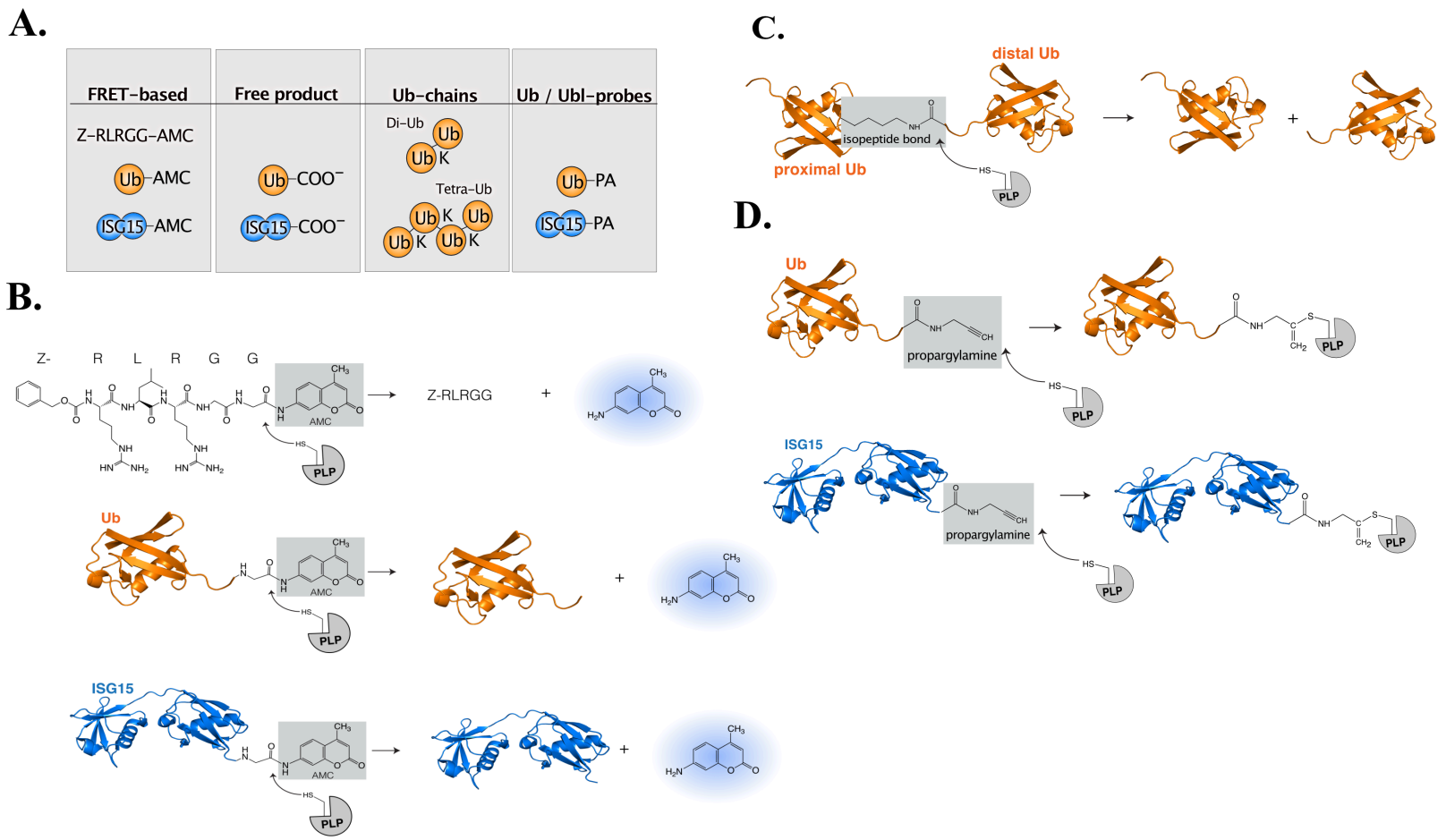


Figure B.1 Summary of Ub/Ubl tools used to characterize PLP enzymatic function

(A) FRET-based, free product, Ub-chains, and Ub/Ubl-probes used to assay protease activity are listed. (B) Enzymatic action of PLP towards commercial FRET-based substrates with the fluorogenic 7-amino-4-methylcourmin group at the C-terminus, which from top to bottom are used to estimate protease (Z-RLRGG-AMC), DUB (Ub-AMC), deISG (ISG15-AMC) activity. (C) Mechanism of PLP cleaving Ub-chains displaying physiological isopeptide bond. (D) PLP reactivity with Ub and ISG15-PA probes displaying the warhead, propargylamine.

APPENDIX C. SCHEMATIC OF PLP MUTANT DESIGN AND CHARACTERIZATION

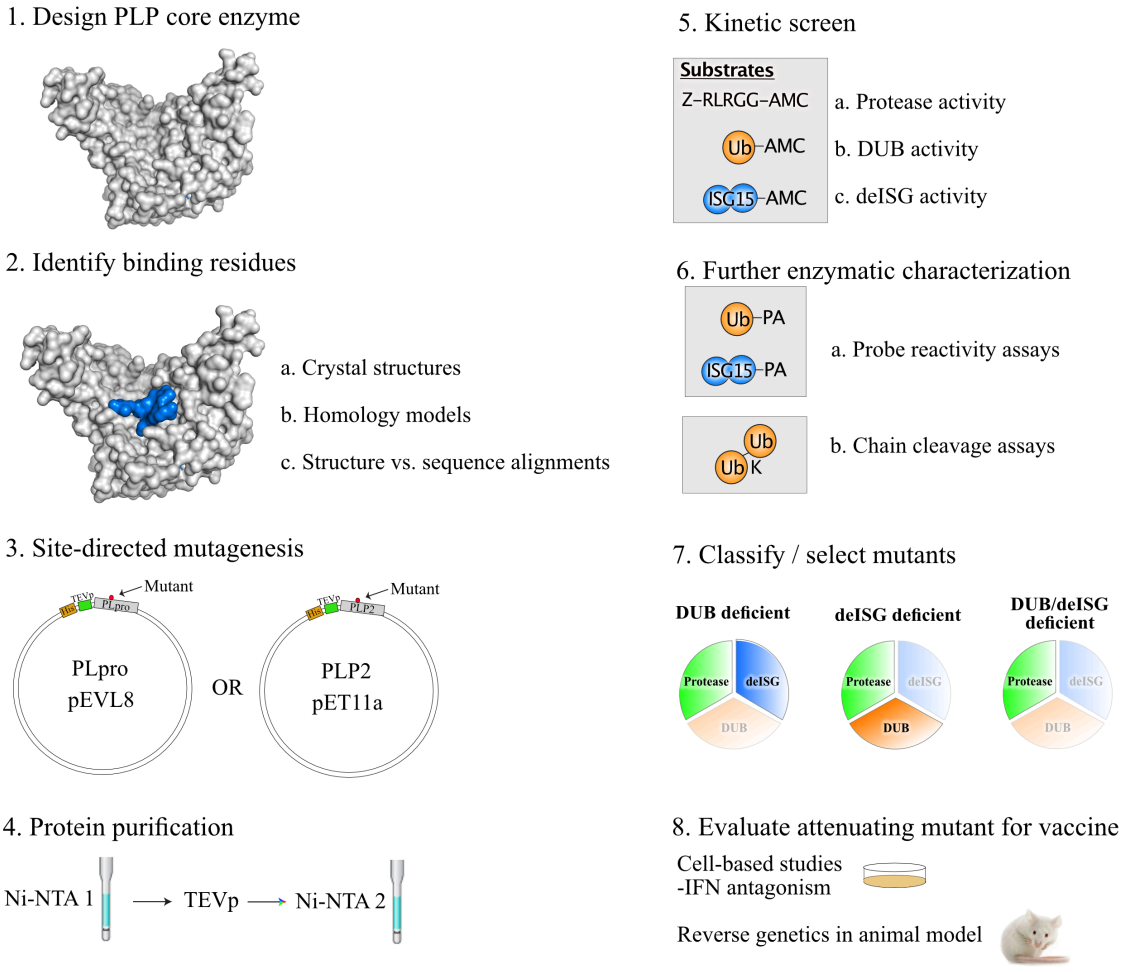


Figure C.1 Schematic of the design, purification, and kinetic screening process for PLP mutant selection and assay for future work

(1) First, the core PLP catalytic domain was designed and (2) Ub/ISG15 binding residues were identified based on (a-c), including structural and sequence alignments (3) Next, specific primers were designed to insert the mutation into the PLP gene via site-directed mutagenesis. (4) Proteins were purified using Ni-NTA gravity flow columns, and untagged PLP2 was used for an initial kinetic screen. (5) The turnover number of each mutant was determined for three FRET-based substrates, RLRGG-AMC (a), Ub-AMC (b), and ISG15-AMC (c) (6) To determine kinetic parameters of mutants, a complete kinetic characterization across varying Ub-AMC and hISG15-AMC was performed. Deficient mutants were further characterized with probe reactivity and chain cleavage assays. (7) Finally, selected mutants were classified as DUB, deISG, or both DUB/deISG deficient based on (5-6). (8) Future directions of this work involve testing mutants in each category for IFN antagonism and evaluate attenuating mutants in an animal model using reverse genetics.

APPENDIX D. SYNTHESIS OF UB/ISG15-PA PROBES

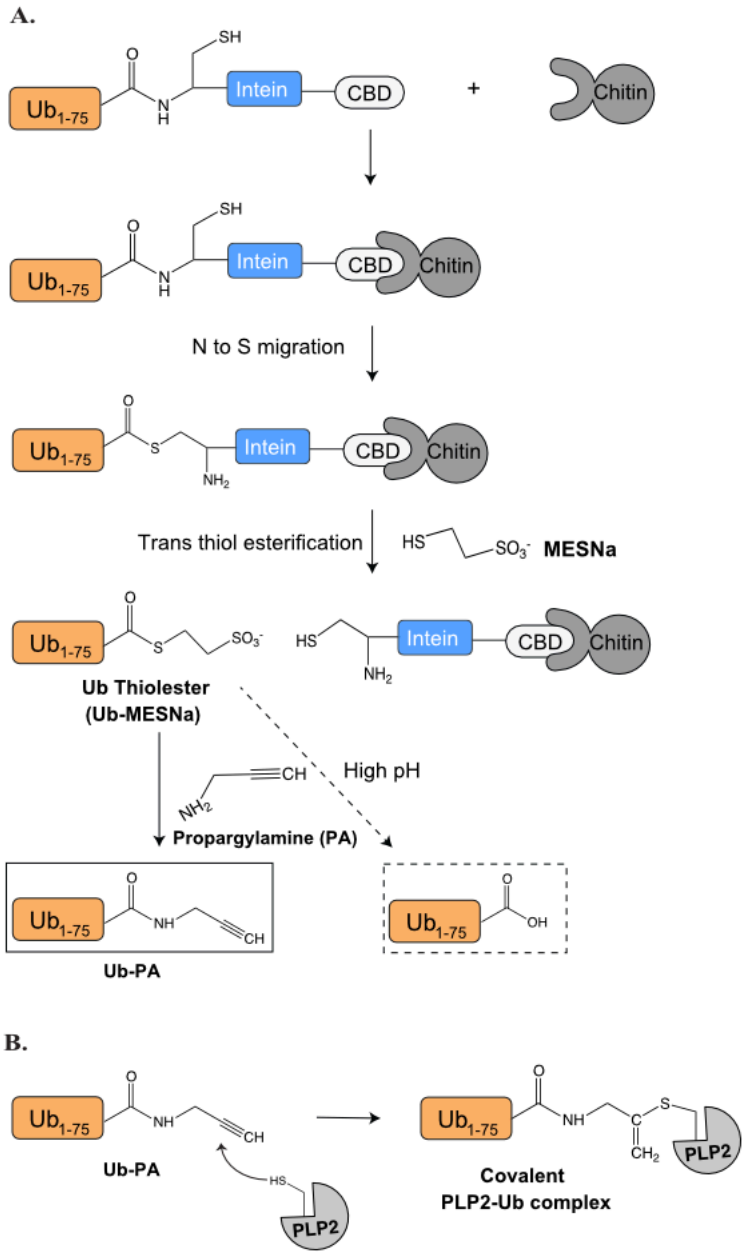


Figure D.1 Schematic of Ub-MESNa purification and generating Ub-PA for reacting with PLP enzymes

(A) Ub_{1-75}, feline $ISG15_{1-155}$, porcine $ISG15_{1-154}$ -intein-chitin binding domain (CBD) were expressed and bound to chitin beads, eluted with MESNa, reacted with PA at high pH according to the chitin-intein system described by [134] and in section 5.2.9. (B) PA probes react with C-terminal alkyne of the PA to generate the covalent vinyl thioether complex.

APPENDIX E. PURIFICATION AND CRYSTALLIZATION OF PEDV PLP2-UB COMPLEX

E.1 Purifying Ub-PA, PLP2, and generating the PEDV PLP2-Ub complex

First, Ub_{1-75}-intein-CBD domain was expressed and purified using the intein-chitin binding domain system as described in 5.2.9 (see procedure in Figure D.1). From a 1 L culture (13.9 g pellet), >100 mg was generated for the reaction with PLP2. Untagged PEDV PLP2 protein was expressed (1 L culture; 6.8 g pellet) and purified with the exact produced described in section 5.2.5. Approximately 20 mg of untagged PLP2 was generated for reaction with Ub-PA. Fresh PEDV PLP2 was used to react with freshly generated unfrozen Ub-PA at a 1:5 ratio in MonoQ buffer A (10 mM Tris, pH 9, 5% glycerol, 10 mM BME). Approximately 20 mg and 28 mg of PLP2 and Ub-PA was used, respectively. The reaction was done in a 65 mL volume using a low PLP2

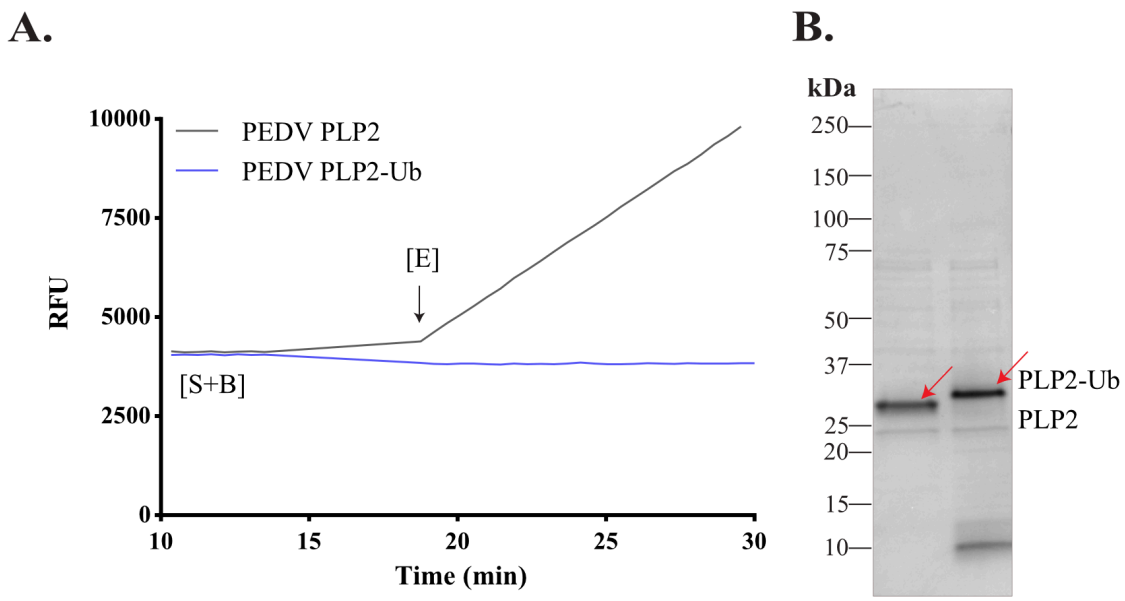


Figure E.1 PEDV PLP2-Ub complex was successfully generated

(A) Time course fluorescence curve of the hydrolysis of active PLP2 (black) versus PLP2-Ub complex (blue) against 50 μM of RLRGG-AMC. Enzyme was loaded at 0.8 μM. (B) SDS-PAGE analysis reveals full conversion of PLP2-Ub complex with excess Ub-PA at the bottom of the gel. Protein was loaded at 5 μM on a 4-20% Mini-PROTEAN TGX gel from Biorad.

concentration of 10 μM to prevent the protein from interacting with itself and forming oligomers with excess Ub-PA at 50 μM. The activity assay and SDS-PAGE gel confirmed that there was complete conversion and reaction of PLP2 with Ub-PA to form the covalent adduct, PLP2-Ub complex (Figure E.1).

After confirming the generation of PLP2-Ub complex, the reaction was dialyzed for another 48 hours and was then loaded onto an 8 mL MonoQ column at 2 mL/min equilibrated with MonoQ buffer A. The column was then washed with 5 CV of buffer A before eluting the complex with a 0-20% buffer B (10 mM Tris, pH 9, 5% glycerol, 500 mM NaCl) gradient across 120 mL at 2.0 mL/min while collecting 5 mL fractions. There was a small peak (purple arrow, Figure E.2) at 16% buffer B then a sharp peak (blue shading, Figure E.2) at 20% buffer B. Then, 20 mL of 30% B was passed over the column, which produced another sharp peak (purple shading). Finally, all other proteins were eluted with 100% buffer B over 20 mL (orange shading). The resulting chromatogram and SDS-PAGE analysis of fractions is shown in Figure E.2 with the final pool. Free Ub-PA was observed in the flow-through and some unreacted PLP2 came off at the 30% buffer B wash while other protein contaminants were observed in the 100% buffer B wash. A summary gel showing the purification of PLP2, reaction with Ub-PA, and purification of the PEDV PLP2-Ub complex is shown in Figure E.3. Interestingly, as observed with unliganded PEDV PLP2, higher order oligomer bands were observed for the complex on the SDS-PAGE gel.

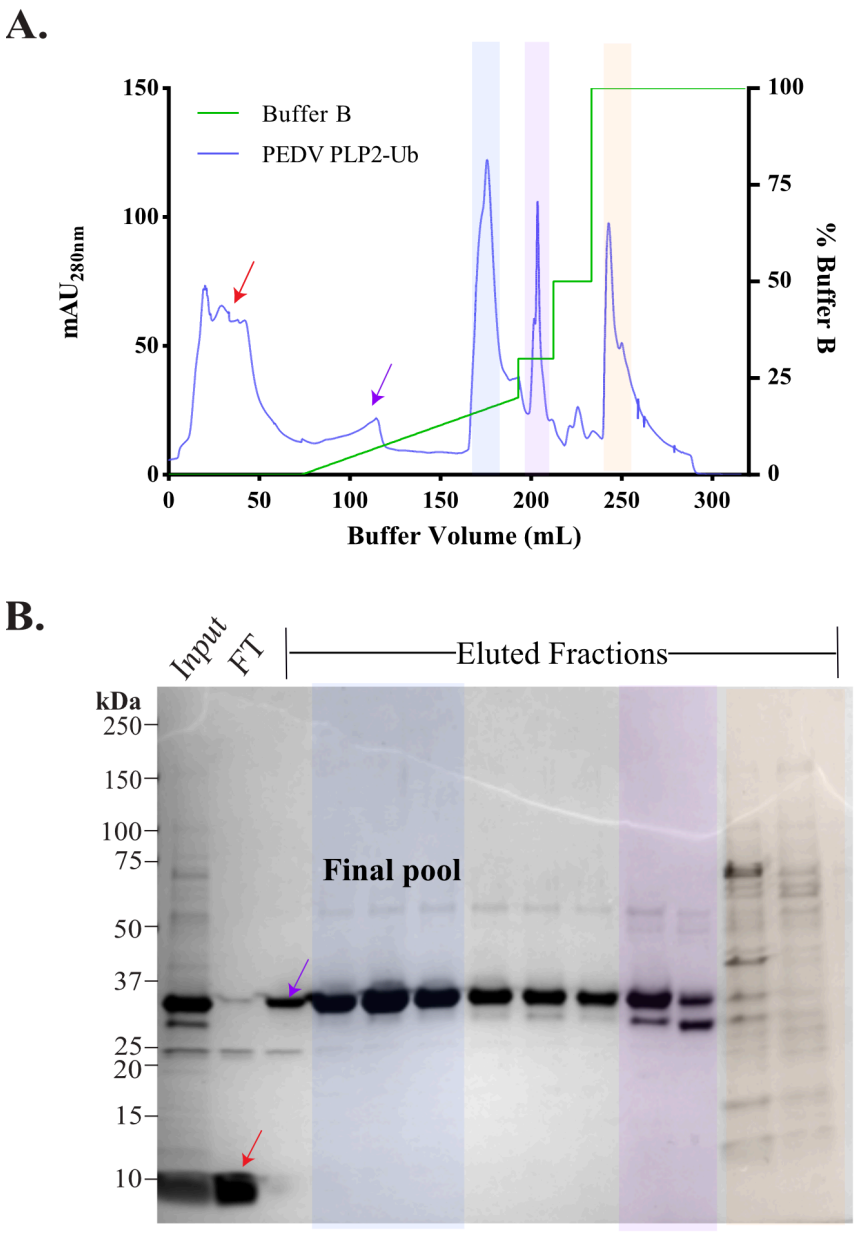


Figure E.2 Purification of PEDV PLP2-Ub complex

(A) FPLC chromatogram plotting buffer volume as a function of milli-absorbance units (mAU) at 280 nm as well as % buffer B. The green line shows the gradient of buffer B, including the linear gradient where the complex was eluted. Red arrow, injection; purple arrow, eluted fraction in panel B. (B) SDS-PAGE analysis of PEDV PLP2-Ub and eluted proteins. Samples were loaded at 10 µg if able. Final pool fractions are indicated, and protein shading corresponds to peaks in panel A.

The final pool was buffer exchanged into storage buffer containing 10 mM Tris, pH 7.5, 100 mM NaCl, 5% glycerol, 10 mM DTT and concentrated using 10 kDa MWCO Amicon device to 30.6 mg/mL PLP2-Ub. Approximately 5 mg of PLP2-Ub complex was generated for crystallization. Protein was screened for crystallization conditions, using MCSG screens 1-3 and Anions from Qiagen, at a 1:1 ratio with three protein concentrations (15 mg/mL, 10 mg/mL, and 5 mg/mL) using 200 nL drop sizes on the Mosquito®Crystal liquid handling robot (TTP Labtech).

Needles of the PEDV PLP2-Ub complex grew within a week at room temperature in similar conditions in the MCSG-2 screen, which both contained the salt, 0.2 M calcium chloride, and buffer 0.1 M Tris, 8.5. One condition contained 5% more PEG4000 (25% compared to 20%). Needles were observed in subwells that contained 15 mg/mL and 10 mg/mL complex. Several optimization attempts were made in a 24-well format at 4°C. Very thin but long (up to 500 μm) needles grew from 20-22% PEG4000 (the same salt/buffer in the initial condition) with 10 mg/mL protein in a 2 μL dropsize (keeping a 1:1 ratio). The same type of very thin needles also grew at lower PEG4000 concentrations down to 18% after one month incubation (Figure E.3). Further optimization was done by lowering the protein concentration to 2.5 mg/mL at the same PEG4000 gradient while other optimization attempts were done lowering the PEG4000 gradient but keeping the same 10 mg/mL protein concentration. Both of these optimizations produced nice 3D cubed crystals, ~50 μm thick after about a week. An example crystal is shown in Figure E.3 from condition 0.1 M Tris, 8.5, 0.2 M calcium chloride, 12% PEG4000. The best crystals so far have been observed in crystallization condition 0.1 M Tris, 8.5, 0.2 M calcium chloride, and 20% PEG4000 at 2.5 mg/mL complex. These crystals were looped, briefly soaked in reservoir supplemented with 20% glycerol, before plunged into liquid nitrogen for storage. Data will be collected on these crystals at LRL-CAT.

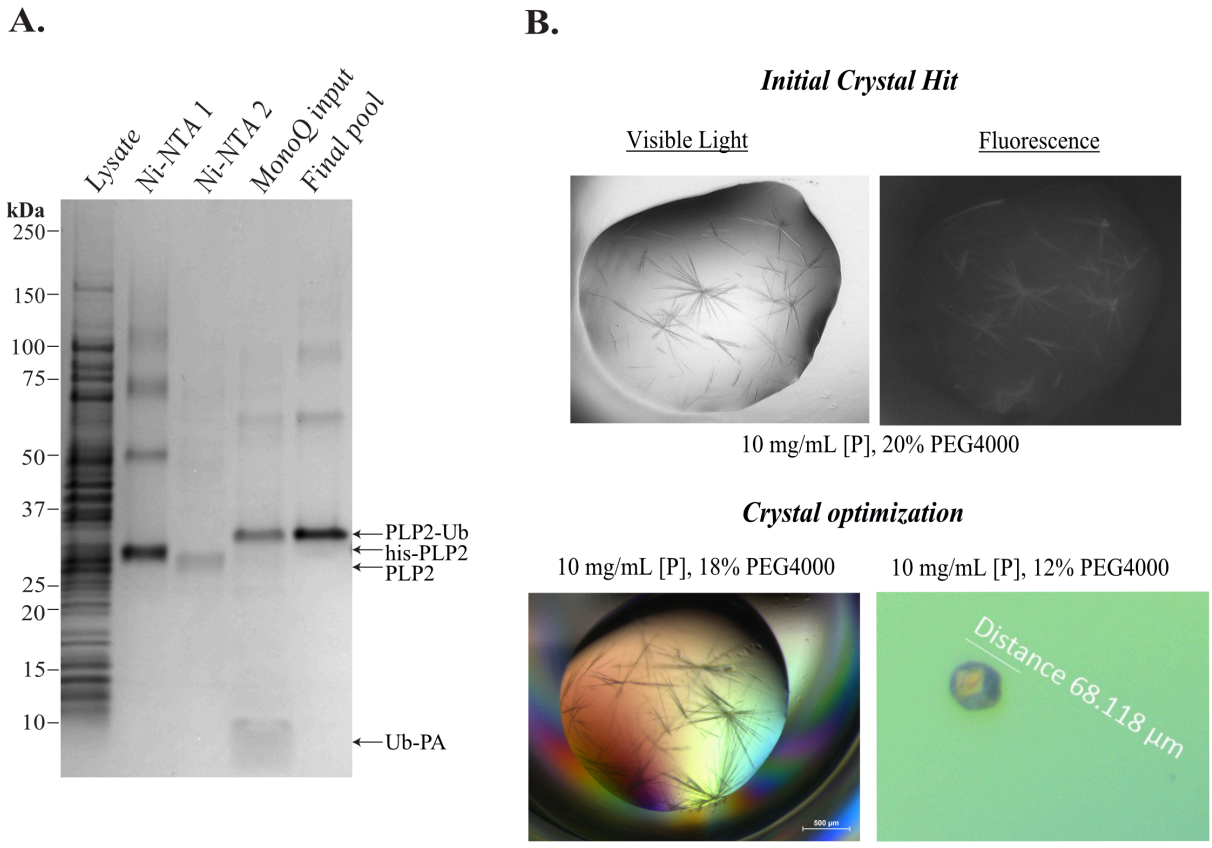


Figure E.3 Summary gel and crystallization hits of PEDV PLP2-Ub complex

(A) SDS-PAGE analysis showing PLP2 purification after Ni-NTA and final pool of PLP2-Ub complex obtained after MonoQ column. Lysate was loaded at 5 µg while other lanes were loaded at 1 µg. (B) Light and fluorescence images of initial crystal hits (top images) in 96-well format versus crystals after optimization at 24-well format. Needles were optimized to 3D crystals by lowering the w/v % of PEG4000.

REFERENCES

[1] C. f. D. C. a. P. (CDC), "Severe acute respiratory syndrome (SARS) and coronavirus testing--United States, 2003," *MMWR Morb Mortal Wkly Rep,* vol. 52, no. 14, pp. 297-302, Apr, 2003.

[2] Y. Guan *et al.*, "Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China," *Science,* vol. 302, no. 5643, pp. 276-8, Oct, 2003.

[3] V. C. Cheng *et al.*, "Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection," *Clin Microbiol Rev,* vol. 20, no. 4, pp. 660-94, Oct, 2007.

[4] M. L. Baker, T. Schountz, and L. F. Wang, "Antiviral immune responses of bats: a review," *Zoonoses Public Health,* vol. 60, no. 1, pp. 104-16, Feb, 2013.

[5] A. P. Dobson, "Virology. What links bats to emerging infectious diseases?," *Science,* vol. 310, no. 5748, pp. 628-9, Oct, 2005.

[6] M. Aly *et al.*, "Occurrence of the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) across the Gulf Corporation Council countries: Four years update," *PLoS One,* vol. 12, no. 10, pp. e0183850, 2017.

[7] X. Chen *et al.*, "Comparative epidemiology of Middle East respiratory syndrome coronavirus (MERS-CoV) in Saudi Arabia and South Korea," *Emerg Microbes Infect,* vol. 6, no. 6, pp. e51, Jun, 2017.

[8] S. R. Bialek *et al.*, "First confirmed cases of Middle East respiratory syndrome coronavirus (MERS-CoV) infection in the United States, updated information on the epidemiology of MERS-CoV infection, and guidance for the public, clinicians, and public health authorities - May 2014," *MMWR Morb Mortal Wkly Rep,* vol. 63, no. 19, pp. 431-6, May, 2014.

[9] V. M. Corman *et al.*, "Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat," *J Virol,* vol. 88, no. 19, pp. 11297-303, Oct, 2014.

[10] J. F. Chan *et al.*, "Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease," *Clin Microbiol Rev,* vol. 28, no. 2, pp. 465-522, Apr, 2015.

[11] C. B. Reusken *et al.*, "Cross host transmission in the emergence of MERS coronavirus," *Curr Opin Virol,* vol. 16, pp. 55-62, Feb, 2016.

[12] I. Eckerle *et al.*, "Replicative Capacity of MERS Coronavirus in Livestock Cell Lines," *Emerg Infect Dis,* vol. 20, no. 2, pp. 276-9, Feb, 2014.

[13] A. R. Fehr, and S. Perlman, "Coronaviruses: an overview of their replication and pathogenesis," *Methods Mol Biol,* vol. 1282, pp. 1-23, 2015.

[14] K. A. Ivanov *et al.*, "Major genetic marker of nidoviruses encodes a replicative endoribonuclease," *Proc Natl Acad Sci U S A,* vol. 101, no. 34, pp. 12694-9, Aug, 2004.
[15] R. L. Graham, E. F. Donaldson, and R. S. Baric, "A decade after SARS: strategies for controlling emerging coronaviruses," *Nat Rev Microbiol,* vol. 11, no. 12, pp. 836-48, Dec, 2013.
[16] G. Gerna *et al.*, "Genetic variability of human coronavirus OC43-, 229E-, and NL63-like strains and their association with lower respiratory tract infections of hospitalized infants and immunocompromised patients," *J Med Virol,* vol. 78, no. 7, pp. 938-49, Jul, 2006.
[17] C. Geller, M. Varbanov, and R. E. Duval, "Human coronaviruses: insights into environmental resistance and its influence on the development of new antiseptic strategies," *Viruses,* vol. 4, no. 11, pp. 3044-68, Nov, 2012.
[18] P. Zhou *et al.*, "Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin," *Nature,* vol. 556, no. 7700, pp. 255-258, Apr, 2018.
[19] D. Song, and B. Park, "Porcine epidemic diarrhoea virus: a comprehensive review of molecular epidemiology, diagnosis, and vaccines," *Virus Genes,* vol. 44, no. 2, pp. 167-75, Apr, 2012.
[20] Q. Chen *et al.*, "Isolation and characterization of porcine epidemic diarrhea viruses associated with the 2013 disease outbreak among swine in the United States," *J Clin Microbiol,* vol. 52, no. 1, pp. 234-43, Jan, 2014.
[21] M. B. Pensaert, and P. de Bouck, "A new coronavirus-like particle associated with diarrhea in swine," *Arch Virol,* vol. 58, no. 3, pp. 243-7, 1978.
[22] S. Wang *et al.*, "Classification of emergent U.S. strains of porcine epidemic diarrhea virus by phylogenetic analysis of nucleocapsid and ORF3 genes," *J Clin Microbiol,* vol. 52, no. 9, pp. 3509-10, Sep, 2014.
[23] G. Tekes, and H. J. Thiel, "Feline Coronaviruses: Pathogenesis of Feline Infectious Peritonitis," *Adv Virus Res,* vol. 96, pp. 193-218, 2016.
[24] K. Hartmann, "Feline infectious peritonitis," *Vet Clin North Am Small Anim Pract,* vol. 35, no. 1, pp. 39-79, vi, Jan, 2005.
[25] Y. Cong, P. Verlhac, and F. Reggiori, "The Interaction between Nidovirales and Autophagy Components," *Viruses,* vol. 9, no. 7, 07, 2017.
[26] K. A. Mihindukulasuriya *et al.*, "Identification of a novel coronavirus from a beluga whale by using a panviral microarray," *J Virol,* vol. 82, no. 10, pp. 5084-8, May, 2008.
[27] P. C. Woo *et al.*, "Discovery of a novel bottlenose dolphin coronavirus reveals a distinct species of marine mammal coronavirus in Gammacoronavirus," *J Virol,* vol. 88, no. 2, pp. 1318-31, Jan, 2014.
[28] K. Mardani *et al.*, "Infectious bronchitis viruses with a novel genomic organization," *J Virol,* vol. 82, no. 4, pp. 2013-24, Feb, 2008.

[29] A. Chamings *et al.*, "Detection and characterisation of coronaviruses in migratory and non-migratory Australian wild birds," *Sci Rep,* vol. 8, no. 1, pp. 5980, Apr, 2018.
[30] Y. Ma *et al.*, "Origin, evolution, and virulence of porcine deltacoronaviruses in the United States," *MBio,* vol. 6, no. 2, pp. e00064, Mar, 2015.
[31] P. C. Woo *et al.*, "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,* vol. 86, no. 7, pp. 3995-4008, Apr, 2012.
[32] X. D. Lin *et al.*, "Extensive diversity of coronaviruses in bats from China," *Virology,* vol. 507, pp. 1-10, 07, 2017.
[33] B. Hu *et al.*, "Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus," *PLoS Pathog,* vol. 13, no. 11, pp. e1006698, Nov, 2017.
[34] S. J. Anthony *et al.*, "Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus," *MBio,* vol. 8, no. 2, 04, 2017.
[35] Y. Yang *et al.*, "Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus," *J Virol,* vol. 89, no. 17, pp. 9119-23, Sep, 2015.
[36] Q. F. Zhang *et al.*, "Morphology and morphogenesis of severe acute respiratory syndrome (SARS)-associated virus," *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai),* vol. 35, no. 6, pp. 587-91, Jun, 2003.
[37] Y. Báez, "Insight into the Substrate Specificity and Inhibition of Human Coronavirus Papain-Like Proteases," Pharmacognosy, University of Illinois at Chicago, 2012.
[38] J. M. White, and G. R. Whittaker, "Fusion of Enveloped Viruses in Endosomes," *Traffic,* vol. 17, no. 6, pp. 593-614, 06, 2016.
[39] T. Takano *et al.*, "TNF-alpha, produced by feline infectious peritonitis virus (FIPV)-infected macrophages, upregulates expression of type II FIPV receptor feline aminopeptidase N in feline macrophages," *Virology,* vol. 364, no. 1, pp. 64-72, Jul, 2007.
[40] Z. Yang *et al.*, "Coronavirus MHV-A59 infects the lung and causes severe pneumonia in C57BL/6 mice," *Virol Sin,* vol. 29, no. 6, pp. 393-402, Dec, 2014.
[41] N. van Doremalen, and V. J. Munster, "Animal models of Middle East respiratory syndrome coronavirus infection," *Antiviral Res,* vol. 122, pp. 28-38, Oct, 2015.

[42] A. M. Mielech *et al.*, "Nidovirus papain-like proteases: multifunctional enzymes with protease, deubiquitinating and deISGylating activities," *Virus Res,* vol. 194, pp. 184-90, Dec, 2014.
[43] B. W. Neuman, "Bioinformatics and functional analyses of coronavirus nonstructural proteins involved in the formation of replicative organelles," *Antiviral Res,* vol. 135, pp. 97-107, 11, 2016.
[44] H. J. Maier *et al.*, "Infectious bronchitis virus generates spherules from zippered endoplasmic reticulum membranes," *MBio,* vol. 4, no. 5, pp. e00801-13, Oct, 2013.
[45] H. J. Maier *et al.*, "Extensive coronavirus-induced membrane rearrangements are not a determinant of pathogenicity," *Sci Rep,* vol. 6, pp. 27126, 06, 2016.
[46] X. Deng *et al.*, "Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages," *Proc Natl Acad Sci U S A,* vol. 114, no. 21, pp. E4251-E4260, 05, 2017.
[47] X. Deng, and S. C. Baker, "An "Old" protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses," *Virology,* vol. 517, pp. 157-163, 04, 2018.
[48] K. Ratia, "Structure, Function, and Inhibition of the Papain-like protease from SARS Coronavirus ", University of Illinois at Chicago, 2008.
[49] J. Lei, Y. Kusov, and R. Hilgenfeld, "Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein," *Antiviral Res,* vol. 149, pp. 58-74, 01, 2018.
[50] A. Putics, A. E. Gorbalenya, and J. Ziebuhr, "Identification of protease and ADP-ribose 1"-monophosphatase activities associated with transmissible gastroenteritis virus non-structural protein 3," *J Gen Virol,* vol. 87, no. Pt 3, pp. 651-6, Mar, 2006.
[51] J. A. Wojdyla *et al.*, "Papain-like protease 1 from transmissible gastroenteritis virus: crystal structure and enzymatic activity toward viral and cellular substrates," *J Virol,* vol. 84, no. 19, pp. 10063-73, Oct, 2010.
[52] X. Hu *et al.*, "Transmissible Gastroenteritis Virus Papain-Like Protease 1 Antagonizes Production of Interferon-," *Biomed Res Int,* vol. 2017, pp. 7089091, 2017.
[53] S. Keep *et al.*, "The ADRP domain from a virulent strain of infectious bronchitis virus is not sufficient to confer a pathogenic phenotype to the attenuated Beaudette strain," *J Gen Virol,* vol. 99, no. 8, pp. 1097-1102, Aug, 2018.
[54] Y. Chen *et al.*, "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,* vol. 290, no. 42, pp. 25293-306, Oct, 2015.

[55] P. Shang *et al.*, "A Naturally Occurring Recombinant Enterovirus Expresses a Torovirus Deubiquitinase," *J Virol,* vol. 91, no. 14, 07, 2017.
[56] P. J. Bonilla, S. A. Hughes, and S. R. Weiss, "Characterization of a second cleavage site and demonstration of activity in trans by the papain-like proteinase of the murine coronavirus mouse hepatitis virus strain A59," *J Virol,* vol. 71, no. 2, pp. 900-9, Feb, 1997.
[57] Z. Chen *et al.*, "Proteolytic processing and deubiquitinating activity of papain-like proteases of human coronavirus NL63," *J Virol,* vol. 81, no. 11, pp. 6007-18, Jun, 2007.
[58] J. Ziebuhr *et al.*, "Human coronavirus 229E papain-like proteases have overlapping specificities but distinct functions in viral replication," *J Virol,* vol. 81, no. 8, pp. 3922-32, Apr, 2007.
[59] M. M. Angelini *et al.*, "Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles," *MBio,* vol. 4, no. 4, Aug, 2013.
[60] M. C. Hagemeijer *et al.*, "Membrane rearrangements mediated by coronavirus nonstructural proteins 3 and 4," *Virology,* vol. 458-459, pp. 125-35, Jun, 2014.
[61] N. Barretto *et al.*, "The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity," *J Virol,* vol. 79, no. 24, pp. 15189-98, Dec, 2005.
[62] H. A. Lindner *et al.*, "The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme," *J Virol,* vol. 79, no. 24, pp. 15199-208, Dec, 2005.
[63] K. Ratia *et al.*, "Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme," *Proc Natl Acad Sci U S A,* vol. 103, no. 15, pp. 5717-22, Apr, 2006.
[64] B. A. Bailey-Elkin *et al.*, "Structure and Function of Viral Deubiquitinating Enzymes," *J Mol Biol,* vol. 429, no. 22, pp. 3441-3470, 11, 2017.
[65] D. Komander, "The emerging complexity of protein ubiquitination," *Biochem Soc Trans,* vol. 37, no. Pt 5, pp. 937-53, Oct, 2009.
[66] X. Jiang, and Z. J. Chen, "The role of ubiquitylation in immune defence and pathogen evasion," *Nat Rev Immunol,* vol. 12, no. 1, pp. 35-48, Jan, 2012.
[67] P. Kumari, and H. Kumar, "Viral deubiquitinases: role in evasion of anti-viral innate immunity," *Crit Rev Microbiol,* vol. 44, no. 3, pp. 304-317, May, 2018.
[68] Y. M. Báez-Santos, S. E. St John, and A. D. Mesecar, "The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds," *Antiviral Res,* vol. 115, pp. 21-38, Mar, 2015.
[69] L. Kong *et al.*, "Structural view and substrate specificity of papain-like protease from avian infectious bronchitis virus," *J Biol Chem,* vol. 290, no. 11, pp. 7160-8, Mar, 2015.

[70] Y. Xing *et al.*, "The papain-like protease of porcine epidemic diarrhea virus negatively regulates type I interferon pathway by acting as a viral deubiquitinase," *J Gen Virol,* vol. 94, no. Pt 7, pp. 1554-67, Jul, 2013.
[71] L. Yuan *et al.*, "p53 degradation by a coronavirus papain-like protease suppresses type I interferon signaling," *J Biol Chem,* vol. 290, no. 5, pp. 3172-82, Jan, 2015.
[72] Y. Ma-Lauer *et al.*, "p53 down-regulates SARS coronavirus replication and is targeted by the SARS-unique domain and PLpro via E3 ubiquitin ligase RCHY1," *Proc Natl Acad Sci U S A,* vol. 113, no. 35, pp. E5192-201, 08, 2016.
[73] L. Cappadocia, and C. D. Lima, "Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism," *Chem Rev,* vol. 118, no. 3, pp. 889-918, 02, 2018.
[74] A. Basters *et al.*, "Structural basis of the specificity of USP18 toward ISG15," *Nat Struct Mol Biol,* vol. 24, no. 3, pp. 270-278, Mar, 2017.
[75] S. G. Devaraj *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,* vol. 282, no. 44, pp. 32208-21, Nov, 2007.
[76] M. Frieman *et al.*, "Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling," *J Virol,* vol. 83, no. 13, pp. 6689-705, Jul, 2009.
[77] K. Ratia *et al.*, "Structural Basis for the Ubiquitin-Linkage Specificity and deISGylating activity of SARS-CoV papain-like protease," *PLoS Pathog,* vol. 10, no. 5, pp. e1004113, May, 2014.
[78] G. Wang *et al.*, "PLP2 of mouse hepatitis virus A59 (MHV-A59) targets TBK1 to negatively regulate cellular type I interferon signaling pathway," *PLoS One,* vol. 6, no. 2, pp. e17192, 2011.
[79] M. A. Clementz *et al.*, "Deubiquitinating and interferon antagonism activities of coronavirus papain-like proteases," *J Virol,* vol. 84, no. 9, pp. 4619-29, May, 2010.
[80] Y. W. Chou *et al.*, "Differential domain structure stability of the severe acute respiratory syndrome coronavirus papain-like protease," *Arch Biochem Biophys,* vol. 520, no. 2, pp. 74-80, Apr, 2012.
[81] M. Békés *et al.*, "Recognition of Lys48-Linked Di-ubiquitin and Deubiquitinating Activities of the SARS Coronavirus Papain-like Protease," *Mol Cell,* vol. 62, no. 4, pp. 572-85, May, 2016.
[82] J. Lei *et al.*, "Crystal structure of the papain-like protease of MERS coronavirus reveals unusual, potentially druggable active-site features," *Antiviral Res,* vol. 109, pp. 72-82, Sep, 2014.

[83] B. A. Bailey-Elkin *et al.*, "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,* vol. 289, no. 50, pp. 34667-82, Dec, 2014.

[84] L. Sun *et al.*, "Coronavirus papain-like proteases negatively regulate antiviral innate immune response through disruption of STING-mediated signaling," *PLoS One,* vol. 7, no. 2, pp. e30802, 2012.

[85] P. B. van Kasteren *et al.*, "Deubiquitinase function of arterivirus papain-like protease 2 suppresses the innate immune response in infected host cells," *Proc Natl Acad Sci U S A,* vol. 110, no. 9, pp. E838-47, Feb, 2013.

[86] D. Zheng *et al.*, "PLP2, a potent deubiquitinase from murine hepatitis virus, strongly inhibits cellular type I interferon production," *Cell Res,* vol. 18, no. 11, pp. 1105-13, Nov, 2008.

[87] A. M. Mielech *et al.*, "MERS-CoV papain-like protease has deISGylating and deubiquitinating activities," *Virology,* vol. 450-451, pp. 64-70, Feb, 2014.

[88] X. Chen *et al.*, "SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex," *Protein Cell,* vol. 5, no. 5, pp. 369-81, May, 2014.

[89] X. Deng *et al.*, "A chimeric virus-mouse model system for evaluating the function and inhibition of papain-like proteases of emerging coronaviruses," *J Virol,* vol. 88, no. 20, pp. 11825-33, Oct, 2014.

[90] X. Yang *et al.*, "Proteolytic processing, deubiquitinase and interferon antagonist activities of Middle East respiratory syndrome coronavirus papain-like protease," *J Gen Virol,* vol. 95, no. Pt 3, pp. 614-26, Mar, 2014.

[91] D. Niemeyer *et al.*, "The papain-like protease determines a virulence trait that varies among members of the SARS-coronavirus species," *PLoS Pathog,* vol. 14, no. 9, pp. e1007296, 09, 2018.

[92] K. Ratia *et al.*, "A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication," *Proc Natl Acad Sci U S A,* vol. 105, no. 42, pp. 16119-24, Oct, 2008.

[93] Y. M. Báez-Santos *et al.*, "X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases," *J Med Chem,* vol. 57, no. 6, pp. 2393-412, Mar, 2014.

[94] Y. M. Báez-Santos *et al.*, "Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus," *J Virol,* vol. 88, no. 21, pp. 12511-27, Nov, 2014.

[95] P. D. Minor, "Live attenuated vaccines: Historical successes and current challenges," *Virology,* vol. 479-480, pp. 379-92, May, 2015.

[96] G. Bijlenga *et al.*, "Development and use of the H strain of avian infectious bronchitis virus from the Netherlands as a vaccine: a review," *Avian Pathol,* vol. 33, no. 6, pp. 550-7, Dec, 2004.
[97] V. Gerdts, and A. Zakhartchouk, "Vaccines for porcine epidemic diarrhea virus and other swine coronaviruses," *Vet Microbiol,* vol. 206, pp. 45-51, Jul, 2017.
[98] S. N. Langel *et al.*, "Lactogenic immunity and vaccines for porcine epidemic diarrhea virus (PEDV): Historical and current concepts," *Virus Res,* vol. 226, pp. 93-107, 12, 2016.
[99] S. J. van Beurden *et al.*, "Recombinant live attenuated avian coronavirus vaccines with deletions in the accessory genes 3ab and/or 5ab protect against infectious bronchitis in chickens," *Vaccine,* vol. 36, no. 8, pp. 1085-1092, 02, 2018.
[100] J. M. Jimenez-Guardeño *et al.*, "Identification of the Mechanisms Causing Reversion to Virulence in an Attenuated SARS-CoV for the Design of a Genetically Stable Vaccine," *PLoS Pathog,* vol. 11, no. 10, pp. e1005215, Oct, 2015.
[101] A. M. Mielech *et al.*, "Murine coronavirus ubiquitin-like domain is important for papain-like protease stability and viral pathogenesis," *J Virol,* vol. 89, no. 9, pp. 4907-17, May, 2015.
[102] Y. Chen, "Structural and Functional studies of the Papain-like protease 2 from Mouse Hepatitis Virus," Purdue University, 2016.
[103] J. R. Clasman *et al.*, "X-ray Structure and Enzymatic Activity Profile of a Core Papain-like Protease of MERS Coronavirus with utility for structure-based drug design," *Sci Rep,* vol. 7, pp. 40292, Jan, 2017.
[104] J. S. Peiris, Y. Guan, and K. Y. Yuen, "Severe acute respiratory syndrome," *Nat Med,* vol. 10, no. 12 Suppl, pp. S88-97, Dec, 2004.
[105] M. Békés *et al.*, "SARS hCoV papain-like protease is a unique Lys48 linkage-specific di-distributive deubiquitinating enzyme," *Biochem J,* vol. 468, no. 2, pp. 215-26, Jun, 2015.
[106] P. Serrano *et al.*, "Nuclear magnetic resonance structure of the N-terminal domain of nonstructural protein 3 from the severe acute respiratory syndrome coronavirus," *J Virol,* vol. 81, no. 21, pp. 12049-60, Nov, 2007.
[107] H. Lee *et al.*, "Inhibitor recognition specificity of MERS-CoV papain-like protease may differ from that of SARS-CoV," *ACS Chem Biol,* vol. 10, no. 6, pp. 1456-65, Jun, 2015.
[108] Z. Otwinowski, and W. Minor, "Processing of X-ray Diffraction Data Collected in Oscillation Mode," *Methods of Enzymology,* vol. 276, no. Macromolecular Crystallography Part A, pp. 307-326, 1997.

[109] P. Emsley, and K. Cowtan, "Coot: model-building tools for molecular graphics," *Acta Crystallogr D Biol Crystallogr,* vol. 60, no. Pt 12 Pt 1, pp. 2126-32, Dec, 2004.
[110] P. V. Afonine *et al.*, "Towards automated crystallographic structure refinement with phenix.refine," *Acta Crystallogr D Biol Crystallogr,* vol. 68, no. Pt 4, pp. 352-67, Apr, 2012.
[111] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods,* vol. 9, no. 7, pp. 671-675, 2012.
[112] A. Kilianski *et al.*, "Assessing activity and inhibition of Middle East respiratory syndrome coronavirus papain-like and 3C-like proteases using luciferase-based biosensors," *J Virol,* vol. 87, no. 21, pp. 11955-62, Nov, 2013.
[113] J. Lei, and R. Hilgenfeld, "Structural and mutational analysis of the interaction between the Middle-East respiratory syndrome coronavirus (MERS-CoV) papain-like protease and human ubiquitin," *Virol Sin*, May, 2016.
[114] X. Zhu, R. Ménard, and T. Sulea, "High incidence of ubiquitin-like domains in human ubiquitin-specific proteases," *Proteins,* vol. 69, no. 1, pp. 1-7, Oct, 2007.
[115] A. Giner-Sorolla, and A. Bendich, "Fluorine-containing Pyrimidines and Purines: Synthesis and Properties of Trifluoromethyl Pyrimidines and Purines," *J. Am. Chem. Soc.,* vol. 80, no. 21, pp. 5744-5752, 1958.
[116] M. Medebielle, S. Fujii, and K. Kato, "An Electrochemical Approach for the Synthesis of Perfluoroalkylated Purine and Indole Analogues of Plant Growth Regulators," *Tetrahedron,* vol. 56, no. 17, pp. 2655-2664, 2000.
[117] H. Lee *et al.*, "Identification of novel drug scaffolds for inhibition of SARS-CoV 3-Chymotrypsin-like protease using virtual and high-throughput screenings," *Bioorg Med Chem,* vol. 22, no. 1, pp. 167-77, Jan, 2014.
[118] J. Baell, and M. A. Walters, "Chemistry: Chemical con artists foil drug discovery," *Nature,* vol. 513, no. 7519, pp. 481-3, Sep, 2014.
[119] C. M. Daczkowski *et al.*, "Structural Insights into the Interaction of Coronavirus Papain-Like Proteases and Interferon-Stimulated Gene Product 15 from Different Species," *J Mol Biol,* vol. 429, no. 11, pp. 1661-1683, Jun, 2017.
[120] Y. M. Baez-Santos *et al.*, "X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases," *J Med Chem,* vol. 57, no. 6, pp. 2393-412, Mar 27, 2014.
[121] Y. M. Baez-Santos *et al.*, "Catalytic function and substrate specificity of the papain-like protease domain of nsp3 from the Middle East respiratory syndrome coronavirus," *J Virol,* vol. 88, no. 21, pp. 12511-27, Nov, 2014.

[122] H. A. Lindner *et al.*, "Selectivity in ISG15 and ubiquitin recognition by the SARS coronavirus papain-like protease," *Arch Biochem Biophys,* vol. 466, no. 1, pp. 8-14, Oct, 2007.
[123] J. Narasimhan *et al.*, "Crystal structure of the interferon-induced ubiquitin-like protein ISG15," *J Biol Chem,* vol. 280, no. 29, pp. 27356-65, Jul, 2005.
[124] L. A. Durfee *et al.*, "The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15," *Mol Cell,* vol. 38, no. 5, pp. 722-32, Jun, 2010.
[125] C. Zhao *et al.*, "Interferon-induced ISG15 pathway: an ongoing virus-host battle," *Trends Microbiol,* vol. 21, no. 4, pp. 181-6, Apr, 2013.
[126] N. V. Giannakopoulos *et al.*, "ISG15 Arg151 and the ISG15-conjugating enzyme UbE1L are important for innate immune control of Sindbis virus," *J Virol,* vol. 83, no. 4, pp. 1602-10, Feb, 2009.
[127] C. Lai *et al.*, "Mice lacking the ISG15 E1 enzyme UbE1L demonstrate increased susceptibility to both mouse-adapted and non-mouse-adapted influenza B virus infection," *J Virol,* vol. 83, no. 2, pp. 1147-51, Jan, 2009.
[128] S. D. Speer *et al.*, "ISG15 deficiency and increased viral resistance in humans but not mice," *Nat Commun,* vol. 7, pp. 11496, May, 2016.
[129] X. Zhang *et al.*, "Human intracellular ISG15 prevents interferon-α/β over-amplification and auto-inflammation," *Nature,* vol. 517, no. 7532, pp. 89-93, Jan, 2015.
[130] L. Chen *et al.*, "Hepatic gene expression discriminates responders and nonresponders in treatment of chronic hepatitis C viral infection," *Gastroenterology,* vol. 128, no. 5, pp. 1437-44, May, 2005.
[131] WHO. "Middle East respiratory syndrome (MERS-CoV)."
[132] A. R. Fehr, R. Channappanavar, and S. Perlman, "Middle East Respiratory Syndrome: Emergence of a Pathogenic Human Coronavirus," *Annu Rev Med,* vol. 68, pp. 387-399, Jan, 2017.
[133] C. M. Daczkowski *et al.*, "Structurally guided removal of deISGylase biochemical activity from papain-Like protease originating from the Middle East Respiratory Syndrome Virus," *J Virol*, Sep, 2017.
[134] K. D. Wilkinson, T. Gan-Erdene, and N. Kolli, "Derivitization of the C-terminus of ubiquitin and ubiquitin-like proteins using intein chemistry: methods and uses," *Methods Enzymol,* vol. 399, pp. 37-51, 2005.
[135] W. Minor *et al.*, "HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes," *Acta Crystallogr D Biol Crystallogr,* vol. 62, no. Pt 8, pp. 859-66, Aug, 2006.
[136] A. J. McCoy *et al.*, "Phaser crystallographic software," *J Appl Crystallogr,* vol. 40, no. Pt 4, pp. 658-674, 08, 2007.

[137] L. Zheng, U. Baumann, and J. L. Reymond, "An efficient one-step site-directed and site-saturation mutagenesis protocol," *Nucleic Acids Res,* vol. 32, no. 14, pp. e115, Aug, 2004.
[138] M. Akutsu *et al.*, "Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains," *Proc Natl Acad Sci U S A,* vol. 108, no. 6, pp. 2228-33, Feb, 2011.
[139] T. W. James *et al.*, "Structural basis for the removal of ubiquitin and interferon-stimulated gene 15 by a viral ovarian tumor domain-containing protease," *Proc Natl Acad Sci U S A,* vol. 108, no. 6, pp. 2222-7, Feb, 2011.
[140] R. Guan *et al.*, "Structural basis for the sequence-specific recognition of human ISG15 by the NS1 protein of influenza B virus," *Proc Natl Acad Sci U S A,* vol. 108, no. 33, pp. 13468-73, Aug, 2011.
[141] D. Komander, and M. Rape, "The ubiquitin code," *Annu Rev Biochem,* vol. 81, pp. 203-29, 2012.
[142] B. Mole, "Deadly pig virus slips through US borders," *Nature,* vol. 499, no. 7459, pp. 388, Jul, 2013.
[143] C. Lee, "Porcine epidemic diarrhea virus: An emerging and re-emerging epizootic swine virus," *Virol J,* vol. 12, pp. 193, Dec, 2015.
[144] S. Subramaniam *et al.*, "Vaccination of sows with a dendritic cell-targeted porcine epidemic diarrhea virus S1 protein-based candidate vaccine reduced viral shedding but exacerbated gross pathological lesions in suckling neonatal piglets," *J Gen Virol,* vol. 99, no. 2, pp. 230-239, Feb, 2018.
[145] D. Song, H. Moon, and B. Kang, "Porcine epidemic diarrhea: a review of current epidemiology and available vaccines," *Clin Exp Vaccine Res,* vol. 4, no. 2, pp. 166-76, Jul, 2015.
[146] S. Perlman, and A. A. Dandekar, "Immunopathogenesis of coronavirus infections: implications for SARS," *Nat Rev Immunol,* vol. 5, no. 12, pp. 917-27, Dec, 2005.
[147] Y. Kim *et al.*, "Broad-spectrum inhibitors against 3C-like proteases of feline coronaviruses and feline caliciviruses," *J Virol,* vol. 89, no. 9, pp. 4942-50, May, 2015.
[148] S. E. St John *et al.*, "X-ray structure and inhibition of the feline infectious peritonitis virus 3C-like protease: Structural implications for drug design," *Bioorg Med Chem Lett,* vol. 25, no. 22, pp. 5072-7, Nov, 2015.
[149] Y. Kim *et al.*, "Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor," *PLoS Pathog,* vol. 12, no. 3, pp. e1005531, Mar, 2016.
[150] C. Dye, and S. G. Siddell, "Genomic RNA sequence of Feline coronavirus strain FIPV WSU-79/1146," *J Gen Virol,* vol. 86, no. Pt 8, pp. 2249-53, Aug, 2005.

[151] T. C. Terwilliger *et al.*, "Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard," *Acta Crystallogr D Biol Crystallogr,* vol. 64, no. Pt 1, pp. 61-9, Jan, 2008.
[152] A. Catic *et al.*, "ElaD, a Deubiquitinating protease expressed by E. coli," *PLoS One,* vol. 2, no. 4, pp. e381, Apr, 2007.
[153] L. M. Iyer, A. M. Burroughs, and L. Aravind, "The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like beta-grasp domains," *Genome Biol,* vol. 7, no. 7, pp. R60, 2006.
[154] M. J. Pearce *et al.*, "Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis," *Science,* vol. 322, no. 5904, pp. 1104-7, Nov, 2008.
[155] R. Ekkebus *et al.*, "On terminal alkynes that can react with active-site cysteine nucleophiles in proteases," *J Am Chem Soc,* vol. 135, no. 8, pp. 2867-70, Feb, 2013.
[156] M. E. Davis, and M. U. Gack, "Ubiquitination in the antiviral immune response," *Virology,* vol. 479-480, pp. 52-65, May, 2015.
[157] A. G. Daniel, and N. P. Farrell, "The dynamics of zinc sites in proteins: electronic basis for coordination sphere expansion at structural sites," *Metallomics,* vol. 6, no. 12, pp. 2230-41, Dec, 2014.
[158] X. Robert, and P. Gouet, "Deciphering key features in protein structures with the new ENDscript server," *Nucleic Acids Res,* vol. 42, no. Web Server issue, pp. W320-4, Jul, 2014.
[159] A. Waterhouse *et al.*, "SWISS-MODEL: homology modelling of protein structures and complexes," *Nucleic Acids Res,* vol. 46, no. W1, pp. W296-W303, Jul, 2018.
[160] D. Flierman *et al.*, "Non-hydrolyzable Diubiquitin Probes Reveal Linkage-Specific Reactivity of Deubiquitylating Enzymes Mediated by S2 Pockets," *Cell Chem Biol,* vol. 23, no. 4, pp. 472-82, 04, 2016.

VITA

Jozlyn R. Clasman

Education

Purdue University — West Lafayette, IN
Ph.D. in Biological Sciences (GPA: 4.0) — 2014 – 2019
Concentration in Biomolecular Structure / Biophysics
Relevant course work: Methods and Measurements in Physical Biochemistry, Cryo-electron microscopy, 3D Reconstruction, Crystallography of Macromolecules

Grand Valley State University (GVSU) — Allendale, MI
B.S. in Chemistry, Magna Cum Laude — 2010 – 2014
Emphasis in Biochemistry / Biotechnology, Minor in Psychology

Publications

Clasman, J.R., Wesenberg, D.T. and Mesecar, A.D. (2018) Evaluating ubiquitin specificity of papain-like protease 2 from alphacoronaviruses FIPV and PEDV using structure-guided engineering. Manuscript in preparation.

Clasman, J.R., Everett, R.K., and Mesecar, A.D. (2018) X-ray structural and mutational analysis of the MERS virus papain-like-protease-ISG15 complex: Another step towards designed coronaviruses for attenuated vaccines. Manuscript in preparation.

Daczkowski, C.M., Dzimianski, J.V., **Clasman, J.R.**, Goodwin, O., Mesecar, A.D., and Pegan, S.D. (2017) 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.

Clasman, J.R., Báez-Santos, Y.M., Mettelman, R.C., O'Brien, A., Baker, S.C., and Mesecar, A.D. (2017) X-ray structure and enzymatic activity profile of a core papain-like protease of MERS coronavirus with utility for structure-based drug design. Nat. Sci. Reports. 7:40292.

Mitchell, J.M., **Clasman, J.R.**, June, C.M., Kaitany, K.J., LaFleur, J., Taracila, M.A., Klinger, N.V., Bonomo, R.A., Wymore, T., Szarecka, A., Powers, R.A., and Leonard, D.A. (2015) The structural basis of activity against aztreonam and extended spectrum cephalosporins for two carbapenem-hydrolyzing class D β-Lactamases from Acinetobacter Baumannii. Biochemistry. 54(10): 1976-87.

Leadership Experience & Outreach

Purdue University — West Lafayette, IN

Graduate Student Advisory Board Life Sciences student representative — 2017 – present

- Selected to liaison to the graduate school and promote social justice seminars and events
- Helped with the Office of Interdisciplinary Graduate Programs (OIGP) spring reception

Structural Biology social coordinator — 2016 – present

- Organized monthly social events for students and faculty in the Structural Biology group
- Collected departmental funds, purchased food/drinks, and helped decide on event themes

Member, Women in Science Program (WISP) — 2016 – present

- Attended meetings to learn/discuss leadership, networking, and life-management topics

Volunteer, The 4th installment of 'The Hitchhiker's Guide to the Biomolecular Galaxy — 2018

- Judged poster competition for the graduate student poster award

Teaching Assistant — Jan 2018 – May 2018

- Assisted Dr. Andrew Mesecar in the Biological & Structural Aspects of Drug Design and Action course
- Co-wrote homework and exam keys, graded exams and homework, and judged final presentations
- Helped and supervised students in weekly office hours and final projects

Volunteer, Lafayette Regional Science and Engineering Fair — 2016

- Judged posters in biochemistry for K-12 students

Research Experience

Purdue University — West Lafayette, IN

PhD researcher – Mesecar lab — May 2015 – present

- Used X-ray crystallography and steady-state enzyme kinetics to evaluate virus-host interactions
- Tested the efficacy of inhibitors against a panel of human and viral deubiquinating (DUB) enzymes
- Determined and analyzed 3 X-ray crystal structures (PDB codes 6BI8, 4KO3, 6NOZ)
- Purified 20+ viral protease site-directed mutants and tested activities by FRET and SDS-PAGE analysis
- Organized Mesecar lab substrate inventory for DUBs
- Wrote one published first author manuscript and drafted 2 first author manuscripts in preparation
- Gave an oral presentation at an international conference in front of 200+ attendees
- Mentored and trained 4 first year Ph.D. rotation students and 3 undergraduate students

Grand Valley State University (GVSU) — Allendale, MI

Undergraduate researcher – Leonard lab — Jan 2013 – July 2014

- Crystallized and solved several acyl-enzyme complexes of β-lactamases bound to β-lactam antibiotics (PDB codes 4X53, 4X55, and 4X56)
- Determined kinetic parameters for β-lactamases towards 3rd generation cephalosporins using UV/Vis spectroscopy and fluorescence anisotropy
- Authored a published manuscript (co-first author)
- Gave a poster presentation at an international conference

Awards & Honors

- Graduate Women in Science Program (WISP) spring 2017 travel grant (2017)
- Science, Technology, Engineering, and Mathematics (S-STEMS) Scholarship (2012 – 2014)
- GVSU Department of General Chemistry, Organic Chemistry, and Biochemistry Awards (2012-2014)

- Student Summer Scholars (S3) program grant (May 2013 – August 2013)
- GVSU Grand Finish Grant (2013)

Skills & Interests

Skills:

- Computer/Technical (Microsoft office, Adobe Photoshop & Illustrator, some experience with Linux)
- Scientific instruments (GE healthcare ÄKTA purifier & pure FPLC systems, BioTek & Clariostar plate readers, UNICORN software)
- Biophysical techniques (FRET, SEC, UV/Vis, CD spectroscopy, some experience with FP)
- Crystallographic and molecular visualization software (HKL2000/3000, CCP4, Phenix, COOT, Pymol, Chimera)
- HTS crystallization screening and automation (Mosquito, DragonFly, Rigaku Minstrel HT plate hotel, Rock Imager)
- Cryo-EM software (some experience with EMAN2, Relion)
- Scientific Data Analysis/Graphing Software (MARS, GraphPAD, SigmaPlot)

Interests: Horseback riding, running, board games, yoga

Presentations

Clasman, J.R., Wesenberg, D.T. and Mesecar, A.D. Evaluating ubiquitin specificity of papain-like protease 2 from alphacoronaviruses FIPV and PEDV using structure-guided engineering. Structural and Computational Biology and Biophysics Seminar Series, West Lafayette, IN, December 2018 (Invited oral presentation).

Clasman, J.R., Everett, R.K., and Mesecar, A.D. X-ray structure of MERS papain-like protease bound with ISG15 facilitates design of PLPs with attenuated substrate specificities. The Annual Biochemistry Research Retreat, West Lafayette, IN, August 2018 (Invited oral presentation).

Clasman, J.R., Everett, R.K., and Mesecar, A.D. X-ray structure of MERS papain-like Protease bound with ISG15 facilitates design of PLPs with attenuated or enhanced substrate specificities. The Protein Society 32nd Annual Symposium, Boston, MA, July 2018 (Poster presentation).

Clasman, J.R., Everett, R.K., and Mesecar, A.D. Structure of MERS viral protease bound to the complete human ISG15 facilitates functional tools for coronavirus attenuation. The 4th installment of 'The Hitchhiker's Guide to the Biomolecular Galaxy', West Lafayette, IN, May 2018 (Oral presentation).

Clasman, J.R., Everett, R.K., and Mesecar, A.D. Structure of MERS viral protease bound to the complete human ISG15 facilitates functional tools for vaccine design. OIGP Spring Reception, West Lafayette, IN, May 2018 (Poster presentation).

Clasman, J.R., Báez-Santos, Y.M., Mettelman, R.C., O'Brien, A., Baker, S.C., and Mesecar, A.D. Structural and functional studies of the core papain-like protease from MERS coronavirus. PULSe student seminar series, West Lafayette, IN, August 2017 (Invited oral presentation).

Clasman, J.R., Báez-Santos, Y.M., Mettelman, R.C., O'Brien, A., Baker, S.C., and Mesecar, A.D. Structure and kinetics of a core MERS papain-like protease with utility for structure-based drug design. XIVth International Nidovirus Symposium, Kansas City, MO, June 2017 (Invited oral presentation).

Clasman, J.R., Báez-Santos, Y.M., Mettelman, R.C., O'Brien, A., Baker, S.C., and Mesecar, A.D. Crystal structure and kinetic profile of a core papain-like protease of MERS coronavirus with utility for structure-based drug design. The 3rd installment of 'The Hitchhiker's Guide to the Biomolecular Galaxy', West Lafayette, IN, May 2017 (Poster presentation). Received best graduate student poster presentation award.

Clasman, J.R., Báez-Santos, Y.M., Mettelman, R.C., O'Brien, A., Baker, S.C., and Mesecar, A.D. Crystal structure and kinetic profile of a core papain-like protease of MERS coronavirus with utility for structure-based drug design. 36th Midwest Enzyme Chemistry Conference, Chicago, IL, October 2016 (Poster presentation).

Clasman, J.R., Báez-Santos, Y.M., and Mesecar, A.D. The ubiquitin-like 2 domain of MERS coronavirus nsp3 is not required for papain-like protease catalysis. OIGP Spring Reception, West Lafayette, IN, September 2016 (Poster presentation).

Clasman, J.R., Báez-Santos, Y.M., and Mesecar, A.D. The ubiquitin-like 2 domain of MERS coronavirus nsp3 is not required for papain-like protease catalysis. The 2nd installment of 'The Hitchhiker's Guide to the Biomolecular Galaxy', West Lafayette, IN, May 2016 (Poster presentation).

Clasman, J.R., Báez-Santos, Y.M., and Mesecar, A.D. The ubiquitin-like 2 domain of MERS coronavirus nsp3 is not required for papain-like protease catalysis. Biomedical Engineering and Interdisciplinary Life Sciences Graduate Research Symposium, West Lafayette, IN, February 2016 (Poster presentation).

Clasman, J.R., Báez-Santos, Y.M., and Mesecar, A.D. The ubiquitin-like domain associated with the papain-like protease from MERS virus is not required for its catalytic function. OIGP Spring Reception, West Lafayette, IN, April 2015 (Poster presentation).

Clasman, J.R., June, C.M., Jackman, B.J., Powers, R.A., and Leonard, D.A. Biochemical basis for the extended spectrum cephalosporinase activity of a clinical AmpC β-lactamase variant. 28th Annual Symposium of the Protein Society, San Diego, CA, July 2014 (Poster presentation).

Clasman, J.R., Mitchell, J.M., June, C.M., Kaitany, K.J., Klinger, N.V., Bonomo, R.A., Powers, R.A., and Leonard, D.A. Understanding the structural basis of activity against aztreonam and expanded-spectrum cephalosporins for two clinically-derived carbapenem-hydrolyzing class D β-Lactamases in Acinetobacter spp. West Michigan Regional Undergraduate Science Research Conference at Van Andel Institute, Grand Rapids, MI, November 2013 (Poster presentation).

Clasman, J.R., Mitchell, J.M., Kaitany, K.J., Klinger, N.V., Baggett, V.L., June, C.M., Powers, R.A., and Leonard, D.A. Crystal structures of a clinically-derived class D β-lactamase variant with cefotaxime, ceftazidime, and aztreonam bound. 33rd Midwest Enzyme Chemistry Conference, Chicago, IL, October 2013 (Poster presentation).

PUBLICATIONS

Clasman, J.R., Wesenberg, D.T. and Mesecar, A.D. (2018) Evaluating ubiquitin specificity of papain-like protease 2 from alphacoronaviruses FIPV and PEDV using structure-guided engineering. Manuscript in preparation.

Clasman, J.R., Everett, R.K., and Mesecar, A.D. (2018) X-ray structural and mutational analysis of the MERS virus papain-like-protease-ISG15 complex: Another step towards designed coronaviruses for attenuated vaccines. Manuscript in preparation.

Daczkowski, C.M., Dzimianski, J.V., **Clasman, J.R.**, Goodwin, O., Mesecar, A.D., and Pegan, S.D. (2017) 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.

Clasman, J.R., Báez-Santos, Y.M., Mettelman, R.C., O'Brien, A., Baker, S.C., and Mesecar, A.D. (2017) X-ray structure and enzymatic activity profile of a core papain-like protease of MERS coronavirus with utility for structure-based drug design. Nat. Sci. Reports. 7:40292.

Mitchell, J.M., **Clasman, J.R.**, June, C.M., Kaitany, K.J., LaFleur, J., Taracila, M.A., Klinger, N.V., Bonomo, R.A., Wymore, T., Szarecka, A., Powers, R.A., and Leonard, D.A. (2015) The structural basis of activity against aztreonam and extended spectrum cephalosporins for two carbapenem-hydrolyzing class D β-Lactamases from Acinetobacter Baumannii. Biochemistry. 54(10): 1976-87.