2015-12-15

Hepatitis C Virus: Structural Insights into Protease Inhibitor Efficacy and Drug Resistance: A Dissertation

Djade I. Soumana
University of Massachusetts Medical School

HEPATITIS C VIRUS: STRUCTURAL INSIGHTS INTO PROTEASE INHIBITOR EFFICACY AND DRUG RESISTANCE

A Dissertation Presented

By

DJADÉ IBRAHIM SOUMANA

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences,
Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

BIOCHEMISTRY & MOLECULAR PHARMACOLOGY

DECEMBER 15, 2015

HEPATITIS C VIRUS: STRUCTURAL INSIGHTS INTO PROTEASE INHIBITOR EFFICACY AND DRUG RESISTANCE

A Dissertation Presented

By

DJADÉ IBRAHIM SOUMANA

The signatures of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

---

CELIA A. SCHIFFER, PH.D., Thesis Advisor

KENDALL KNIGHT, PH.D., Member of Committee

DAVID G. LAMBRIGHT, PH.D., Member of Committee

WILLIAM E. ROYER, PH.D., Member of Committee

DARIA HAZUDA, PH.D., Member of Committee

JEANNE HARDY, PH.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

---

FRANCESCA MASSI, PH.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school

---

ANTHONY CARRUTHERS, PH.D.
Dean of the Graduate School of Biomedical Sciences

BIOCHEMISTRY & MOLECULAR PHARMACOLOGY

DECEMBER 15, 2015

Table of Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

5172-Linear ..... MK-5172's linear analog

5172mc$P_1 - P_3$ ..... MK-5172's $P_1 - P_3$ macrocyclic analog

ASV ..... Asunaprevir

ASVmc ..... Asunaprevir's $P_1 - P_3$ macrocyclic analog

ADME ..... Absorption, distribution, metabolism, and excretion

BOC ..... Boceprevir

DAA ..... Direct acting antiviral

DAN ..... Danoprevir

DENV ..... Dengue virus

FC ..... Fold change

FDA ..... Food and Drug Administration

GAG ..... Glycosaminoglycans

GRZ ..... Grazoprevir

GT ..... Genotype

H-bonds ..... Hydrogen bonds

HCV ..... Hepatitis C virus

HIV ..... Human Immunodeficiency Virus

ITC ..... Isothermal titration calorimetry

LDLs ..... Low-density lipoproteins

MAVS ..... Mitochondrial antiviral signaling

MD ..... Molecular dynamics

MDR ..... Multi-drug resistant

MES ..... 2-(N-Morpholino)ethanesulfonic acid

NS ..... Non structural protein

ORF ..... Open reading frame

Peg-INF ..... Pegylated interferonα

PI ..... Protease inhibitors

RAV ..... Resistance associated variant

RBV ..... Ribavirin

RMSD ..... Root mean square deviations

RMSF ..... Root mean square fluctuations

RVR ..... Rapid virological response

SAR ..... Structureactivity relationship

SBDD ..... Structure based drug design

SEH ..... Substrate envelope hypothesis

SIM ..... Simeprevir

SOC ..... Standard of care

SVR ..... Sustained virological response

SVR12 ..... Sustained virological response achieved after 12 weeks of treatment

SVR24 ..... Sustained virological response achieved after 24 weeks of treatment

TCEP ..... Tris(2-carboxyethyl)phosphine

TEL ..... Telaprevir

UTR ..... Untranslated region

VAN ..... Vaniprevir

WNV ..... West nile virus

WT ..... Wild type

THIRD PARTY LICENSES

The Following were adapted from journals where permission is required

Chapter/Figure	Publisher	License number
Chapter 1, Figure 1	Nature Publishing Group	3757820940035
Appendix A, Figure 1 1	ChemMedChem	376392136662

The Following were adapted from journals where NO permission required

Figure/Chapter	Publisher	Article	Authors
Chpater 1	PLOS Pathogen	The molecular basis of drug resistance against hepatitis C virus NS3/4A protease inhibitors	Romano et al
Chapter 1	ACS Chemical Biology	Evaluating the role of macrocycles in the susceptibility of hepatitis C virus NS3/4A protease inhibitors to drug resistance	Ali et al
Figure 2	ACS Med Chem Lett	The HCV Revolution Did Not Happen Overnight	Kwong et al
Chapter 2	ACS Chemical Biology	Structural Analysis of Asunaprevir Resistance in HCV NS3/4A Protease	Soumana et al
Chapter 3	ACS Chemical Biology	Structural and Thermodynamic Effects of Macrocyclization in HCV Inhibitor Grazoprevir (MK-5172)	Soumana et al

Table 0.1: Third Party licenses

Cœur qui soupire n'a pas

ce qu'il désire

— *Blaise de Monluc*

In loving memory of

my mother and guardian angel

Assamaou Larabou Soumana

[April 1947– December 1997]

ACKNOWLEDGEMENTS

”And, when you want something, all the universe conspires in helping you to achieve.”

Paulo Coelho,

The Alchemist

A graduate program in biomedical science is similar to running a marathon. It requires a lot of endurance and a tremendous amount of support, where people continuously cheer you on, where they pick you up when you trip, and guide you if you ever get sidetracked. But most importantly, they constantly remind you during the hardest times why you started the marathon. Completing my marathon would not have been possible without first the support and encouragement from my friends and family. I would like to acknowledge the endless support I have received from my father, Soumana Ousseini, and my sisters Rekia, Miriam, Aichatou, and Fatih. I must also acknowledge my closest friends, my defacto brothers, Rachid and Soufiane, who have continuously cheered me on step by step. My adoptive family– Madhavi Kolli, Sagar Kathuria, Kuan-Hung Lin and Tania Silvas– deserve my deepest gratitude for keeping a close eye on me and making sure that my smile never faded. My finacé, Kristina Prachanronarong, was and continues to be a prominent role mode and an incredible support system of mine.

I will forever be indebted to my thesis advisory committee for their active participation in my scientific development. Thank you very much to Dr. Francesca Massi for agreeing

to be both my TRAC and DEC chairwoman. You were particularly instrumental when I was initially beginning to do ITC. When I was learning ITC, you took time out of your busy schedule to sit down with me and walk me through my experiments. A big thank you also goes to both Drs. Bill Royer and David Lambright, both of whom have never hesitated to get their hands wet and help me at the bench. I will never forget going to Biotech II and having Dr. Lambright help me shoot my very important crystals, nor will I forget the times when Dr. Royer would sit down with me at the computer and help me work through diffraction patterns and spacegroups. I must absolutely show my gratitude to Dr. Hazuda for taking time out of her very busy schedule to be my external examiner.

A very special thank you also goes to Dr. Nese Kurt-Yilmaz for constantly making herself available to discuss data interpretation and for helping me refine my scientific writing. A very special thank you goes to Ellen Nalivaika, Dr. Hong Cao, Dr. Mohan Somasundaran, the Schiffer labmembers past & present, and especially Dr. Akbar Ali. You have all significantly contributed to my scientific growth and professional success. Last but not least, I must recognize my mentor, Dr. Celia Schiffer. Five years ago when I walked into your office and told you that I had no experience in crystallography or biophysics and that I absolutely had to join your lab, you did not hesitate. You gave me an opportunity where others in your shoes would have turned me away for my lack of experience in the field. You have cheered me on for the last five years and encouraged me to achieve my personal career goals. Thanks to you, I would consider my graduate experience a true success, something certainly beyond my wildest dreams.

ABSTRACT

The Hepatitis C Virus (HCV) is a global health problem as it afflicts an estimated 170 million people worldwide and is the major cause of viral hepatitis, cirrhosis and liver cancer. HCV is a rapidly evolving virus, with 6 major genotypes and multiple subtypes. Over the past 20 years, HCV therapeutic efforts have focused on identifying the best-in-class direct acting antiviral (DAA) targeting crucial components of the viral lifecycle, The NS3/4A protease is responsible for processing the viral polyprotein, a crucial step in viral maturation, and for cleaving host factors involved in activating immunity. Thus targeting the NS3/4A constitutes a dual strategy of restoring the immune response and halting viral maturation. This high priority target has 4 FDA approved inhibitors as well as several others in clinical development. Unfortunately, the heterogeneity of the virus causes seriously therapeutic challenges, particularly the NS3/4A protease inhibitors (PIs), which suffer from both the rapid emergence of drug resistant mutants as well as a lack of pan-genotypic activity.

My thesis research focused on filling two critical gaps in our structural understanding of inhibitor binding modes. The first gap in knowledge is the molecular basis by which macrocyclization of PIs improves antiviral activity. Macrocycles are hydrophobic chains used to link neighboring chemical moieties within an inhibitor and create a structurally pre-organized ligand. In HCV PIs, macrocycle come in two forms: a $P_1 - P_3$ and $P_2 - P_4$ strategy. I investigated the structural and thermodynamic basis of the role of macrocyclization in reducing resistance susceptibility. For a rigorous comparison, we designed and synthesized both a $P_1 - P_3$ and a linear analog of grazoprevir, a $P_2 - P_4$ inhibitor. I found that, while the $P_2 - P_4$ strategy is more favorable for achieving potency, it does not allow the inhibitor sufficient flexibility to accommodate resistance mutations. On the other hand, the $P_1 - P_3$ strategy strikes a better balance between potency and resistance barrier.

The second gap my thesis addresses is elucidating the structural basis by which highly

potent protease inhibitors function in genotype 1 but not in genotype 3, despite having an 87% sequence similarity. After mapping the amino acids responsible for this differential efficacy in genotypes 1 and 3, I engineered a 1a3a chimeric protease for crystallographic studies. My structural characterization of three PIs in complex with both the 1a3a and genotype 1 protease revealed that the loss of inhibitor efficacy in the 1a3a and GT-3 proteases is a consequence of disrupted electrostatic interactions between amino acids 168 and 155, which is critical for potent binding of quinoline and isoindoline based PIs. Here, I have revealed details of molecular and structural basis for the lack of PI efficacy against GT-3, which are needed for design of pan-genotypic inhibitors.

chapter I

Introduction

Chapter I

Introduction

1.1 Hepatitis C Virus

Hepatitis C Virus (HCV) infects an estimated 150 million people worldwide and is the leading cause of liver cancer and cirrhosis [1]. HCV is highly diverse with six genotypes (classified 1-6) further subcategorized into different subtypes (a, b, c, etc.) [2, 3]. Different genotypes have distinct patterns of global geographical distribution. Genotype 1 is the major genotype in the world, especially in continental America and Europe. Genotype 3, the second predominant genotype, is mostly observed in Southeastern Asia and countries including Russia, England and Australia. Other genotypes are relatively minor, but distinct geographical patterns of spread has been observed: genotype 2 is more dominant in sub-saharan Africa, genotype 4 in Middle East, genotype 5 in South Africa and genotype 6 in Asia. These genotypes and subtypes are distinct both genetically and epidemiologically; relative nucleotide differences are around 30% at genotype level and around 25% at subtype level [2, 4].

1.1.1 Viral lifecycle and viral proteins

The HCV life cycle consists of entry, uncoating, translation, replication, repackaging and exit phases. HCV circulates in the host organism in various ways, either in association with low-density lipoproteins (LDLs), immunoglobulins or as a free virus particle. Association of the envelope protein E2 with Tetraspanin (CD81) and Scavenger Receptor Class B Type I (SR-BI), in combination with LDL receptors (LDL- R), glycosaminoglycans (GAG), Claudin-I and Occludin, facilitate fusion with human hepatocytes. Upon undergo-

ing clatherin mediated enocytosis [5], the release of viral genome is facilitated by the acidic environment of the endoplasm [6].

The 9.6 kb genome of HCV encodes an open reading frame (ORF) flanked by 5' and 3' highly structured untranslated regions (UTRs). The 5' UTR contains the internal ribosome entry site (IRES), which is essential for genome translation. The 3' UTR plays a key role in viral propagation. HCV positive sense RNA is translated by the host replication machinery in a cap-independent manner, producing a single polyprotein composed of structural proteins (Core, E1, E2), p7 and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A, NS5B) [7]. The structural proteins and p7 are inserted into the endoplasmic reticulum lumen and cleaved into functional units by host signal signalases and host signal peptide peptidases [8]. Subsequent polyprotein processing is perform in cis and trans by the NS2/3 and NS3/4A [9, 10](Fig. 1.1)

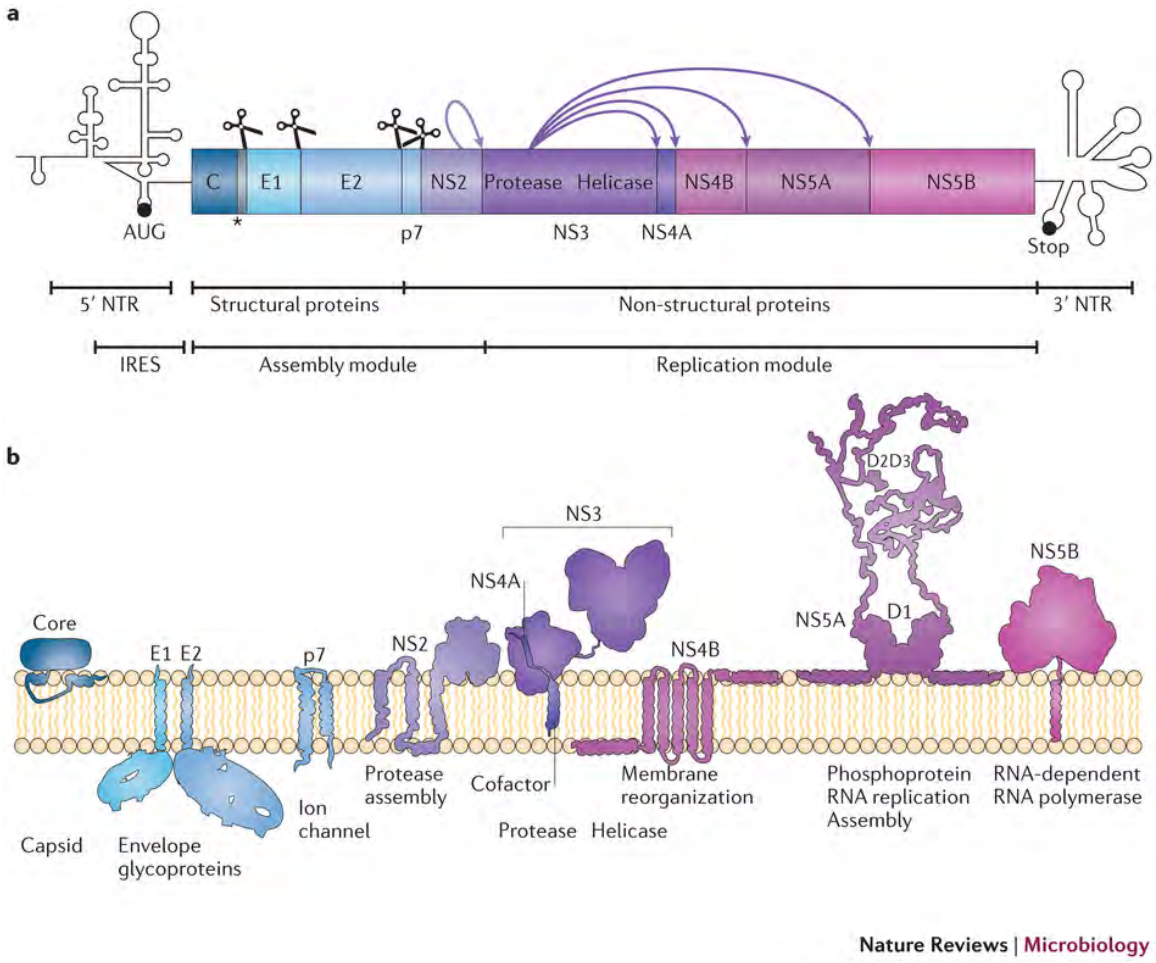


Figure 1.1: Schematic representation of HCV genome
the viral genome is shown with its various structural (blue) and non-structural proteins (purple) before **(A)** and after polyprotein processing **(B)**

Most of these proteins perform multiple functions required for viral propagation, host immune suppression and pathogenesis. Core and envelope proteins are an integral part of the virus particle, forming the viral capsid, facilitating virus–host cell interaction and infection. Core protein forms the viral capsid that binds the viral RNA [11] and also alters the host cell lipid mechanism. Highly glycosylated envelope proteins E1 and E2 form a heterodimeric complex that facilitates host–virus interactions [12]. The p7, a member of viroporins, has been shown to form oligomeric ion channels in vitro and to be critical for infectivity in vivo [13] [14] [15]. Non-structural proteins form the viral replication machinery and are also responsible for packaging of infectious virus particles. NS2 is implicated in virus assembly and polyprotein processing [16] [17]. In combination with N-terminal domain of NS3 protease, NS2/3 cysteine protease cleaves the NS2-NS3 junction [18]. NS3/4A is a multifunctional protease–helicase complex that processes the remaining viral polyprotein, aids in replication by unwinding RNA in ATP-dependent fashion and suppresses immune response [16] [19] [20]. NS4A cofactor, an integral part of the NS3/4A protease, anchors NS3/4A to the membrane and aids in the processivity and translocation efficiency of NS3 helicase [21] [22]. NS4B is implicated in the formation of membranous webs, where viral replication happens [23] [24]. NS5A is an essential phosphoprotein that binds to RNA and implicated in replication, virus assembly and regulation of host processes [25]. NS5B is the RNA-dependent RNA polymerase that replicates the viral genome [10].

1.2 Evolution of Standard of Care

Before the advent of direct-acting antivirals (DAAs), the standard of care (SOC) for HCV infection consisted of pegylated interferonα (Peg-IFN) and/or ribavirin (RBV). Peg-IFN often causes severe side effects and need to be administered intravenously, causing patient burden and limiting adherence. As only 40% of patients receiving a 48 week long Peg-IFN/RBV treatment achieve undetectable viremia post treatment completion, the urgent need for robust direct acting antivirals was clear.

The first DAA developed against HCV targeted the NS3/4A protease, which is a prime therapeutic target that is responsible for cleaving the viral polyprotein at diverse sequences 3-4A, 4A4B, 4B5A, and 5A5B and host cell proteins of the innate immune response, including TRIF and MAVS [26–28]. Thus, inhibiting NS3/4A serves a dual purpose by preventing viral maturation and restoring the immune response. The proof-of-concept for the first direct-acting antiviral (DAA) inhibiting HCV NS3/4A has been provided by the discovery of BILN-2061 [29] and its ability to reduce plasma HCV RNA levels by 2 to 4 orders of magnitude within 48 hours of treatment [30]. Despite its promise, BILN-2061's clinical development was stopped due to toxicity.

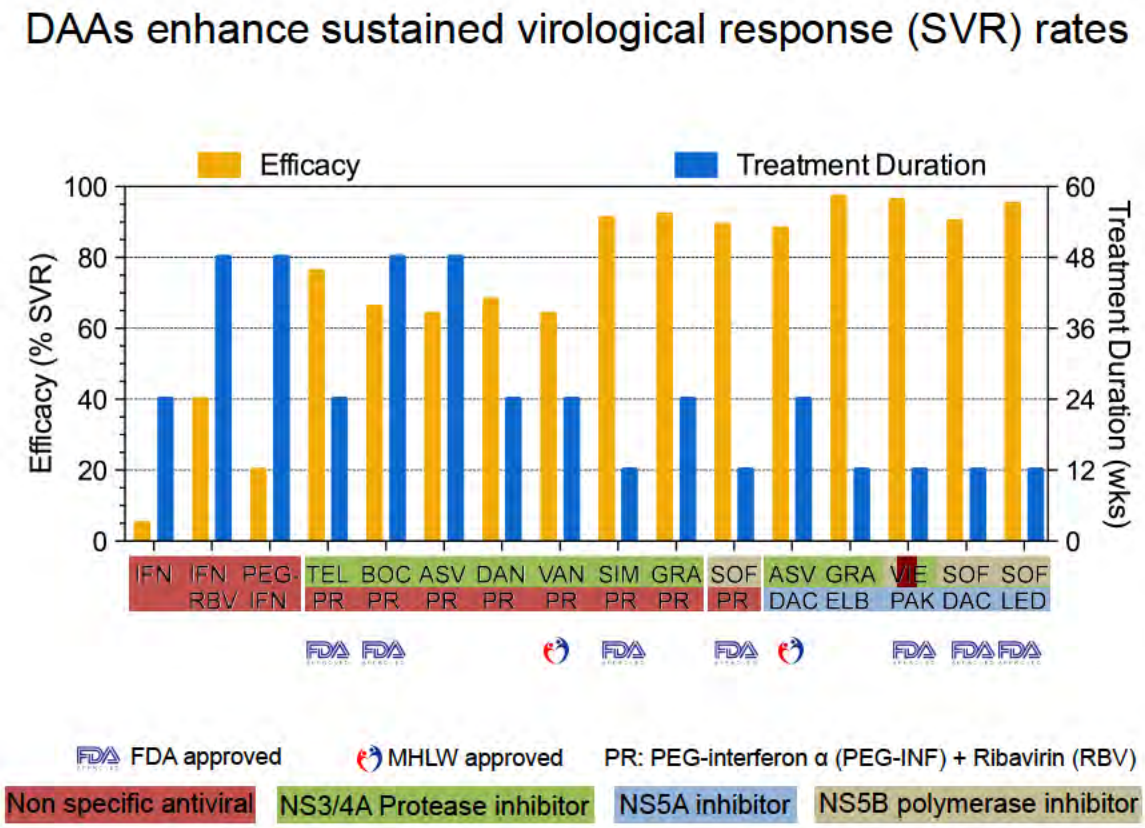
Telaprevir (TEL) is a first approved DAA [31, 32]. This NS3/4A protease inhibitor (PI) is chemically different from BILN-2061, but when administered with Peg-IFN/RBV, it demonstrated an ability to achieve 67% sutained virological response rates (SVR) [33, 34] with a 36 week treatment duration. Unfortunately, a high percentage of TEL patients experienced adverse events (21%, vs. 11% for the SOC).

Boceprevir (BOC) is the second HCV DAA to be approved [35]. BOC and TEL share a common chemical makeup and mode of action. in a SPRINT-1 clinical trial, 70% of patients receiving BOC/Peg-IFN/RBV for 24 weeks achieved SVR [36]. The clinical success of both TEL and BOC was challenged by a rapid emergence of overlapping drug resis-

tance [37].

The next wave of PIs to reach clinical development were asunaprevir (ASV) [38], simeprevir (SIM) [39], paritaprevir (PAR) [40]), danoprevir (DAN) and vaniprevir (VAN) [41]. All five PIs are approved for clinical use (in the US, Japan or UK) in combination therapies with other DAAs (NS5B RNA-dependent RNA polymerase and NS5A inhibitors). Thanks to DAAs, the current HCV SOC is Peg-IFN/RBV free, short in duration (12 weeks) and efficacious (SVR up to 80–90% depending on genotype) (Fig. 1.2).

Figure 1.2: HCV therapeutic landscape



1.3 Protease Inhibitors and Drug Resistance

Protease Inhibitors represent the most studied class of DAAs for HCV treatment. Upon observing that the N-terminal product of substrate sequences acted as competitive inhibitors, $P_6 - P_1$ substrate based sequences became the backbone of inhibitor design [42–44]. Consequently, HCV PIs are peptidomimetics that share a common core spanning P_3 through P_1, but contain different chemical groups or moieties at the P_4, P_2, and $P_{1'}$ positions, and often contain inter-moiety connections called macrocycles. PIs can thus be categorized in three broad groups based on their chemical composition and mode of action: linear covalent inhibitors, linear non-covalent inhibitors, and macrocyclics (Fig. 1.3).

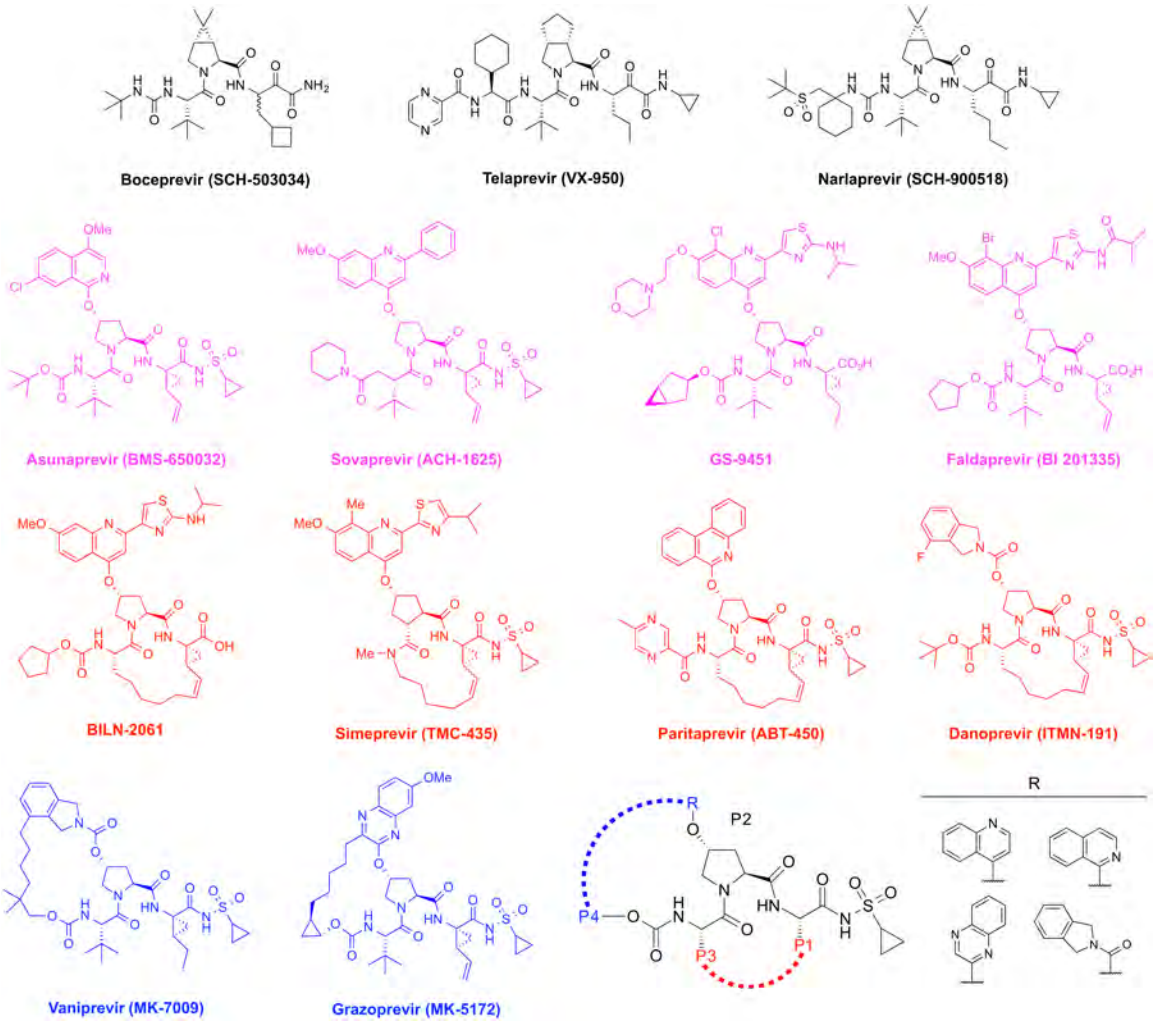


Figure 1.3: Schematic representation of NS3/4A protease inhibitors
The three broad classes of HCV PIs are shown: linear α-ketoamide covalent (top-black); linear non covalent (second row-pink); and macrocyclic ($P_1 - P_3$ and $P_2 - P_4$ in red and blue respectively.

The linear covalent inhibitors (TEL, BOC and narlaprevir) take advantage of the NS3/4A's catalytic mechanism. Proteolytic activity is carried out by the His57-Asp81-Ser139 catalytic triad via general acid-base catalysis chemistry, whereby the carboxyl oxygens of Asp81 coordinate the imidazole ring of His57 and positions the former near nucleophilic Ser139. Through a six-step reaction, the catalytic Ser139 forms a reversible covalent bond with the enzyme's substrate. By presenting a carbonyl carbon to the catalytic Ser139, PI can form reversible protease–inhibitor covalent linkage, which mimics the enzyme–substrate intermediate state. One major concern with covalent inhibition is off target effects. Proteins of similar evolutionary lineage often share common tertiary structures and can be inhibited by the same compound. TEL was reported to also target various host serine protease [45], representing an added complexity for therapeutic use.

1.3.1 Linear inhibitors

Linear non-covalent inhibitors (asunaprevir, sovaprevir, GS-9451, and faldaprevir) lack the R2-(C=O) warhead of their covalent counterpart, but instead, contain an extended P_2 moiety based on quinolone derivatives. Early structure activity relationship studies of the P_2 motif on a hexapeptide demonstrated improved antiviral activity when bulkier substituents were used [46, 47]. During the discovery of BILN-2061, the IC_{50} of Ac-DDIVP-Nva-OH hexapeptide was improved 21 fold with the addition of a benzyloxy motif at the γ carbon and by 400 fold with the larger naphthen-1-ylmeth-oxy derivative addition (150 µM, 7 µM, and 0.39 µM respectively) [46]. This naphthalene P_2 moiety set the stage for P_2 moieties such as quinoline/isoquinoline (asunaprevir), isoindoline (danoprevir-vaniprevir), and quinoxaline (grazoprevir) [48]. These inhibitors are commonly referred to as 1^{st} generation, 2^{nd} wave inhibitors because of their low-moderate resistance barrier.

1.3.2 Macrocyclic inhibitors

In HCV NS3/4A protease, macrocyclic drug candidates come in two forms: linking $P_1 - P_3$ or $P_2 - P_4$ (Fig. 1.3). Once again, the discovery of BILN-2061 was fundamental in the establishing the first example of an HCV macrocyclic inhibitor. Through molecular modeling of BILN-2061 on the full-length apo NS3/4A, several key observations were made that led to the development of the $P_2 - P_4$ macrocycle: (1) the P_2 thiazolylquinoline moiety packed against a featureless subpocket of the enzyme near residues Asp81, Arg155 and Asp168. As this moiety significantly increased ligand potency, a nonspecific interaction could be designed in this pocket; (2) the presence of a unique chemical space between the P_4 cyclopentane and the P_2 quinoline ring that could accommodate a heterocyclic connection. Coincidentally, examination of a full-length apo structure showed the side chain of Glu628 of the N-terminal substrate product overlapped with the proposed chemical space [49]. Thus the first reported $P_2 - P_4$ macrocyclic inhibitor possessed a P_2 isoquinoline connected to the P_4 carbamate by a six-carbon linker. Subsequent lead optimizations gave rise to the discovery of vaniprevir (MK-7009), a subnanomolar inhibitor with an outstanding pharmacokinetic profile [50]. Up until this point, PIs were performing well against GT-1 but not so well in non-1 viruses.

Grazoprevir (MK-5172) is a second generation PI, specifically developed to achieve a broad pan-genotypic activity, and robustness resistance profile [51]. Compared to its predecessors– VAN, DAN, and SIM– Grazoprevir (GRZ) displays 1 to 3 orders of magnitude improvement in both potency and resistance barrier. However, GRZ is not immune to resistance nor has it shown efficacy against GT-3 virus [51].

1.3.3 Drug resistance

Lacking a proof reading activity, the NS5B RNA dependent RNA polymerase misincorporation rate is estimated to be 10^{-5} to 10^{-4} per copied nucleotide [52]. For a rapidly evolving virus like HCV, this means that of the 10^{12} new virions produced daily, 91% will carry no mutation, whereas 8.7% (8.7 billion virions) and 0.42% (4.2 billion virions) will carry a single or double mutation, respectively [53]. Consequently, infected individuals will harbor a heterogeneous pool of viruses with diverse genomes, known as quasispecies [54], where the dominant or wildtype (WT) strain has a high replicative fitness [55, 56]. The diverse pool of quasispecies exists in treatment naïve patients. Once DAA administration begins, the dynamics of the viral pool change. The selective pressure of the antiviral quickly eliminates the WT virus in a process known as rapid virological response (RVR) but fails to affect some mutant variants [57]. Many reports describe the gradual emergence of drug resistant variants over the course of clinical and in vivo studies; and while several viral and host factors contribute to drug resistance, all DAAs are susceptible to drug resistance [58], including HCV NS3/4A PIs.

HCV protease inhibitors are highly potent against the wildtype GT-1 enzyme with low nanomolar to subnanomolar affinity, typically reducing the viral load by 3 to 5 orders of magnitude within 48 hours of PI administration. Despite this remarkable reduction in viremia, viral breakthrough is observed in all PI based monotherapy and Peg-INF/RBV combination therapies. Mechanistically, NS3/4A drug resistance arises when a single or series of amino acid changes occur within the protease leading to an unfavorabe inhibitor binding, a decrease in inhibitor potency, and a restoration of viral substrate processing (Fig. 1.4).

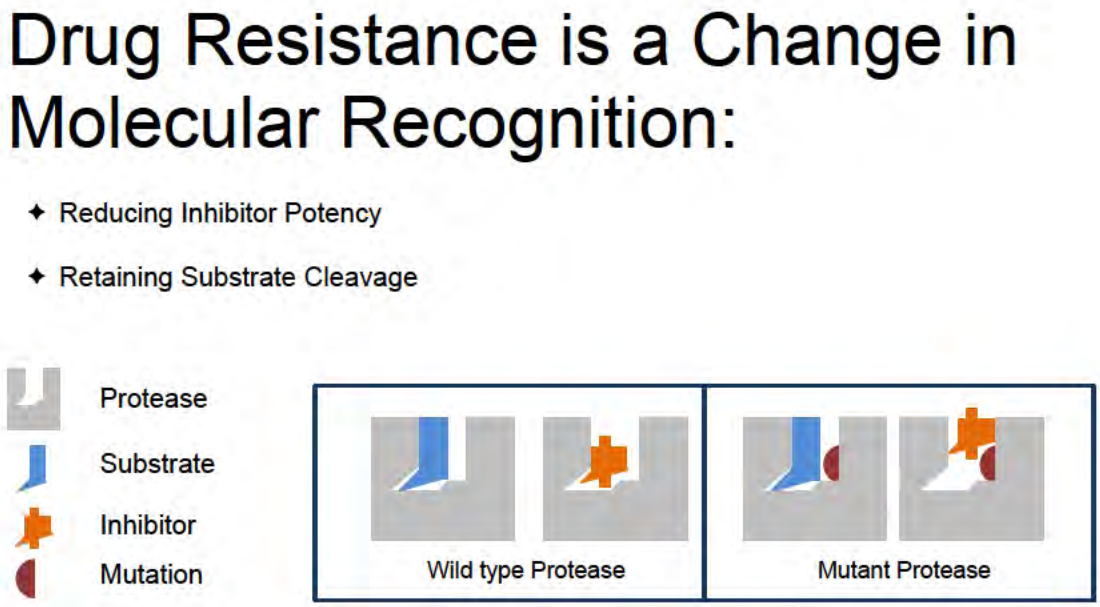


Figure 1.4: Molecular mechanism for drug resistance

The various reported NS3/4A drug resistance mutations cluster in the protease domain of the bifunctional enzyme in a pattern distinct to the PI's chemical structure and the enzyme's genotype [59]. For instance, R155K mutation confers resistance to virtually all PIs in GT-1a but is rarely observed in the GT-1b subtype. Instead, distinct resistance mutations arise in genotype 1b patients depending on the class of protease inhibitor used: A156 mutates in response to treatment with linear ketoamide protease inhibitors [60–62], whereas macrocyclic inhibitors more commonly select for D168A and R155K variants [63–66] (Fig. 1.5). Mutations at V36, T54, and V36+A155 are also associated with resistance to ketoamide inhibitors [60–62] with R155, A156 and D168 mutations conferring multidrug resistance. Variations in the patterns of resistance mutations arise from the complex interplay between genotype, replication rates, mutation rates, and the resulting effect of mutations on viral fitness and drug potency. Clearly, despite the benefits of combination therapy in improving SVR rates, the emergence of resistance challenges the long-term efficacy of NS3/4A protease inhibitors.

Major NS3/4A protease resistance associated variants (RAVs) in HCV genotype 1

RAVs	Mutation	Drug(s) susceptible
V36	***A,M,L,G***	***Boceprevir, telaprevir***
Q41	***R***	***Boceprevir, danoprevir***
F43	***S,C,V,I***	***Boceprevir, telaprevir, danoprevir, simeprevir***
V55	***A***	***Boceprevir***
T54	A,S	***Boceprevir, telaprevir, grazoprevir***
Q80	K,R,H,G,L	***Simeprevir, asunaprevir, faldaprevir, danoprevir, boceprevir, telaprevir***
S138	***T***	***Danoprevir, simeprevir***
R155	***K,T,I,M,G,L,S,Q***	***Faldaprevir, danoprevir, paritaprevir, vaniprevir, sovaprevir, GS-9451,boceprevir, telaprevir***
A156	***V,T,S,I,G***	***Faldaprevir, vaniprevir, simeprevir, grazoprevir, boceprevir, telaprevir***
V158	***I***	***Boceprevir***
D168	***A,V,E,G,N,T,Y,H,I***	***Danoprevir, vaniprevir, Asunaprevir, Paritaprevir, Sovaprevir, GS-9451, grazoprevir***
V170	***A***	***Boceprevir, telaprevir***
M175	***L***	***Boceprevir***

Active site mutations

Figure 1.5: NS3/4A resistance associated variants
HCV NS3/4A is susceptible to both non active site and active site mutations [59]. Active site residues 155, 156 and 168 are prone to multi-drug resistance

1.3.4 Substrate recognition versus inhibitor binding

Drug resistance at the molecular level is a change in the balance of recognition events between the relative affinity of the enzyme to bind inhibitors versus ability to process substrates. To better understand drug resistance, it is thus important to elucidate the structural basis for substrate recognition and inhibitor binding to the target protein.

An alignment of the host and viral substrate sequences recognized by the NS3/4A revealed conservation at $P_{1'}$ (serine), P_1 (cysteine), and P_6 (acid or small residue); however, P_5 through P_2 and $P_{2'}$ through $P_{4'}$ displayed a high degree of heterogeneity (Fig. 1.6). As this heterogeneity rules out amino acid sequence as the substrate recognition motif, a more careful examination was needed. Crystal structures of HCV NS3/4A protease bound to peptides corresponding to substrate cleavage sites revealed that, similar to HIV-1 protease, the substrates adopt a conserved shape defined as the substrate envelope [67]. Active site resistance mutations occur where the inhibitors protrude beyond this substrate envelope and contact the protease. Those positions are optimal for multi-drug resistance, as they are more important for inhibitor binding than for substrate recognition. Most resistance mutations in HCV NS3/4A cluster around the S_2 subsite where the relatively large P_2 moieties of PIs protrude beyond the substrate envelope and contact residues R155K, A156T, D168A, which mutate to confer resistance (Fig. 1.7) [68].

Substrate	P6	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'
3-4A	D	L	E	V	V	T	S	T	W	V
4A4B	D	E	M	E	E	C	S	Q	H	L
4B5A	E	C	T	T	P	C	S	G	S	W
5A5B	E	D	V	V	C	C	S	M	S	Y
TRIF	P	S	S	T	P	C	S	A	H	L
MAVS	E	R	E	V	P	C	H	R	P	S

Figure 1.6: Alignment of HCV HS3/4A substrate sequences
Genotype 1a NS3/4A viral substrates and host-cell substrates TRIF and MAVS. Sequences are shown from P_6 residues (N-terminus) to $P_{4'}$ (C-terminus) with conserved residues highlighted in yellow and scissile bond depicted in red dashed line.

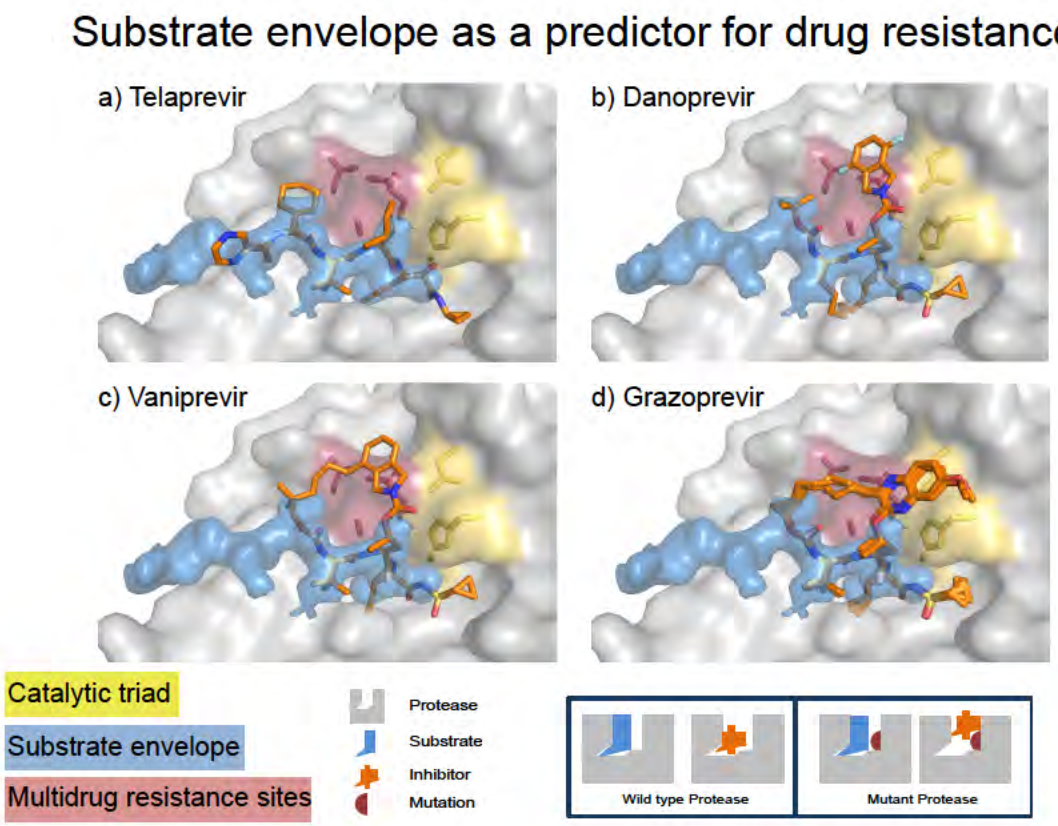


Figure 1.7: NS3/4A substrate envelope
HCV NS3/4A (grey) substrate envelope analysis of four inhibitors (orange sticks) complexes: **(A)** telaprevir, **(B)** danoprevir, **(C)** vaniprevir and **(D)** grazoprevir; catalytic triad and S_2 subsite are highlighted in yellow and red resptectively [68].

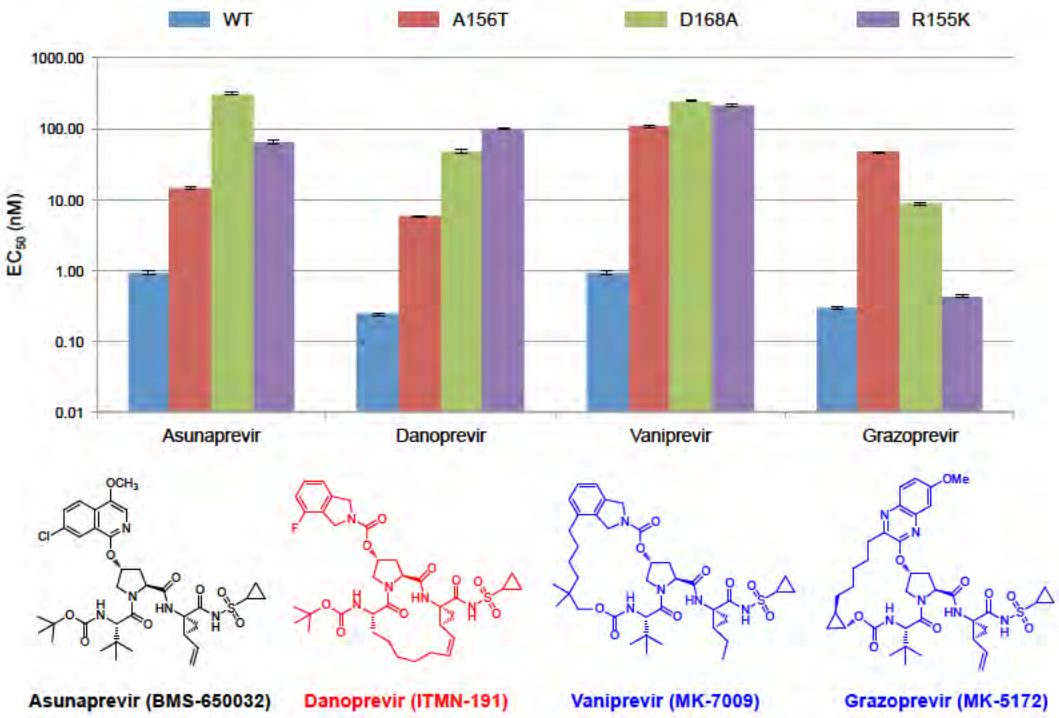


Figure 1.8: Effects of macrocyclization on NS3/4A inhibitor efficacy
The synergistic effect of chemical moieties within an inhibitor complicates our evaluation of the efficacy of these compounds.

1.4 Thesis Scope

Therapeutic advances in HCV PIs have generated a wealth of both structure–activity relationship (SAR) information, and structural perspectives. For instance, extended P_2 moieties such as isoquinoline, quinoxaline, and isoindoline are known to have potency enhancement properties, while the addition of a macrocycle to the inhibitor scaffold reduces the resistance profile.

This thesis attempts to fill critical gaps in our structural understanding of inhibitor binding modes as they relate to drug resistance and elucidate the basis for inhibitor failure in non-1 genotypes, specifically genotype 3. First, I introduce the concept of the structural basis for drug resistance in HCV protease by comparing the binding mode of ASV to wild-type protease and the multidrug resistance variant R155K. I demonstrate that very subtle rearrangements in active site side chains destabilize protease–inhibitor contacts thus reducing binding favorability. Next, I evaluate the role of macrocyclization in drug resistance by investigating the structural, dynamic and biophysical basis for enhanced resistance barrier of a $P_1 - P_3$ analogue of MK-5172 relative to the parent MK-5172 and linear analogue. Next, I investigate the molecular basis for the differential efficacy of PIs in GT-3 relative to GT-1 by determining the necessary and sufficient amino acid changes to GT-1 protease to recapitulate GT-3 activity and inhibition. Finally, I present novel strategies for robust inhibitor design by presenting a structural tool, which compiles our structural knowledge to predict drug efficacy in GT-3 and well as two novel inhibitor design strategies with promising applications for broad genotypic targeting.

PREFACE TO : CHAPTER II

This chapter of the thesis has been published in *American Chemical Society, Chemical Biology*

Soumana, D., Ali, A., and Schiffer, C. A. **Structural Analysis of Asunaprevir Resistance in HCV NS3/4A Protease** *ACS. Chem. Biol.* 2014 Nov 21;9(11):2485-90. doi: 10.1021/cb5006118

Author contributions : DS, AA and CAS designed the study. AA made the inhibitors. **DS** performed all the experiments, analyses and figure constructions. **DS**, AA and CAS interpreted the data and wrote the manuscript.

chapter II

Structural Analysis of Asunaprevir Resistance in HCV NS3/4A Protease

Chapter II

Structural Analysis of Asunaprevir Resistance in HCV NS3/4A Protease

2.1 Abstract

Asunaprevir (ASV), an isoquinoline-based competitive inhibitor targeting the hepatitis C virus (HCV) NS3/4A protease, is very potent *in vivo*. However, the potency is significantly compromised by the drug resistance mutations R155K and D168A. In this study three crystal structures of ASV and an analog were determined to analyze the structural basis of drug resistance susceptibility. These structures revealed that ASV makes extensive contacts with Arg155 outside the substrate envelope. Arg155 in turn is stabilized by Asp168, and thus when either residue is mutated, the enzyme's interaction with ASV's P_2 isoquinoline is disrupted. Adding a $\mathrm{P}_1 - \mathrm{P}_3$ macrocycle to ASV enhances the inhibitor's resistance barrier, likely due to poising the inhibitor to its bound conformation. Macrocyclic inhibitors with P_2 extension moieties avoiding interaction with the protease S_2 residues including Arg155 must be chosen for future design of more robust protease inhibitors.

2.2 Introduction

The hepatitis C virus (HCV) is the causal agent of viral hepatitis, cirrhosis and hepatocellular carcinoma. In the United States, HCV infects an estimated 3.2 million individuals with 17 thousand new cases annually. Globally, the HCV pandemic is a true public health problem and touches all corners of the world, infecting an estimated 170 million

people [69, 70].

While most of the steps in the viral lifecycle are currently targeted in antiviral therapy, NS3/4A protease is a key target especially as combination therapy becomes the paradigm in patient therapy [71]. The NS3/4A protease is a bifunctional enzyme consisting of an N-terminal serine protease domain, and a C-terminal domain that is a member of the DExH/D-box helicase superfamily II with ATPase, nucleic acid binding and helicase unwinding activities. The protease is responsible for processing the viral polyprotein and cleaving host factors involved in the immune response [27, 72, 73]. Thus, inhibiting the NS3/4A serves a dual purpose by preventing viral maturation and restoring the immune response.

There are currently three protease inhibitors (PIs) approved by the FDA (boceprevir, telaprevir, and most recently simeprevir) and several in advanced clinical trails, which are of a variety of chemical classes. Boceprevir and telaprevir are linear ketoamide compounds that form a reversible, covalent bond with the catalytic serine of NS3/4A protease [32, 74]. Noncovalent inhibitors include both linear (asunaprevir (ASV) [75], BI 201335 [76]) and macrocyclic compounds, containing either a $\mathrm{P}_1 - \mathrm{P}_3$ (danoprevir [77], simeprevir [78], or a $\mathrm{P}_2 - \mathrm{P}_4$ (vaniprevir [50], MK-5172 [79]) macrocycle. However, despite often having nanomolar potency against the wildtype (WT) enzyme, drug resistant mutants rapidly emerge [59]. With the current therapy handicapped by drug resistance, suboptimal pharmacokinetics and a lack of cross-genotype activity, the need for robust PIs with high barriers to resistance is paramount. Drug resistance in HCV NS3/4A is a complex interplay of molecular events whereby a change in the protease results in a decrease in inhibitor potency while retaining substrate processing and viral maturation. Most HCV PIs have a common peptidomimetic $\mathrm{P}_1 - \mathrm{P}_4$ scaffold but differ in their P_2 moiety. We have previously shown that the nature of this P_{2^*} extension moiety, which often protrudes from the substrate envelope [80], accounts for much of the resistance conferred by single-site mutations at residues Arg155, Ala156, and Asp168. In addition, we have characterized how

resistance susceptibility and potency of the NS3/4A inhibitors are dependent on both the P_2 extension and the macrocyclization [81].

ASV, a potent linear PI, is currently under development by Bristol-Myers Squibb and in phase III clinical trials. ASV is chemically characterized by a $\mathrm{P}_{1'}$ acylsulfonamide-linked cyclopropyl moiety, P_2 isoquinoline, and a $\mathrm{P}_1 - \mathrm{P}_4$ pepdidomimetic backbone shared by most PIs [82]. ASV has promising pharmacokinetics and as monotherapy lowers viremia by 2.8 to 3.5 orders of magnitude within 48 hours of treatment [83]. However, resistance challenges ASV due to the rapid evolution of the virus as relapse or viral breakthrough occurs mid-therapy. R155K, D168G, and I170T conferred low to moderate in genotype 1a and mutations at D168 conferred high resistance (16- to 280-fold) levels in genetype 1b [82, 83]. Thus, ASV needs to be supplemented with Peg-IFNα/RBV or another direct-acting antiviral (DAA) [84]. Consequently, understanding ASV's binding mode in the WT and multi-drug resistant (MDR) variants of NS3/4A translates into unraveling the molecular basis for isoquinoline inhibitors' drug resistance profile and sets the stage for improved structure-based drug design efforts.

In our previous study, we studied the impact of drug resistance mutations on the potency of ASV and ASV's $\mathrm{P}_1 - \mathrm{P}_3$ macrocyclic analogue (ASVmc) [81] to WT, R155K, R155K/V36M, A156T and D168 variants in both enzymatic studies and viral replicon assays. Both compounds had extremely good potency in replicon assays with 0.9 and 0.23 nM EC_{50}s, but lost 70/35- (R155K) and 341/84- (D168A) fold in affinity for ASV and ASVmc, respectively. In this study, we report the crystal structures of WT-ASV, R155K-ASV and WT-ASVmc complexes. ASV makes extensive intermolecular contacts with the catalytic residues His57 and Asp81, and the P_2 isoquinoline forms an aromatic stacking on Arg155 with a salt bridge between Arg155-Asp168. Drug resistance occurs when this electrostatic network is disrupted, thus triggering a domino-like effect to alter the van der Waals (vdw) contacts, especially at Asp81, and compromises the cation–π stacking inter-

action. We also show that adding a $P_1 - P_3$ macrocyle to ASV enhances the inhibitor's resistance barrier by restraining flexibility to poise the inhibitor for binding. Through a detailed structural analysis, we describe the atomic basis for ASV's resistance to R155K and show how macrocyclizing this inhibitor enhances its resistance barrier.

2.3 Results and Discussion

To structurally characterize the binding of ASV and assess the impact of mutations and macrocyclization on drug resistance, the crystal structures of three HCV-NS3/4A genotype 1a protease complexes were determined. These were WT-ASV, R155K-ASV, and WT-ASVmc, which diffracted to resolutions of 1.66-2.70 Å, respectively, in the $P2_12_12_1$ space group (Table 2.3.1).

2.3.1 Detailed Structural Analysis of ASV Binding

The HCV NS3/4A inhibitor ASV binds in the active site extensively packing around the catalytic His57 and Asp81 (Fig. 2.1b) with the P_2 isoquinoline moiety stacking both with the catalytic D81 and R155. ASV makes contacts with 27 protease residues (3 of which are minor) contributing to a total vdW energy of -44.5 kcal/mol (Fig. 2.2 and 2.3). While contact with His57 is the most extensive, Asp81, Ser139, Arg155, Ala156, and Ala157 also form extensive contacts, followed by residues 135-137. In contrast Asp168 makes relatively little direct contact with the inhibitor, suggesting that the large loss of affinity to the D168A variant results from indirect effects.

Complexes/crystal	WT-ASV	R155K-ASV	WT-ASVmc
PDB ID	4WF8	4WH6	4WH8
Resolution (Å)	1.6	1.9	2.7
Space Group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
Molecules in AU	1	1	2
Cell Dimensions			
a	39.35	54.92	58.43
b	60.89	58.41	54.96
c	80.87	60.09	59.88
β (°)	90	90	90
Completeness (%)	92.7	91.7	99.8
Measured Reflections	170394	61851	26199
Unique Reflections	22158	12673	6134
Average I/σ	11.6 (7)	9.2 (4.3)	10.5
Redundency	7.7	4.9	4.3
R_{sym} (%)	6.8 (31.5)	10 (32)	6.8 (21.2)
RMSD in			
Bonds (Å)	0.013	0.006	0.015
Angles (°)	1.67	1.5	1.6
R_{factor} (%)	19.2	16.8	17.6
R_{free} (%)	21.8	21.9	24.8

Table 2.1: Crystallography statistics for HCV NS3/4A–Asunaprevir complexes.

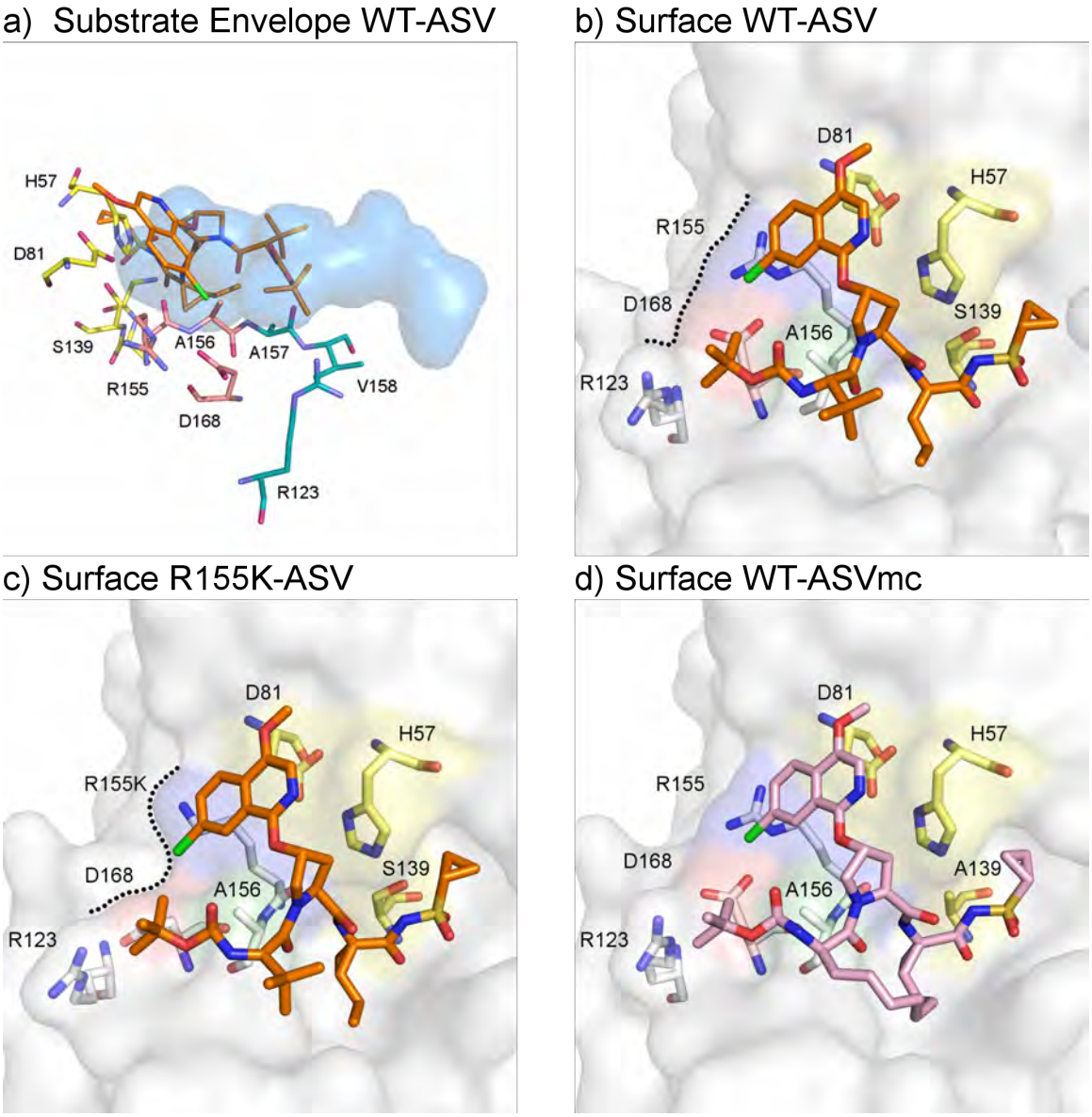


Figure 2.1: Structural binding mode of ASV and ASVmc
Asunaprevir's binding mode is reliant on protease S_2 residues: structure of HCV NS3/4A protease in complex with ASV. (a) ASV P_2 isoquinoline protrudes from the substrate envelope (blue volume), interacting with protease S_2 residues Arg155 and Asp168 (orange). (b) In the WT complex (PDB id: 4WF8), ASV (orange sticks) engages in stacking interactions with the catalytic D81 and S_2 residue R155. (c) In the R155K complex (PDB id: 4WH6), Lys155 is no longer stabilized by Asp168, which rotates toward R123 for electrostatic interaction. Consequently, ASV's P_2 isoquinoline loses important interaction binding surface (black dashed lines) and is destabilized. (d) ASVmc (PDB id: 4WH8, pink sticks) adopts a similar binding mode as ASV.

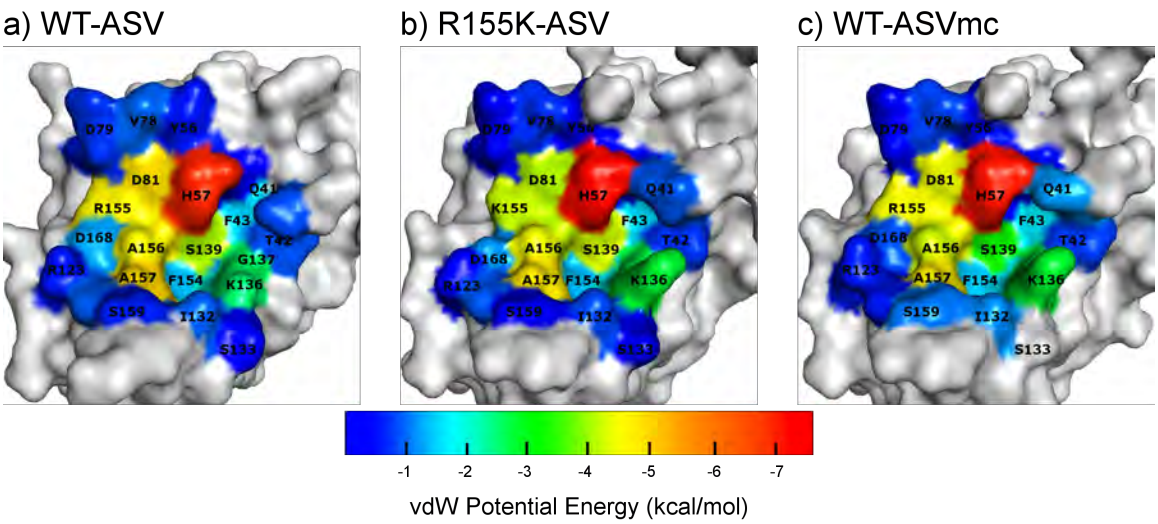


Figure 2.2: van der Waals contact energy landscape of ASV and ASVmc
vdW contacts are reduced in the drug-resistant mutant structure but maintained in ASVmc. Surface representation of protease-inhibitor vdW contact energy, where cool and warm colors indicate low and high vdW energy respectively. (a) In the WT-ASV structure, high contact energy is exhibited with the catalytic His57, Asp81 and S_2 R155, while low to moderate contact energy is observed in the $P_{1'}$-P_3 pocket. (b) In addition to altering this energy landscape by reducing inhibitor-protease vdW contacts at Asp81 and Lys155, the R155K mutation also changes the amino acid packing around Gln41, Val78, Asp168 and Arg123. (c) WT-ASVmc exhibits similar amino acid packing as WT-ASV around the S_2 and $P_{1'}$-P_3 pockets.

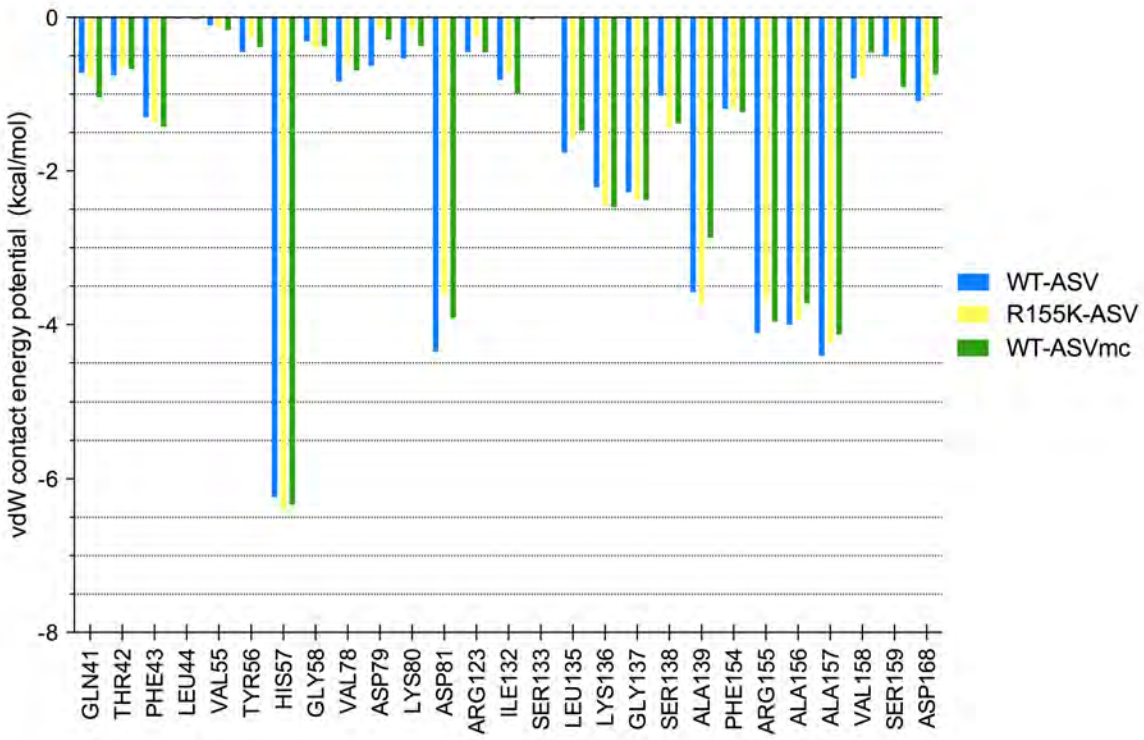


Figure 2.3: van der Waals contact energy of active site residues in ASV/ASVmc
The van der Waals contact energy of protease residues in WT-ASV (blue), R155K-ASV (yellow) and WT-ASVmc (green) structures, which are mapped onto the structure in the figure above. ASV's binding mode is characterized by high vdW contact energy with catalytic residues His57, Asp81 as well as S_2 residues Arg155, Ala156 and Ala157.

Eleven direct intermolecular and three water-mediated hydrogen bonds (H-bonds) are formed between the protease and ASV (Table 2.2), in particular at $\mathrm{P}_{1'}$, P_1 and P_3 moieties. As in other HCV NS3/4A PIs, the $\mathrm{P}_{1'}$ acylsulfonamide oxygens accept Hbonds from Ser139 and the oxyanion hole residue Gly137, while the nitrogen (N45) provides a hydrogen for His57 Nε2. The ASV P_1 carbonyl oxygen accepts three H-bonds from Ser139, Gly137 and Ser138 while ASV's P_1 amide nitrogen is a donor for Arg155's backbone oxygen. Finally, the P_3 carbonyl oxygen and amide hydrogen atoms both participate in H-bonds with the protease Ala157 backbone NH and CO respectively. Additionally, three hydrogen bonds are formed between neighboring water molecules and ASV's P_2 carbonyl oxygens, acylsulfonamide oxygens and P_4 carbamate oxygen.

Upon binding, ligands often induce conformational changes in the protein. Compared to the unliganded apo crystal structure (PDB id: 3RC6) [85] such changes are observed throughout the WT-ASV complex, specifically within the backbone Cα, illustrated by the presence of major peaks in the Cα distance difference plot (Fig. 2.4). In the WT-ASV complex, Asp168 has two ionic interactions with Arg155, in contrast to the apo structure of HCV NS3/4A where Asp168 is rotated away from Arg155 to interact with Arg123 instead (Fig. 2.5a). Thus in the WT-ASV complex, Asp168 stabilizes Arg155 thereby providing a binding surface for the isoquinoline and facilitating the aromatic stacking on the catalytic Asp81.

Table 2.2: Intra and intermolecular hydrogen bond of ASV complexes
Intra and intermolecular hydrogen bond analysis of protease-inhibitor complexes

	Complex Structure								
	WT-ASV			R155K-ASV			WT-ASVmc		
S2 Subsite Residues	Donnor Atom	Acceptor Atom	Distance (Å)	Donnor Atom	Acceptor Atom	Distance (Å)	Donnor Atom	Acceptor Atom	Distance (Å)
	His57 H_	Asp81 CO_1	1.8	His57 NH	Asp81 CO_1	1.9	His57 NH	Asp81 CO_1	1.7
	His57 NH	Asp81 CO_2	1.6	His57 H_1	Asp81 CO_2	1.7	His57 NH_1	Asp81 CO_2	1.6
	Arg155 NH	Asp168 CO	1.9	Lys155 NH	Asp168 CO	2.0	Arg155 NH	Asp168 CO	1.9
	Arg155 N_H	Asp168 O_1	1.6	Arg123 N_H	Asp168 CO_1	1.6	Arg155 NH12	Asp168 CO_1	2.0
	Arg123 N_H	Asp168 O_2	2.7	Arg123 NH21	Asp81 CO_2	2.1	Arg155 N_H	Asp168 CO_1	1.7
							Arg123 NH22	Asp168 CO_2	2.9
							Arg123 NH_	Asp168 CO_2	1.6
	Protease Atom	Inhibitor Atom	Distance (Å)	Protease Atom	Inhibitor Atom	Distance (Å)	Protease Atom	Inhibitor Atom	Distance (Å)
Protease Backbone	Ala157 CO	HN17	1.9	Ala157 CO	NHBA	1.9	His57 N_2	NH39	2.1
	Ala157 NH	O15	1.9	Ala157 NH	OAL	2.0	Ala157 NH	O34	2.1
	Arg155 CO	NH06	1.9	Lys155 CO	NHBB	2.0	Arg155 CO	H35	4.0
	His57 N_2	NH45	2.2	His57 N_2	NHBC	2.4	Arg155 CO	H13	2.3
	Ser138 NH	O44	2.9	Ser138 NH	OAK	2.6	Ser138 NH	O38	3.0
	Gly137 NH	O44	2.0	Gly137 NH	OAK	2.1	Gly137 NH	O38	2.3
	Gly137 NH	O48	2.3	Gly137 NH	OAM	2.2	Ala139 NH	O38	2.8
	Ser139 NH	O44	2.6	Ser139 NH	OAK	2.5	Gly137 NH	O41	2.4
	Ser139 NH	O47	3.1	Ser139 NH	OAN	2.7	Gly137 NH	O42	2.5
	Ser139 H_	O47	2.1	Ser139 H_	OAN	1.8	Ala139 NH	O42	2.7
	Water Molecule	Inhibitor Atom	Distance (Å)	Water Molecule	Inhibitor Atom	Distance (Å)			
Water Molecules	HW1 HOH 26	O20	2.1	HW1 HOH 29	OBF	1.9			
	HW1 HOH 122	O08	1.8	HW1 HOH 142	NAZ	2.0			
	HW2 HOH 135	O48	2.6	HW1 HOH 145	O	1.9			
				HW1 HOH 139	O	1.9			
				HW2 HOH 139	OAM	2.0			
				HW2 HOH 60	OAM	2.1			

2.3.2 Structural Changes Leading to Resistance

ASV is most susceptible to resistance mutations at amino acids 155, 156, and 168. These amino acids are all located at the NS3/4A S_2 binding pocket, where the inhibitor's isoquinoline moiety protrudes from the substrate envelope [68](Fig. 2.1a) and interacts extensively with the protease. In the R155K-ASV structure (Fig. 2.1b), the absence of a Nε on amino acid 155 prevents Lys155 from interacting with Asp168's Oδ, thus disrupting the 168-155 electrostatics, which provides an additional binding surface for the isoquinoline. As a result, Asp168 shifts towards Arg123 for electrostatic interactions (Fig. 2.5b). This alteration affects the entire binding interface with ASV as reflected in the altered van der Waals interactions (Fig. 2.2 and 2.6), while the inhibitor-protease hydrogen bonding network is maintained (Table 2.2). The extensive disruption of the ASV binding interface in the R155K-ASV complex is in agreement with inhibitor potency drop from 2.7 nM to 142.7 nM. A similar mechanism may also explain the loss of inhibitor potency in the D168A variant. While attempts to crystallize the ASV in complex with the D168A variant were unsuccessful, the interactions of the inhibitor are expected to be significantly disrupted as the Ala168-Arg155 ionic interaction will not exist. Consequently, ASV's affinity for the D168A variant drops by three orders of magnitude [81]. This structural resistance mechanism helps to explain the observed low to moderate resistance patterns in genotype 1a. Resistance to mutations at residues Arg155 and Asp168 is not unique to ASV, but rather these are signatures of resistance for many NS3/4A inhibitors such as danoprevir, vaniprevir, sovaprevir and faldaprevir [63, 86, 87]. However, in other genotypes, such as genotype 1b, the patterns of resistance are altered; for instance, to acquire the R155K mutation two nucleotide substitutions would be required [88], thus genetically altering the resistance barrier.

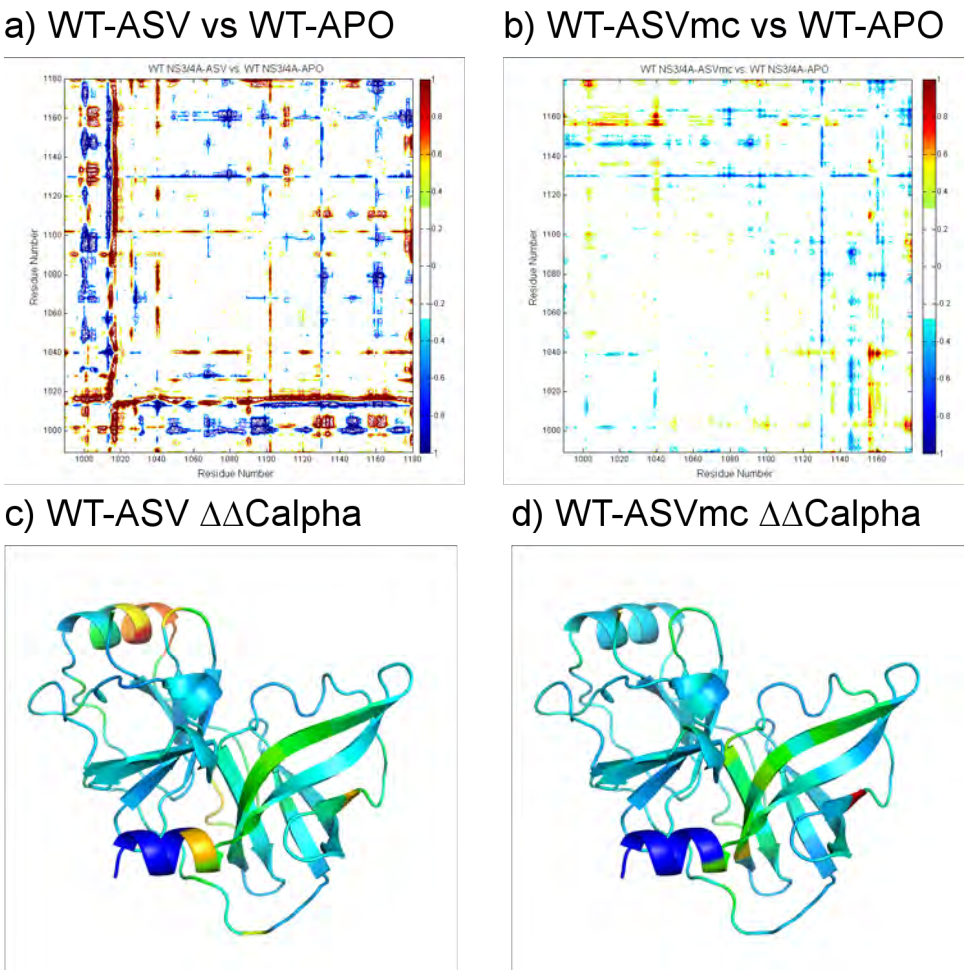


Figure 2.4: Structural impact of ASV macrocyclization
$P_1 - P_3$ macrocyclization of Asunaprevir retrains protease conformation: Cα displacement analysis of inhibitor complexes relative to the unliganded WT protease (PDB id: 3RC6). The backbone in WT-ASVmc (b) is less disturbed than that in the WT-ASV complex (a). Mapping the displacement onto the apo protease structure shows the various secondary structural elements whose conformations are allosterically affected by inhibitor binding (c and d).

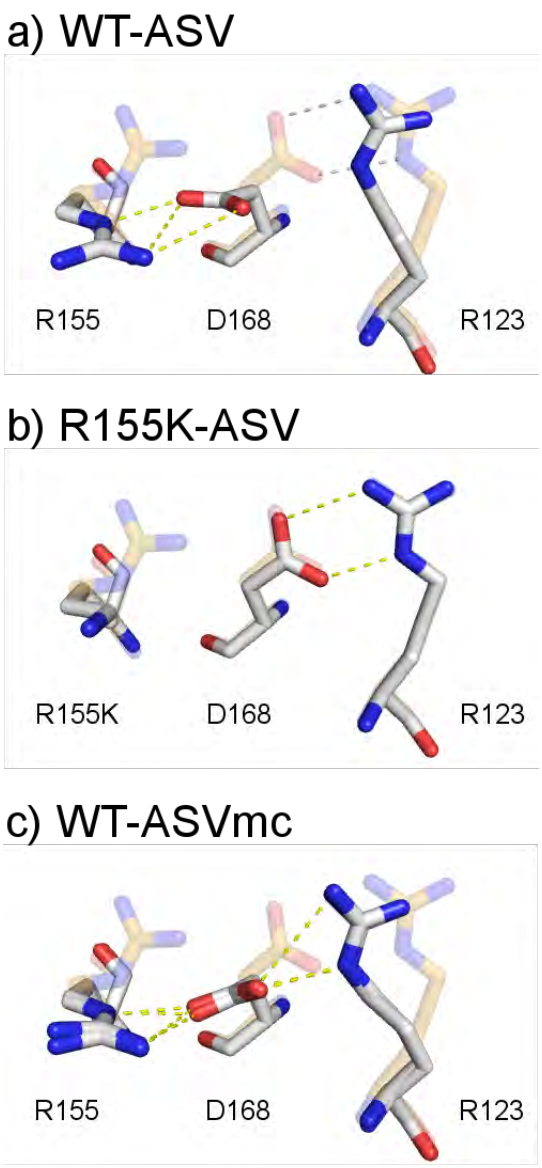


Figure 2.5: Active site electrostatic network in ASV complexes
Electrostatic network spanning Arg155-Asp168-Arg123 is not conserved in the mutant structure: (a) in the apo NS3/4A (orange transparent sticks, PDB: 3RC6), Asp 168 is oriented toward Arg 123 for interactions (grey dashed lines). The WT-ASV complex (white sticks) has Asp168's carbonyl oxygens oriented towards Arg155's nitrogens forming a salt-bridge (yellow dashed lines). (b) In the R155K complex (white sticks), Asp168 rotates away from Lys155 and towards Arg123 for interaction, as is observed in the apo structure. (c) In the WT-ASVmc (white sticks), Asp168's position enables an extended salt-bridge formation spanning Arg155-Asp168-Arg123. Both conformers of the crystal structure are represented.

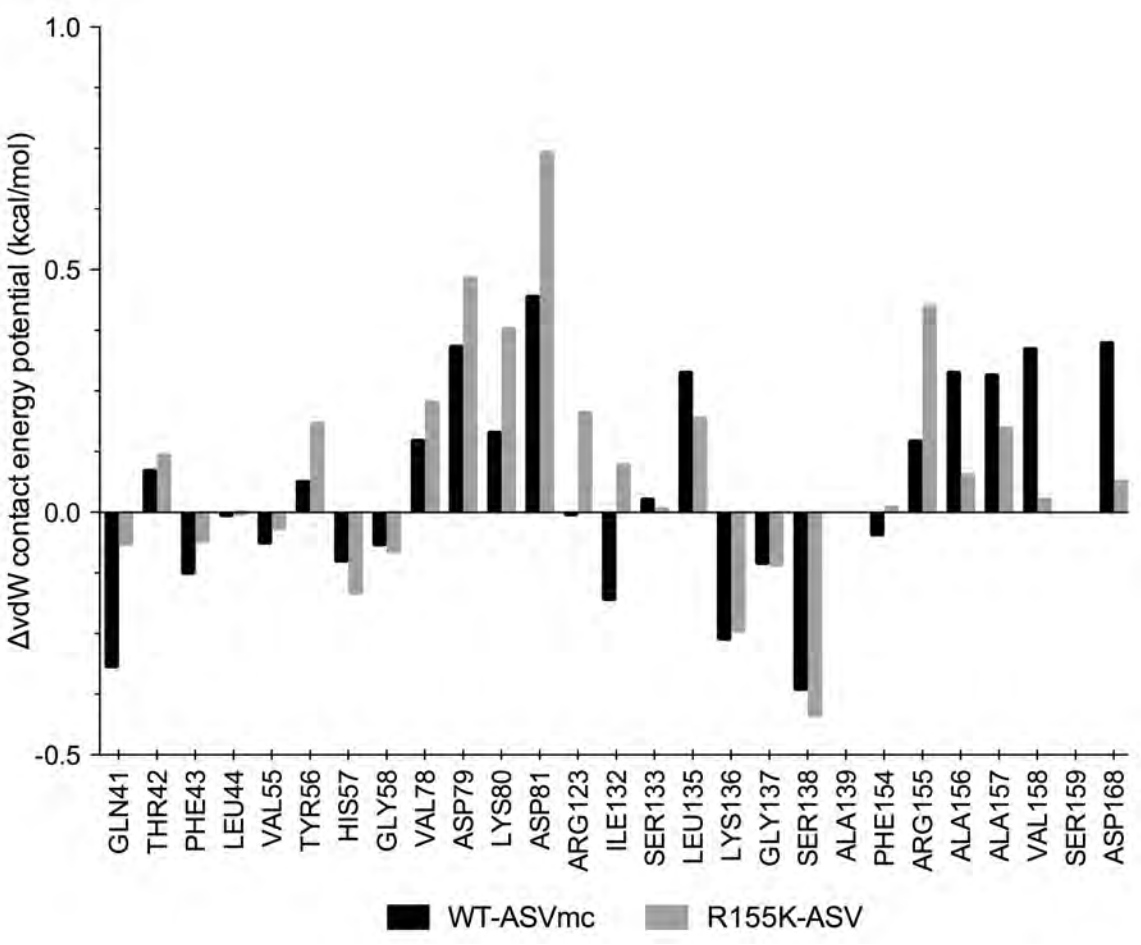


Figure 2.6: Change in van der Waals contact energy in ASV complexes
Change in van der Waals contact energy of protease residues with the inhibitor in WT-ASVmc (black) and R155K-ASV (grey) relative to the WT-ASV structure.

2.3.3 Improving Asunaprevir's Resistance Barrier

Although the crystal structure of WT protease with the $P_1 - P_3$ ASV macrocyclic analog revealed a similar binding mode to the WT-ASV complex (Fig. 2.1d), several differences were observed. Compared with the apo structure, the WT-ASVmc complex does not show as extensive changes in the protease backbone relative to the WT-ASV structure, perhaps accounting for some of the enhanced binding affinity (Fig. 2.4). Compared to ASV, the macrocyclic ASVmc has altered vdW interactions at the binding site (Fig. 2.2). Specifically subtle alterations between the structures were observed at $P_{1'}$, where the cyclopropyl group rotates toward the catalytic histidine, the P_{2^*} proline bears a less puckered pyrrolidine ring and the $P_3 - P_4$ peptidomimetic bond is displaced out of the active site by 1 Å. Within the active site, the major difference lies in the S_2 pocket electrostatic network. While both carbonyl oxygens of Asp168 in WT-ASV are oriented towards R155's Nε and Nη (Fig. 2.5A), Asp168 in WT-ASVmc rotates by 8.7° about the Cα-Cβ-O angle, leading to the formation of an extensive electrostatic network connecting Arg155-Asp168A-Arg123 (Fig. 2.5c).

In agreement with our structural data, ASV and ASVmc displayed similar enzyme inhibition constants with the WT protease (2.6 and 1.2 nM, respectively) [81]. While not significantly improving potency to the WT protease, the addition of this macrocycle considerably improved the inhibitor's resistance barrier. Indeed, all three MDR variants R155K, A156T and D168A experienced reduced drug resistance susceptibility with lower fold changes in EC_{50}'s [81]. In the free state, small molecules typically adopt a dynamic ensemble of conformations; this is particularly true for those containing proline amide groups, which have the ability to form *cis* and *trans* conformations. The $P_1 - P_3$ macrocycle restricts ASVmc in a *trans* conformation, thereby potentially decreasing the entropic penalty associated with conformational reconfiguration required for binding compared to the linear ASV.

2.4 Conclusion

In summary, asunaprevir is a potent competitive inhibitor targeting the HCV NS3/4A protease, with clinical and *in vitro* resistance susceptibility to protease mutations at Arg155, Asp168 and Ala156. Similar to other inhibitors that we and others have anlyzed [59,68,80], this resistance profile is the result of the inhibitors reliance on packing to the S_2 site residues including Arg155 beyond the substrate envelope. Avoiding such extensive S_2 interactions and restraining the *trans* conformation via macrocyclization may be viable strategies to design inhibitors that are more robust against resistance.

2.5 Materials and Methods

2.5.1 Protein Handling and Data Collection

Protein expression, crystallization, data collection and processing was performed as previously reported [68]. Briefly, cells were harvested after 5 hours of expression, pelleted, and frozen at -80 °C for storage. Cell pellets were thawed, resuspended in 5 mL/g of resuspension buffer (50 mM phosphate buffer, 500 mM NaCl, 10% glycerol, 2 mM β-ME, pH 7.5) and lysed with a cell disruptor. The soluble fraction was retained, applied to a nickel column (Qiagen), washed with resuspension buffer, and eluted with resuspension buffer supplemented with 200 mM imidazole. The eluent was dialyzed overnight (MWCO 10 kD) to remove the imidazole, and the His-tag was simultaneously removed with thrombin treatment. The nickel-purified protein was then flash frozen and stored at -80 °C.

2.5.2 Crystallization

The above-mentioned protein solution was thawed, concentrated to 3 mg/mL and loaded on a HiLoad Superdex75 16/60 column equilibrated with gel filtration buffer, 25 mM 2-(N-Morpholino)ethanesulfonic (MES), 500 mM NaCl, 10% glycerol, 30 mM zinc chloride, and 2 mM 1,4-Dithiothreitol, pH 6.5. The protease fractions were pooled and concentrated to 20-25 mg/mL with an Amicon Ultra-15 10 kD device (Millipore). The concentrated samples were incubated for 1 hour with 1-3 molar excess of inhibitor. Diffraction-quality crystals were obtained by mixing equal volume of concentrated protein solution with precipitant solution (20-26% PEG-3350, 0.1 M sodium MES buffer, 4% ammonium sulfate, pH 6.5) in 24-well VDX hanging drop trays.

2.5.3 Data Collection and Structure Solution

X-ray diffraction data were collected either at Advanced Photon Source LS-CAT 21-ID-F, at our in-house Rigaku Saturn x-ray system. Diffraction intensities were indexed, integrated and scaled using the program HKL2000. All structure solutions were generated using molecular replacement with PHASER [89]. The B chain model of viral substrate product 4A-4B (3M5M) was used as the starting model for all structure solutions. Initial refinement was carried out in the absence of modeled ligand, which was subsequently built in during later stages of refinement. Subsequent crystallographic refinement was carried out within the Phenix program suite [90], with iterative rounds of TLS or restrained refinement until convergence was achieved. The final structures were evaluated with MolProbity [91] prior to deposition in the Protein Data Bank. To limit the possibility of model bias throughout the refinement process, 5% of the data were reserved for the free R-value calculation. Interactive model building and electron density viewing was carried out using the program.

2.5.4 Structural Analysis

Superpositions were performed in PyMOL using the Cα atoms of the active site protease residues 137-139 and 154-160. The A chain of WT-ASV complex was used as the reference structure for each alignment. During the preparation of this manuscript, a WT-ASV structure (PDB id: 4NWL) similar to our WT-ASV was released [92]. We chose to use our structure for analysis because of its improved statistics and resolution.

Van der Waals contact energy potential and distance difference plots were determined as previously described [93, 94].

PREFACE TO : CHAPTER III

This chapter of the thesis has been published in

American Chemical Society, Chemical Biology

Soumana, D., Yilmaz N. K., Prachanronarong, K., Aydin, C., Ali, A., and Schiffer, C. A. (2015). **Structural and Thermodynamic Effects of Macrocyclization in HCV Inhibitor Grazoprevir (MK-5172)** *ACS. Chem. Biol.* 2015 Dec.

Author contributions : **DS**, AA, CAS designed the study. AA made the inhibitors. **DS** performed all the experiments (including MD simulations), analyses and figure constructions. CA contributed preliminary data, KP provided technical assistance. **DS**, NYK and CAS interpreted the data and wrote the manuscript.

chapter III

Structural and Thermodynamic Effects of Macrocyclization in HCV Inhibitor Grazoprevir (MK-5172)

Chapter III

Structural and Thermodynamic Effects of Macrocyclization in HCV Inhibitor Grazoprevir (MK-5172)

3.1 Abstract

Recent advances in direct-acting antivirals against Hepatitis C Virus (HCV) have led to the development of potent inhibitors, including MK-5172, that target the viral NS3/4A protease with relatively low susceptibility to resistance. MK-5172 has a $P_2 - P_4$ macrocycle and a unique binding mode among current protease inhibitors where the P_2 quinoxaline packs against the catalytic residues H57 and D81. However, the effect of macrocyclization on this binding mode is not clear, as is the relation between macrocyclization, thermodynamic stabilization, and susceptibility to the resistance mutation A156T. We have determined high-resolution crystal structures of linear and $P_1 - P_3$ macrocyclic analogs of MK-5172 bound to WT and A156T protease and compared these structures, their molecular dynamics and experimental binding thermodynamics to the parent compound. We find that the unique binding mode of MK-5172 is conserved even when the $P_2 - P_4$ macrocycle is removed or replaced with a $P_1 - P_3$ macrocycle. While beneficial to decreasing the entropic penalty associated with binding, the constraint exerted by the $P_2 - P_4$ macrocycle prevents efficient rearrangement to accommodate the A156T mutation, a deficit alleviated in the linear and $P_1 - P_3$ analogs. Design of macrocyclic inhibitors against NS3/4A needs to achieve the best balance between exerting optimal conformational constraint for enhancing potency, fitting within the substrate envelope and allowing adaptability to be robust

against resistance mutations.

3.2 Introduction

Hepatitis C virus (HCV) causes chronic liver infection that affects about 3% of the global population and is the main cause of hepatitis, cirrhosis and liver cancer [1, 69, 70]. HCV has highly error prone replication and therefore is a rapidly evolving, highly diverse virus with six known genotypes (gt) and multiple subtypes [2, 3]. Before the recent availability of direct-acting antivirals (DAAs), the standard of care consisted of pegylated-interferon (Peg-INF) and ribavirin (RBV) with moderate to low rates of cure across genotypes and low tolerability [1, 95, 96]. Current efforts aim to determine the best-in-class DAAs that target several viral proteins including the viral entry protein, the NS3/4A protease, the NS5A and NS5B non structural proteins [97] and host microRNAs [98] either individually or in combination. Four FDA-approved HCV inhibitors (telaprevir [31, 32], boceprevir [35], simeprevir [39], and most recently, paritaprevir [40]) target the NS3/4A protease. The NS3/4A protein is a bifunctional enzyme containing an N-terminal serine protease domain (amino acids 1–180) with the classic catalytic triad (S139, H57, D81) of the chymotrypsin superfamily, and a C-terminal DExH/D-box helicase of superfamily II with NTPase activity [99–102]. The NS3/4A protease is responsible for cleaving the viral polyprotein and host factors involved in the innate immune response, including TRIF and MAVS. Thus, targeting the NS3/4A protease achieves a two-pronged attack on the virus by preventing viral maturation and restoring the immune response [26, 27, 72, 103].

As the HCV NS3/4A inhibitors are a key component of combination therapy and increasing the cure rate across HCV genotypes, many more are currently in advanced clinical trials. Among these inhibitors, MK-5172 stands out with relatively high pan-genotypic potency [51]. MK-5172 shares the same peptidomimetic core $P_{1'}$-P_3 scaffold as several

other HCV PIs (danoprevir, asunaprevir and vaniprevir) but is distinct in its P_4 capping, P_2 quinoxaline moiety connected to the P_2 proline via an ether linkage, and finally, its $P_2 - P_4$ macrocycle (Fig. 3.3.1 A) [49, 79]. While the majority of NS3/4A inhibitors are susceptible to single site mutations R155K, A156T and D168A, MK-5172 is more robust against resistance with the exception of A156T in vitro [51, 68]. These single site mutations are able to confer resistance to most inhibitors, depending mostly on the inhibitor P_2 moiety [51, 68, 81]. The large P_2 moiety of inhibitors protrude out the substrate envelope and extensively contact R155 and D168 residues that mutate to confer resistance [68, 80]. The co-crystal structure we determined showed that MK-5172 thwarts resistance at these sites as the P_2 quinoxaline moiety largely interacts with residues of the catalytic triad [68]. This binding mode explains MK-5172's relatively high potency against R155K and D168A and provides a unique barrier to resistance, as the catalytic residues cannot mutate without compromising proteolytic activity.

Another distinct chemical feature of MK-5172 is the macrocycle that connects the P_2 quinoxaline moiety to the P_4 capping group. Macrocyclization is a general strategy employed to increase inhibitor potency in a wide array of drug targets [104, 105]. By introducing a conformational constraint to the inhibitor, the structure can be pre-organized in a binding-competent conformation to decrease the entropic penalty associated with target association. However the effect of macrocyclization on the energetics of binding for MK-5172 has not been quantified. In fact, there is no previously published data on the thermodynamics of binding for any active-site/ competitive HCV protease inhibitor. For NS3/4A protease, we have previously found by enzymatic and antiviral assays that $P_1 - P_3$ and $P_2 - P_4$ macrocyclic inhibitors are more potent compared to their linear analogs, and that the location of macrocycle significantly affects resistance profiles [81]. In particular, the $P_1 - P_3$ macrocyclic analog of MK-5172 is more robust against the A156T resistance mutation, however whether the unique binding mode on the catalytic triad is retained and the

thermodynamic and energetic implications of alternative macrocyclization status were not known.

To elucidate the structural basis of resistance and activity for MK-5172, its linear (5172-linear) and $\mathrm{P}_1 - \mathrm{P}_3$ macrocyclic (5172-mc$\mathrm{P}_1 - \mathrm{P}_3$) analogs (Fig. 3.3.1A) and their susceptibility to A156T, we determined high-resolution inhibitor co-crystal structures with both WT and A156T mutant genotype 1a NS3/4A protease domain. In addition to analyzing the structures of the six protease–inhibitor complexes, we determined how binding energetics and conformational dynamics change with macrocyclization status using calorimetry and molecular dynamics simulations. We find that the unique binding mode of MK-5172 is independent of the macrocycle, and that the rigidity imposed by the $\mathrm{P}_2 - \mathrm{P}_4$ macrocycle prevents inhibitor rearrangement to be able to efficiently accommodate the A156T mutation. The flexibility of the linear analog allows subtle conformational rearrangements to accommodate the mutation but the associated entropic disadvantage results in lower potency. Lastly, 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ benefits from the entropic advantage of macrocyclization while still allowing conformational rearrangements to reduce susceptibility to A156T. The preservation of the parent compounds binding mode in the absence of the $\mathrm{P}_2 - \mathrm{P}_4$ macrocycle provides further opportunities to optimize this class of inhibitors, while avoiding known sites of drug resistance.

Inhibitors	5172-mc$P_1 - P_3$		5172-Linear	
Complexes	WT	A156T	WT	A156T
PDB Code	5EPN	5EPY	5EQQ	5ETX
Resolution (Å)	2.3	2.3	1.6	2.3
Space Group	P212121	P212121	P212121	P21
Twin Law	-	-	-	-k,-h,-l
Twin Fraction (%)	-	-	-	0.1
Molecules in AU	1	1	1	4
Cell Dimensions				
a	55.2	54.9	54.71	65.4
b	58.5	58.4	58.42	63.2
c	60.1	60.0	59.96	92.3
$\beta(^\circ)$	90	90	90	91.6
Completeness (%)	96.2	99.9	99	99.9
Measured Reflections	38720	39511	156955	129154
Unique Reflections	8793	9070	23941	31801
Average I/sigma	9	9.1	9.1	18.4
Redundancy	4.4	4.1	6.6	4.1
Rsym (%)	6.7 (18.1)	6.8 (20)	7.9 (32.9)	7.2 (11.3)
RMSD in				
Bonds (Å)	0.005	0.011	0.014	0.015
Angles ($^\circ$)	1.3	1.3	1.7	1.7
Rfactor (%)	15.2	17.3	16.3	20
Rfree (%)	19.9	21.6	19.7	25.4

Table 3.1: Crystallographic statistics of MK-5172 analogs
Structural comparison was performed with 3SUD (WT–MK-5172) and 3SUD (A156T–MK-5172)

3.3 Results

We have determined high-resolution crystal structures of MK-5172 analogs in complex with the WT genotype 1a NS3/4A protease and A156T variant (WT–5172-linear, A156T–5172-linear, WT–5172-mc$\mathrm{P_1-P_3}$ and A156T–5172-mc$\mathrm{P_1-P_3}$) (Fig. 3.1). Three of these structures are in the same P212121 space group while the A156T–5172-linear is in P21 with four molecules in the asymmetric unit (only one molecule was chosen for analysis; see Materials and Methods). These four new and the two previously determined crystal structures of MK-5172 (PDB IDs 3SUD and 3SUG, for WT and A156T variants, respectively) [68] were analyzed comparatively, specifically for details of inhibitor–protease interactions.

3.3.1 The unique binding mode of MK5172 is independent of the macrocycle

Unlike all other HCV PIs with known co-crystal structures, MK-5172 interacts with the catalytic triad in a unique conformation where the $\mathrm{P_2}$ quinoxaline moiety packs largely against the catalytic residues H57 and D81 (Fig. 3.3.1) . The $\mathrm{P_1-P_4}$ peptidomimetic inhibitor scaffold spans the $\mathrm{S_1-S_4}$ binding pockets interacting with the carbonyl oxygens of R155 and A157 as well as the N of A157. The $P_{1'}$ acylsulfonamide is positioned in the oxyanion hole and hydrogen bonds to H57, G137 and S139. This binding mode is unchanged when the $\mathrm{P_2-P_4}$ macrocycle is removed (5172-linear) or replaced with a $\mathrm{P_1-P_3}$ macrocycle (5172-mc$\mathrm{P_1-P_3}$). Therefore the binding mode of MK-5172 is a function of the $\mathrm{P_2}$ moiety rather than the macrocycle.

Despite conservation of the overall binding mode, the potency of MK-5172 and its analogs vary greatly against WT and A156T variants [68, 81]. MK-5172 inhibits WT protease with a K_i of 0.14 nM, lower than those of linear and $\mathrm{P_1-P_3}$ analogs (9.5 and 2.0 nM

respectively), but loses 74-fold potency against the A156T variant. The $\mathrm{P}_1 - \mathrm{P}_3$ and linear analogs lose 29 and 230-fold potency against A156T relative to WT protease. To gain insights into the molecular basis of this difference, next we analyzed how the details of protease–inhibitor interactions are affected by the macrocyclization status and the A156T drug resistance mutation.

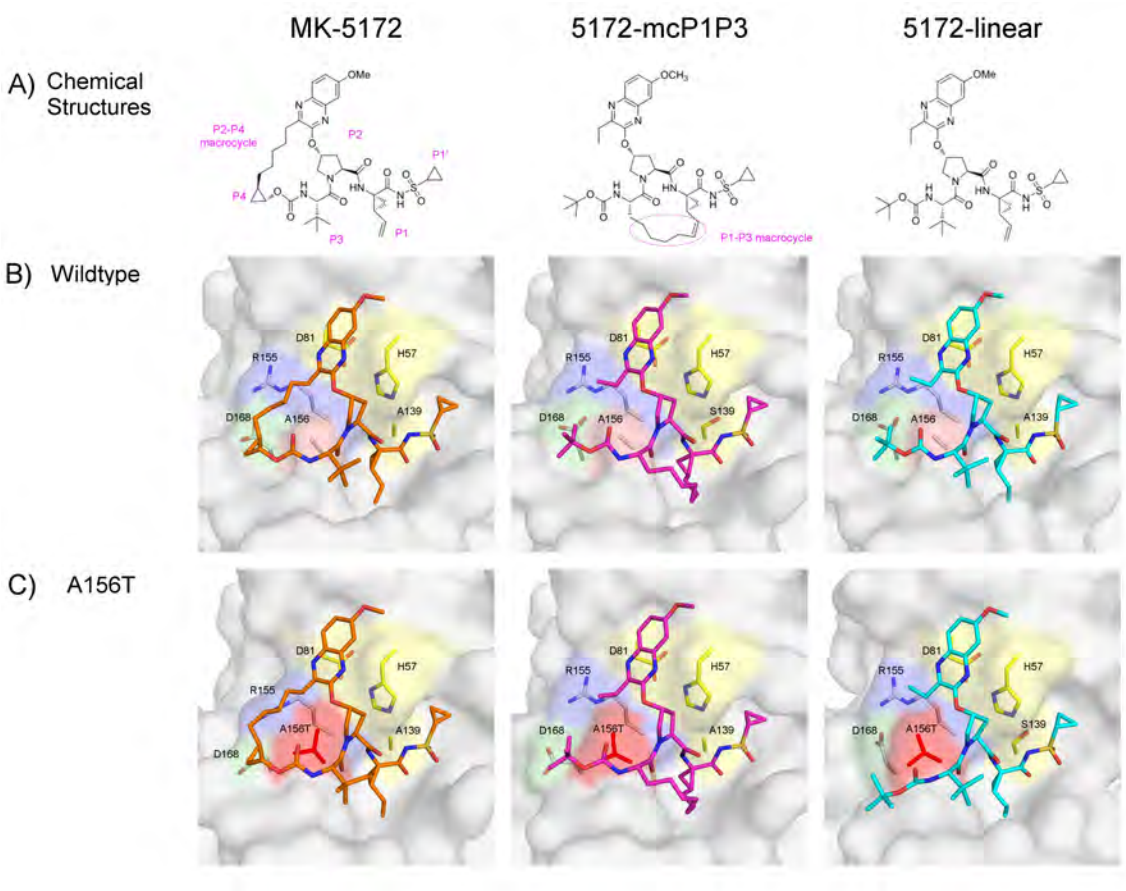


Figure 3.1: MK-5172 and analogs binding to HCV NS3/4A protease
(**A**) The chemical structures of MK-5172, 5172-mc$P_1 - P_3$ and 5172-linear. The $P_{1'}$ to P_4 moieties are indicated on the MK-5172 structure. Binding of the compounds at the active site in (**B**) WT and (**C**) A156T protease. The absence of the $P_2 - P_4$ macrocycle does not disrupt the P_2 quinoxalines interactions with the catalytic aspartate/histidine residues. The catalytic triad is highlighted in yellow, R155 in blue, D168 in green and A/T156 in red.

3.3.2 A156T alters the packing of the macrocyclic analogs at the active site

To analyze the details of inhibitor packing at the active site, the relative van der Waals (vdW) interaction energies between the inhibitor and protease residues in the crystal structures were calculated (Fig. 3.2 and 3.3). In line with the conservation of the overall binding mode, the strongest inhibitor–protease interaction occurs with the catalytic H57 and D81 in all WT complex structures (Fig. 3.2B), with similar total vdW interaction energies (WT–MK-5172: -43.2, WT–5172-linear: -42.3, WT–5172-mc$\mathrm{P_1 - P_3}$: -45.1 kcal/mol). Nevertheless, there are subtle alterations in the details of inhibitor packing. For instance, compared to the parent compound, vdW interactions with D81 are 0.8 kcal/mol weaker in the linear analog, indicating subtle changes in the stacking of the $\mathrm{P_2}$ quinoxaline moiety. In addition, inhibitor contacts with R155, Y56 and Q41 are weaker in the analogs compared to the parent MK-5172, in agreement with decreased potency against the WT protease (Fig. 3.3).

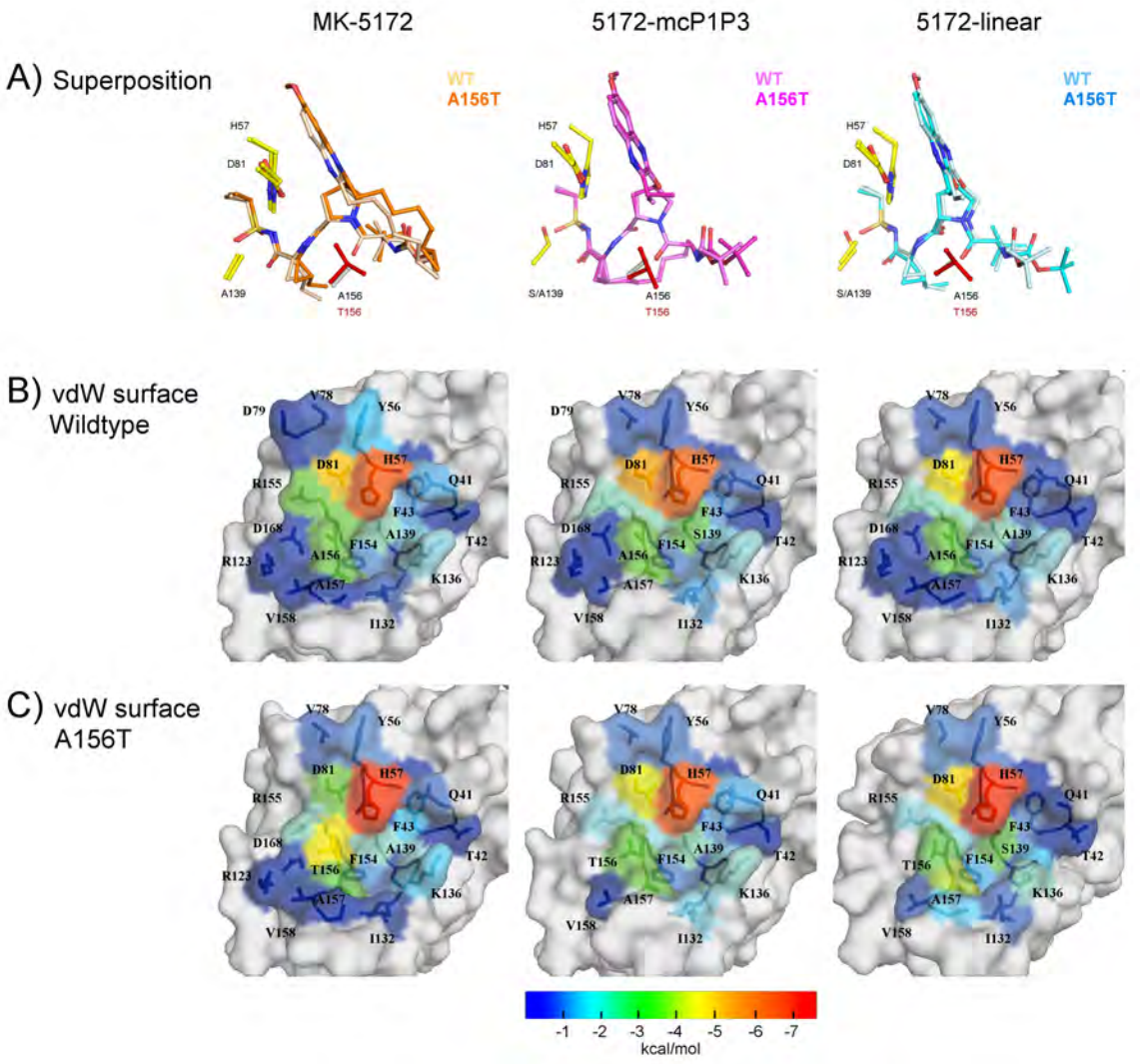


Figure 3.2: Surface representation of van der Wals contacts in MK-5172 analogs Changes in the packing of the inhibitors due to A156T at the HCV NS3/4A active site evaluated by the vdW contact energies. (**A**) Superposition of the inhibitor bound to WT and A156T protease. The vdW contact energy with the inhibitor mapped onto the protease surface in (**B**) WT and (**C**) A156T variants. The warmer (red) and cooler (blue) colors indicate more and less contacts with the inhibitor, respectively.

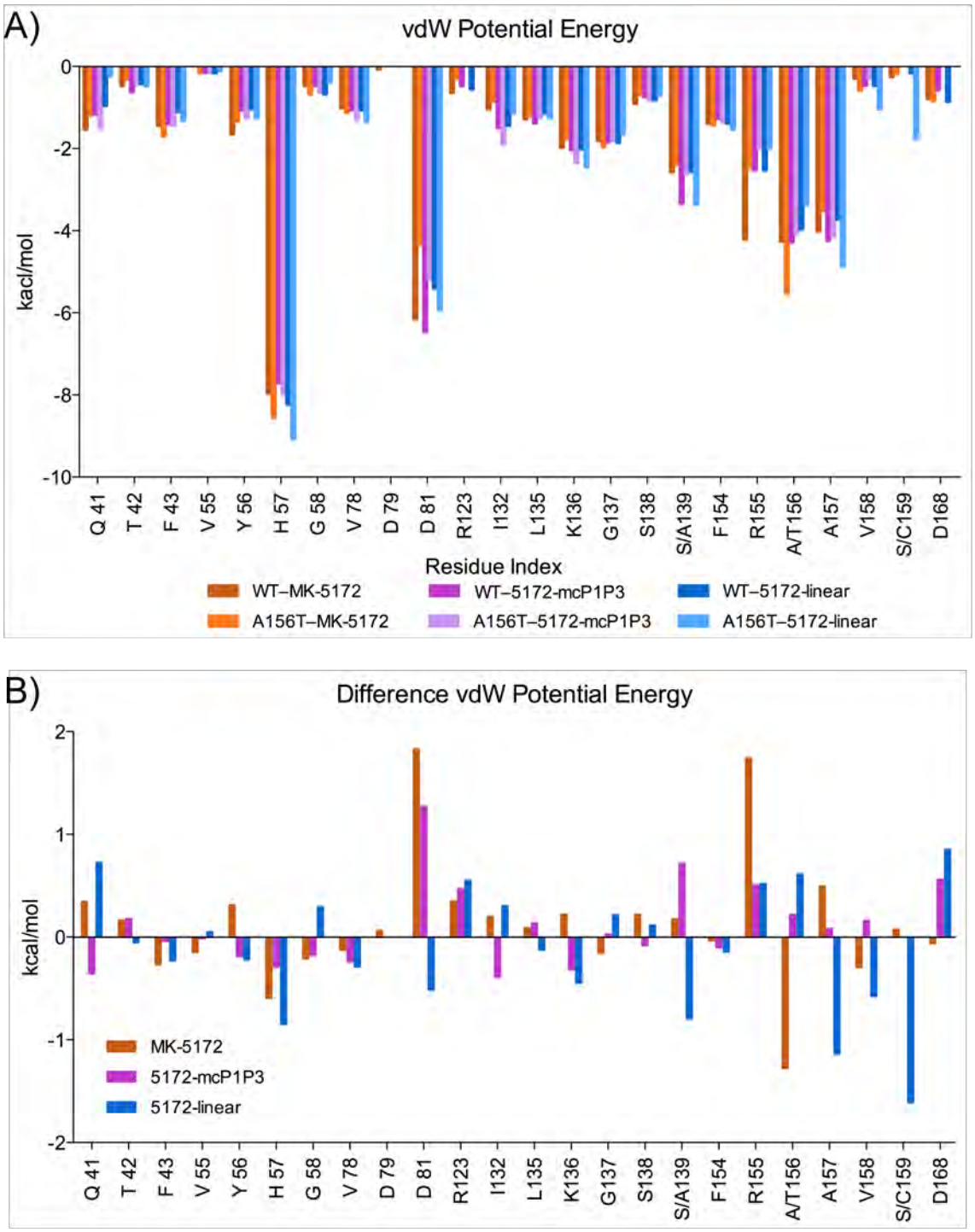


Figure 3.3: van der Waals contact energy of protease–inhibitor interactions (**A**) van der Waals contact energy of active site resides in MK-5172 (orange), 5172-mcP1P3 (magenta) and 5172-Linear (blue), with wild type and A156T complexes shown in dark and light shading, respectively. (**B**) Changes in vdW potential observed in A156T complexes relative to wild type.

The primary drug resistance mutation A156T alters packing of MK-5172 and its analogs at the active site to varying extents (Fig. 3.2C). In complexes with the mutant protease, the inhibitor needs to rearrange to avoid a steric clash with the larger Thr sidechain. In the WT complex conformation, the $\mathrm{P}_2 - \mathrm{P}_4$ macrocycle of MK-5172 would sterically clash with the larger Thr sidechain. As a result, the entire inhibitor shifts relative to the WT complex in the A156T–MK-5172 structure, with the macrocycle moving away from residues R155, A157 and the P_2 quinoxaline away from D81 toward the catalytic histidine (Fig. 3.2, 3.4). We have previously described this steric effect as the molecular mechanism behind MK-5172's susceptibility to A156T [68]. When the $\mathrm{P}_2 - \mathrm{P}_4$ macrocycle is removed, the shift due to the A156T mutation is much more subtle with the 5172-linear's P_2 quinoxaline maintaining stacking interactions with H57 and D81. Instead, the P_4 tert-butyl moiety adapts a different conformation to accommodate the larger Thr sidechain and loses packing against residues 155 and 156, but establishes more contacts with residues 157–159. Similarly, in the A156T–5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ complex, rearrangement of the P_4 moiety is able to accommodate the larger Thr sidechain to avoid a steric clash without severely compromising the stacking interactions of the P_2 quinoxaline. Overall, 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ maintains WT-like contacts at the protease active site with the exception of decreased packing against D81, but to a lesser extent than that of the parent MK-5172. Hence, 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ analog benefits from stabilization by the $\mathrm{P}_1 - \mathrm{P}_3$ macrocyclization without paying a steric penalty due to the Thr mutation at position 156.

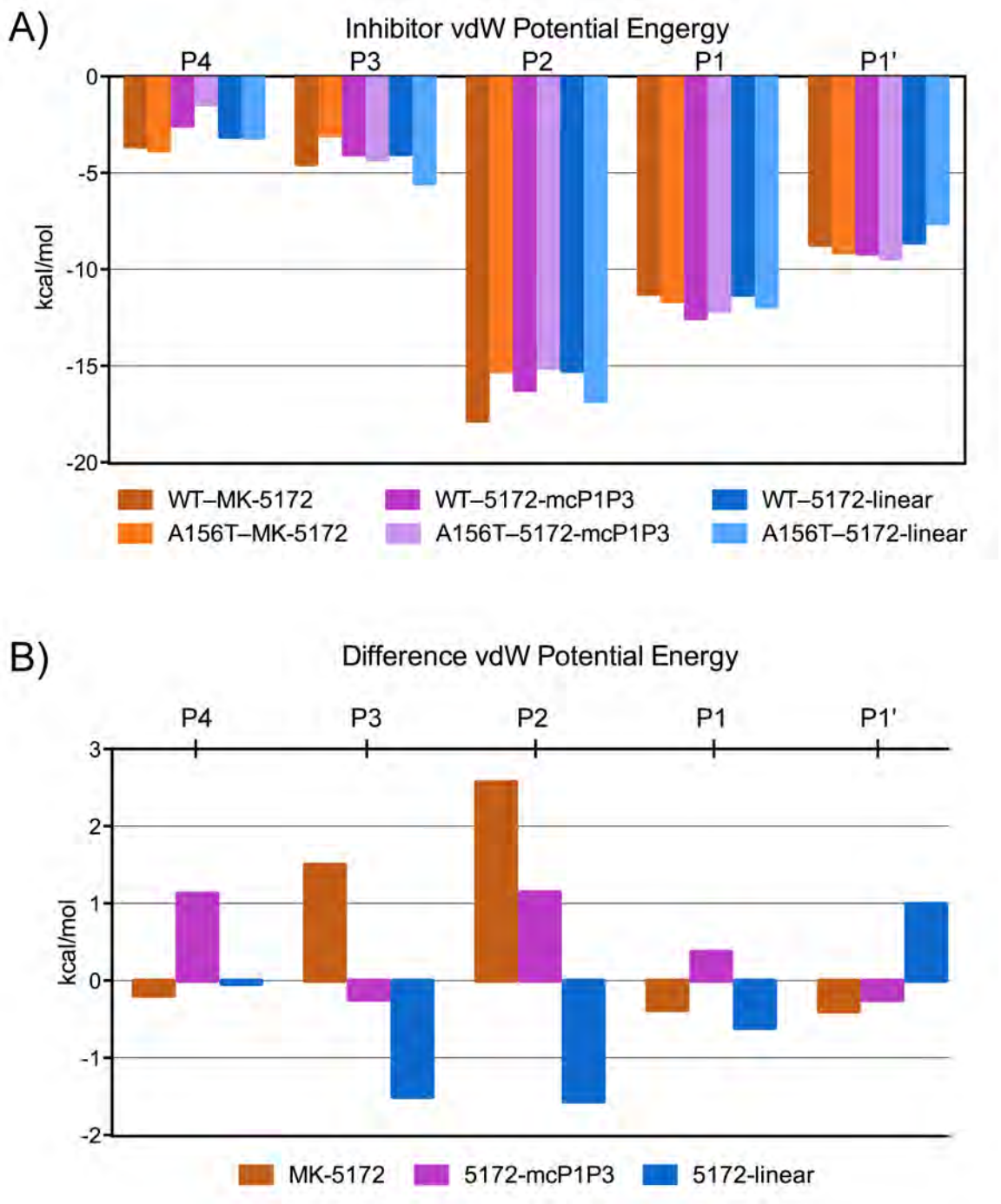


Figure 3.4: Inhibitor packing at the active site in HCV NS3/4A protease crystal structures (**A**) Relative van der Waals interactions of $P_{1'} - P_4$ inhibitor moieties (**B**) changes in vdW interactions in A156T relative to WT protease.

3.3.3 Active site hydrogen bonding pattern is dependent on the macrocyclization status and gets disrupted with A156T mutation

Unlike in the apo state or when binding weak ligands, high-affinity substrates and inhibitors stabilize an extensive electrostatic network at the HCV protease active site, involving sidechains of protease residues D81, R155, D168, and R123 [68, 85]. Within this network, R155–D168 interaction is particularly key for maintaining an electrostatic surface that accommodates high-affinity ligand binding [106]. This interaction is present in all WT complex crystal structures with two hydrogen bonds (H-bonds) between the sidechains of R155 and D168 (Fig. 3.5 and 3.3). In the A156T–MK-5172 protease complex, however, R155 loses both these H-bonds and D168 instead interacts with R123. A similar rearrangement of sidechains underlies the resistance mechanism of asunaprevir to R155K mutation [106]. Mutation of A156, while not directly part of this electrostatic network, still results in a similar disruption of the active site binding surface in the MK-5172 complex, correlating with MK-5172's susceptibility to A156T.

To assess the stability of the active site hydrogen bonding, we calculated the average time H-bonds were observed during the molecular dynamics (MD) simulations. Although present in the MK-5172 crystal structure, the H-bonds between D168–R155 does not persist during the MD simulations, suggesting that this interaction may not be critical to high-affinity inhibitor binding. In contrast, the sidechains of R155 and D168 form highly stable H-bonds in all WT complexes (present more than 89% of the time), consistent with its key role in optimal active site electrostatics. The MD simulations also revealed the persistence of protease–inhibitor H-bonds over time, elucidating the relative bond stability and overall contribution to intermolecular interactions. Most intermolecular H-bonds are at the $\mathrm{P}_1 - \mathrm{P}_{1'}$ moieties (Fig. 3.5 and 3.2), where the inhibitors also have strong vdW interactions with the active site (Fig. 3.2A). Although the hydrogen-bonding patterns of the three

analogs with the active site residues are very similar in the complex crystal structures (Table 3.3), the stability of specific H-bonds varies (Fig. 3.5). MK-5172s P_1 amide N forms a H-bonds with R155 (83% time). This interaction is lost in 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ but present in 5172-linear (68% time). At the $P_{1'}$ moiety, MK-5172 forms four H-bonds with H57, S139 and G137 (77%, 63%, 38% and 52%, respectively). 5172-linear forms the same four H-bonds as MK-5172 but at a higher frequency (50%, 87%, 90% and 86%). This improved stability of H-bonds in 5172-linear correlates with the more favorable enthalpy of binding seen in 5172-linear relative to both macrocyclic analogs (Fig. 3.6A).

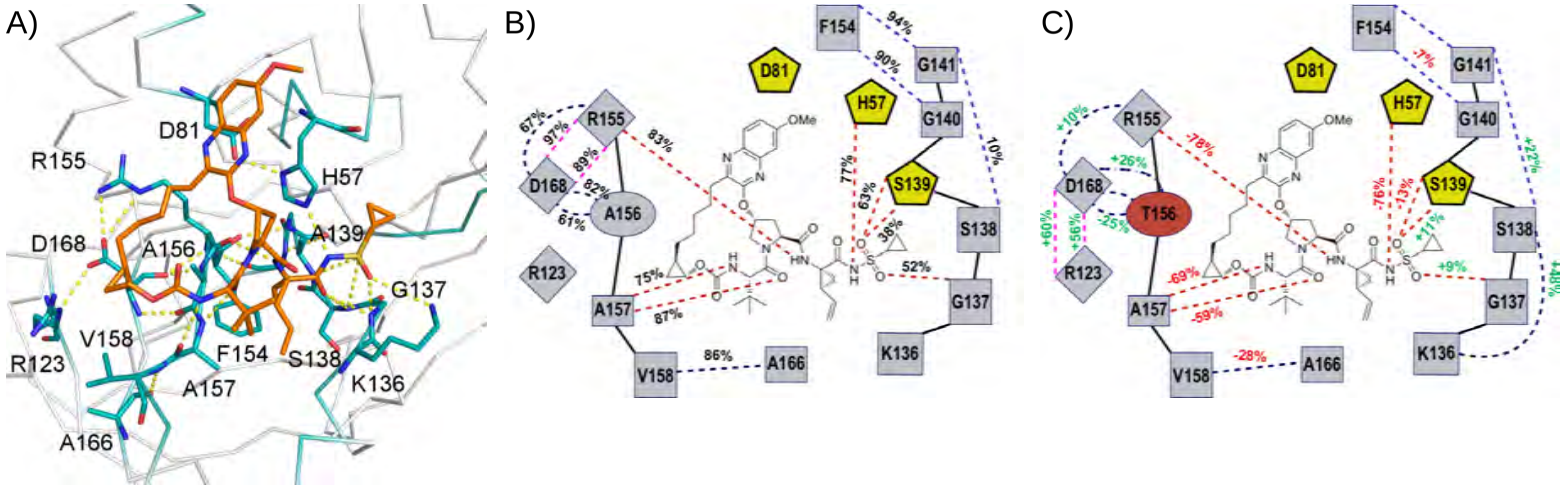


Figure 3.5: Changes in hydrogen bonding patterns in MK-5172 across 100 ns MD simulations.
(**A**) 3D Schematic representation of inter- and intra-molecular (yellow dashed lines) hydrogen bonding interactions in WT–MK-5172 static structure; and 2D representation of hydrogen bonding fingerprint during MD simulation with the percentage of time the bond is present shown in the WT–MK-5172 (**B**), with the percent changes shown in the A156T–MK-5172 (**C**).

The presence of the A156T drug resistance mutation considerably alters the network of hydrogen bonding patterns at the active site. Just as in the crystal structure, during the MD simulations the A156T–MK-5172 complex loses R155–D168 interactions for the less stable D168–R123 H-bonds (97% and 89% versus 60% and 56%, respectively) (Fig. 3.5C). In contrast to the parent compound, both 5172-mc$\mathrm{P_1 - P_3}$ and 5172-linear analogs maintain the key R155–D168 H-bonds (Fig. 3.5A), which partially correlates with the lower fold-loss in potency due to A156T mutation.

Compared to the protease active site changes, loses in intermolecular hydrogen bonding in the mutant complexes are more pronounced across all three inhibitors. MK-5172 loses all H-bonds at $\mathrm{P_3}$ and $\mathrm{P_1}$, and has a weaker $\mathrm{P_{1'}}$ amide N–H57 interaction. 5172-linear experiences the worst drops in H-bonds stability, affecting all of its intermolecular interactions. 5172-mc$\mathrm{P_1 - P_3}$, while experiencing decreases in H-bonds stability throughout the inhibitor, is able to maintain WT-like interactions. The A156T drug resistance mutation therefore has a disruptive effect on both inter and intramolecular hydrogen bonding, however, the inhibitors macrocyclization status can differentially attenuate this effect. Overall, the $\mathrm{P_1 - P_3}$ analog is able to maintain a more WT-like hydrogen-bonding pattern.

3.3.4 Enthalpic loss due to A156T underlies differences in inhibitor susceptibility

The thermodynamics of protease binding for MK-5172 and its analogs were assessed to interrogate the underlying energetic effects of macrocyclization on inhibitor potency and susceptibility to A156T mutation (Fig. 3.6A and Table 3.4) using isothermal titration calorimetry. Although a weaker binder, interestingly the linear analog had more favorable enthalpic interactions. The macrocyclic analogs MK-5172 and 5172-mc$\mathrm{P_1 - P_3}$ have a more favorable (negative) entropic contribution than 5172-linear (-TΔS of -6.6 and -6.0, respectively compared to -3.1 kcal/mol) in binding WT protease, presumably due to the lack of entropic penalty from the rigidity of the macrocyclization. This enhancement in entropy more than compensates for the decrease in the enthalpy of binding, underlying the increased potency of macrocyclic inhibitors compared to their linear counterparts.

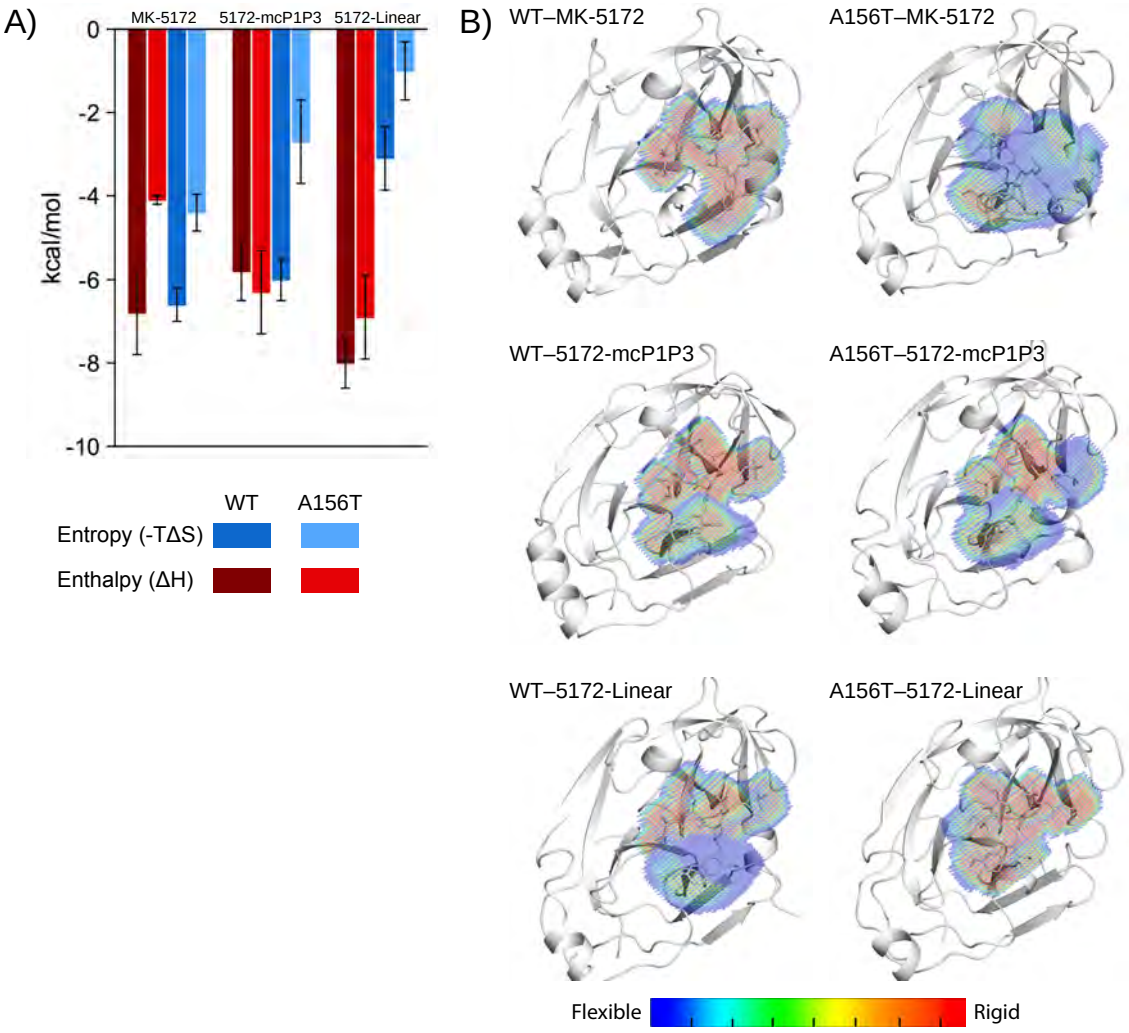


Figure 3.6: Energetic and dynamic effects of macrocyclization on inhibitor binding. (**A**) The entropy and enthalpy of binding of MK-5172 and analogs to WT and A156T protease variants. (**B**) The dynamic inhibitor envelope when bound to WT and A156T protease. Red and blue indicate less and more flexible regions of the inhibitor, respectively.

All three inhibitors lose considerable potency in the presence of the A156T mutation compared to binding the WT protease. The entropic loss due to this mutation is similar for all inhibitors (2.2, 3.3 and 2.3 kcal/mol for MK-5172, 5172-mc$\mathrm{P_1 - P_3}$ and 5172-linear, respectively Table 3.4), suggesting the loss may be partially related to the greater loss in the degrees of freedom of the larger Thr sidechain compared to Ala. Unlike entropy, the enthalpic changes vary greatly among the three inhibitors (2.7, -0.5 and 1.1 kcal/mol for MK-5172, 5172-mc$\mathrm{P_1 - P_3}$, and 5172-linear, respectively Table 3.4) and largely correlate with the changes in inhibitor packing presented above, and susceptibility to A156T. Although 5172-mc$\mathrm{P_1 - P_3}$ loses similar amount of entropy due to the A156T mutation, unlike the parent MK-5172, the enthalpic contribution to binding actually is better (-5.8 and -6.3 kcal/mol, respectively for binding WT vs. A156T protease). 5172-mc$\mathrm{P_1 - P_3}$ can better accommodate the larger Thr sidechain to enhance inhibitor packing at the active site, which results in maintaining the favorable binding enthalpy and hence potency against A156T. Thus binding thermodynamics is consistent with the steric clash of A156T with MK-5172 causing the greatest loss of affinity.

3.3.5 Enzyme and inhibitor conformational dynamics correlate with binding thermodynamics

To further access the impact of A156T mutation on inhibitor binding and reveal any dynamic changes in the complexes underlying potency loss, we carried out fully solvated and extensive MD simulations of all six complexes starting from each of the crystal structures we determined.

The dynamics of inhibitors bound to the protease were assessed by the dynamic inhibitor envelope [107], which can be visualized as a probability distribution for inhibitor occupancy at the active site during MD trajectories (Fig. 3.6B). When the inhibitor is

rigid the occupancy is high (red), while more flexible moieties sample different conformations resulting in low occupancy values (blue). The conformational flexibility of the bound inhibitor can affect both the entropy and enthalpy, and we find that the changes in flexibility mostly correlate with the enthalpy of binding. Comparing the two macrocyclic inhibitors binding to the WT protease, MK-5172 is more rigid and maintains more stable inter-molecular interactions, which correlates with its more favorable enthalpy compared to 5172-mc$\mathrm{P_1 - P_3}$. In the presence of the A156T mutation, these interactions are destabilized resulting in a less occupancy (more blue) dynamic inhibitor envelope, and worse enthalpy of binding for MK-5172 (-6.8 versus -4.1 kcal/mol). This destabilization is less pronounced for the $\mathrm{P_1 - P_3}$ analog (more rigid (or red) in (Fig. 3.6B), which maintains its enthalpy of binding. As expected, the linear analog is more flexible even in the bound state in the WT complex compared to the macrocyclic analogs. However, the linear analog pays a higher entropic penalty upon binding due to the lack of a macrocycle constraining conformational flexibility in the free state (as revealed by the smaller magnitude of favorable entropy than those of the macrocyclic compounds). However, although not favorable for entropic considerations, the lack of a macrocyclic constraint allows 5172-linear to adapt to the mutated surface for optimal packing. This adaptation likely results in the largely maintained favorable enthalpy of binding (6.9 kcal/mol) in the presence of A156T mutation. Indeed, MD simulations revealed that the linear analog adapts to the A156T mutation by sampling a substantially different conformation than that in the WT complex or the starting crystallographic binding pose (Fig. 3.11). By sampling this alternate conformation, 5172-linear sacrifices entropically favorable flexibility of the $\mathrm{P_2}$ group in the bound state, resulting in a high-occupancy (red) dynamic inhibitor envelope (Fig. 3.6B). Hence, alterations in conformational flexibility and related changes in the entropy and enthalpy of binding due to A156T depend on the macrocyclization status of the analogs.

In addition to inhibitor flexibility, the backbone dynamics of the protein were assessed

by Cα root mean square fluctuations (RMSF) during the MD simulations (Fig. 3.7 and 3.10). Nine highly flexible structural regions were identified: (1) the N-terminal NS4A linker, (2) α0 helix spanning residues 12–15, the loops (3), (5), (7) and (8) spanning residues 37–43, 67–72, 88–93 and 157–163 respectively; (4) α1 helix spanning residues 56–61 and containing the catalytic His; (6) F1 β strand spanning residues 76–82 and containing the catalytic Asp; and (9) the C terminal α3 helix spanning residues 173–179. These flexible regions mainly correspond to turns/loops or terminal portions of secondary structural elements throughout the protein structure. The WT protease displays similar backbone dynamics with flexible N and C termini and a stable core when bound to MK-5172 and the two analogs. However, the introduction of the A156T mutation alters these dynamic signatures. In MK-5172 (Fig. 3.7 and 3.10A), the A156T variant increases the backbone flexibility at the N-terminus as well as regions 2, 3, 4, 6, 7 and 8. In 5172-mc$\mathrm{P_1-P_3}$ complex (Fig. 3.10B), the A156T mutation has the opposite effect of reducing flexibility, in particular at regions 1 and 3, relative to the WT complex. As the $\mathrm{P_1-P_3}$ macrocycle does not sterically clash with the larger Thr sidechain, the backbone fluctuations of the mutant complex are not considerably affected by this mutation. In 5172-linear (Fig. 3.10C), the A156T mutation perturbs the backbone Cα at the N and C termini as well as at regions 2, 3, 6, 7. Thus, the mutation causes disturbance to the protein backbone dynamics both at and away from the active site in complexes of MK-5172 and the linear analog, while the protease dynamics is more robust against the dynamic effects of the mutation when bound to the $\mathrm{P_1-P_3}$ analog. Overall, changes in both inhibitor and protease dynamics underpin the observed impact of A156T mutation on the thermodynamics of binding and loses in inhibitor affinity.

3.4 Discussion

The high potency of MK-5172 (sub-nM for WT protease) and robustness to R155K and D168A mutations depend on MK-5172s unique binding mode where the P_2 quinoxaline moiety packs against the catalytic H57 and D81. Albeit beneficial for potency, the $P_2 - P_4$ macrocycle can be regarded as MK-5172s Achilles heel as the rigidity of this region sterically clashes with the multidrug-resistant (MDR) mutation A156T resulting in 3800 fold change (FC) in K_i relative to WT. However, whether the macrocycle could be removed without affecting the binding mode was unclear. In this study the structural and thermodynamic analyses were probed with the linear and $P_1 - P_3$ analogs of MK-5172 we previously designed [81], revealing that macrocyclization does not alter the packing of the quinoxaline moiety but influences susceptibility to the A156T mutation.

The HCV NS3/4A protease structure has an extensive electrostatic network at the active site connecting the catalytic H57 and D81 with residues R155, D168 and R123 [58, 68, 80, 94, 108]. Drug resistance mutations such as R155K and D168A are detrimental because they (1) disrupt this electrostatic network reducing the proteins binding ability, and (2) weaken inhibitor binding by destabilizing the interaction of the inhibitors P_2 heterocyclic moiety. PIs with P_2 moieties stacking against the electrostatic network, such as vaniprevir, danoprevir and asunaprevir, experience decreases in inhibitor affinity when this subsite is altered by mutations at R155 and D168. MK-5172, however, has a P_2 moiety that stacks against the catalytic residues, making the inhibitor less susceptible to the canonical resistance mutations. This binding mode is beneficial for the pan-genotypic activity of MK-5172, as the triad is invariant across genotypes. Our crystal structures show that the stacking of the P_2 quinoxaline in MK-5172 is maintained in linear and $P_1 - P_3$ macrocyclic analogs. Thus, P_2 stacking against the catalytic triad is driven primarily by the quinoxaline moiety and does not rely on the restraint exerted by the $P_2 - P_4$ macrocycle

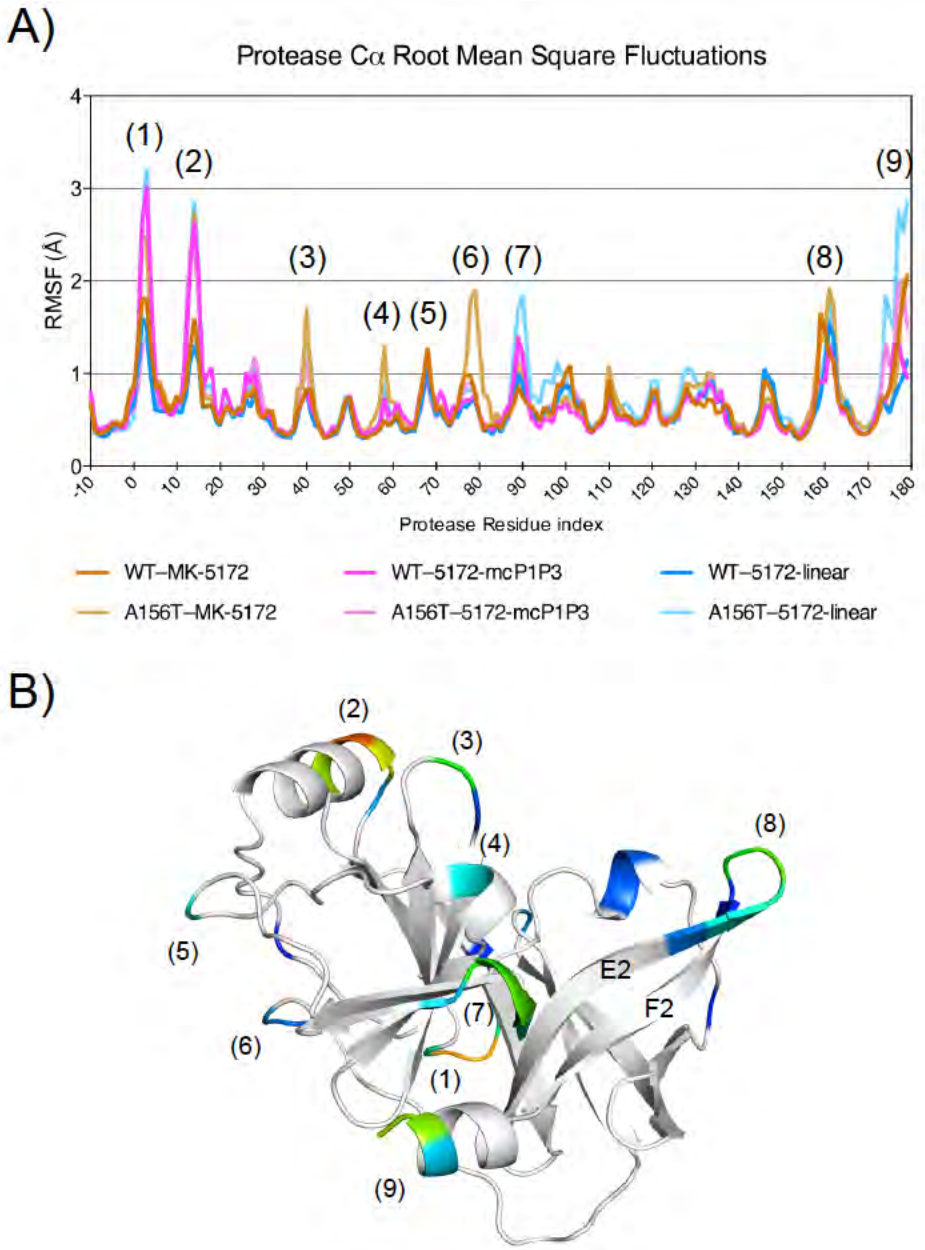


Figure 3.7: Effects of macrocyclization status on protein backbone dynamics
(A) Root square mean fluctuations (RMSF) of backbone Cα during 100 ns MD simulation are shown, with regions in the protease displaying high C rmsf numbered. Residues -10 through 0 represent the NS4A cofactor and linker. **(B)** RMSF mapped onto the NS3/4A structure, with stable core (¡0.7 Å rmsf) in white and regions with peak in rmsf in color. Numbers in parentheses correspond to RMSF peaks in (A).

on the quinoxaline.

Macrocyclization increases inhibitor potency mostly by decreasing the entropic penalty of binding by constraining the unbound inhibitors degrees of freedom and stabilizing the inhibitor in a binding-competent conformation. Even without any stabilizing macrocycle, the 5172 core, 5172-linear, has a K_i of 9.5 nM against WT protease (Table 3.4). This low nanomolar inhibition constant relies on the highly favorable enthalpy of binding due to the preservation of the parent compounds binding mode and efficient packing of the quinoxaline moiety at the catalytic site. Without the macrocycle though, the linear analog is likely highly flexible in the unbound state. The presence of a macrocycle prearranges the inhibitor in binding competent conformation, favorably contributing to the entropy of binding to further increase the potency to 1.96 nM in 5172-mc$\mathrm{P_1 - P_3}$ and 0.14 nM in MK-5172. In addition to contributing directly to inhibitor–protease hydrophobic interactions, as is the case for the $\mathrm{P_2 - P_4}$ macrocycle in MK-5172, the hydrophobic character of the macrocycle can increase lipophilicity, and improve pharmacokinetic properties such as cellular permeability and stability [104, 109, 110].

Protease inhibitors with a $\mathrm{P_2 - P_4}$ macrocycle, such as MK-5172 and vaniprevir, are susceptible to A156T mutation as the larger Thr sidechain can engage in a steric clash with the macrocycle, displacing the inhibitor in the mutant complex structures [68]. In the A156T–MK-5172 complex, the whole inhibitor including the $\mathrm{P_2 - P_4}$ macrocycle is pushed away from the active site due to A156T. Furthermore, the quinoxaline moiety loses critical stacking interactions with H57 and D81. When the $\mathrm{P_2 - P_4}$ macrocycle is not present, in the A156T–5172-linear and A156T–5172-mc$\mathrm{P_1 - P_3}$ complexes, the inhibitor can conformationally adapt to accommodate the A156T mutation. Thus, the 5172-mc$\mathrm{P_1 - P_3}$ analog has the advantage of macrocyclization while still maintaining critical contacts with the catalytic residues.

Second generation protease inhibitors, such as MK-5172, have broader antiviral activity

and higher barriers to resistance than first generation protease inhibitors [111]. However even these inhibitors are not completely robust against drug resistance. In a recent clinical trial [112], A156T variants with an additional V36M/V or I170V mutation were detected in patients who failed therapy. Although the A156T viral variant is debilitated for replication in vitro [113] coexisting compensatory mutations, as often seen in HIV-1 protease [114], may recover viral fitness and lead to virological failure. Therefore, the development of future NS3/4A PIs needs to encompass optimal strategies to best balance inhibitor potency and barrier to resistance. The 5172-mc$\mathrm{P_1-P_3}$ analog provides opportunities for further optimization of compounds with stable stacking interactions with the catalytic residues, while staying within the substrate envelope. As the unique binding mode of MK-5172 is preserved even without the $\mathrm{P_2-P_4}$ macrocycle, the removal of the macrocyclic linkage at the $\mathrm{P_2}$ quinoxaline enables exploration of the chemical space to modify this group for an even more potent inhibitor robust against resistance.

In rational drug design, the effect of macrocyclization on the entropy and enthalpy of binding is not straightforward to predict, especially in the presence of resistance mutations. While the macrocyclic constraint is beneficial for entropy, as we have demonstrated here for MK-5172, the location of the macrocycle should be designed for optimal enthalpy as well, to preserve the desired binding mode and ideally stay within the substrate envelope [80,85]. The protrusion of the $\mathrm{P_2-P_4}$ macrocycle in MK-5172 outside the substrate envelope effectively surrounding A156 causes vulnerability to mutations at this residue. In the presence of A156T, while the rigidity brought about by the macrocycle is still beneficial for entropy, the same rigidity is detrimental to enthalpy. Design of macrocycles should not only fit within the substrate envelope but also consider both inhibitor rigidity and adaptability to potential resistance mutations, to achieve optimal entropy–enthalpy balance required for robustness to avoid susceptibility to resistant mutations.

3.5 Materials and Methods

3.5.1 Protein Constructs

The HCV genotype 1a NS3/4A protease domain gene [68] was synthesized by GenScript and cloned into the pET28a expression vector (Novagen). The highly soluble single-chain construct consists of NS3/4A protease domain (residues 4–181) fused to a fragment of the cofactor NS4A (residues 12–23) via an SD linkage. A similar protease construct exhibited catalytic activity comparable to that of the authentic full-length protein [115]. All protease variants were generated using the QuikChange Site-Directed Mutagenesis Kit from Stratagene.

3.5.2 Protein Expression and Purification

Protein expression and purification were carried out as previously described [85]. Briefly, transformed BL21(DE3) E. coli cells were grown at $37\,^{\circ}\mathrm{C}$ and induced at an optical density of 0.6 by adding 1 mM Isopropyl β-D-1-thiogalactopyranoside. Cells were harvested after 5 hours of expression, pelleted, and frozen at -$80\,^{\circ}\mathrm{C}$ for storage. Cell pellets were thawed, resuspended in 5 mL/g of resuspension buffer (50 mM phosphate buffer, 500 mM NaCl, 10% glycerol, 2 mM β-ME, pH 7.5) and lysed with a cell disruptor. The soluble fraction was retained, applied to a nickel column (Qiagen), washed with resuspension buffer, and eluted with resuspension buffer supplemented with 200 mM imidazole. The eluent was dialyzed overnight (MWCO 10 kD) to remove the imidazole, and the His-tag was simultaneously removed with thrombin treatment. The nickel-purified protein was then flash frozen and stored at -$80\,^{\circ}\mathrm{C}$.

3.5.3 Crystallization

The above-mentioned protein solution was thawed, concentrated to 3 mg/mL and loaded on a HiLoad Superdex75 16/60 column equilibrated with gel filtration buffer, 25 mM 2-(N-Morpholino)ethanesulfonic acid (MES), 500 mM NaCl, 10% glycerol, 30 mM zinc chloride, and 2 mM 1,4-Dithiothreitol, pH 6.5. The protease fractions were pooled and concentrated to 20-25mg/mL with an Amicon Ultra-15 10 kD device (Millipore). The concentrated samples were incubated for 1 hour with 1-3 molar excess of inhibitor. Diffraction-quality crystals were obtained by mixing equal volume of concentrated protein solution with precipitant solution (20-26% PEG-3350, 0.1 M sodium MES buffer, 4% ammonium sulfate, pH 6.5) in 24-well VDX hanging drop trays.

3.5.4 Data Collection and Structure Solution

X-ray diffraction data were collected either at Advanced Photon Source LS-CAT 21-ID-F, or at our in-house Rigaku Saturn X-ray system. Diffraction intensities were indexed, integrated and scaled using the program HKL2000 [116]. All structure solutions were generated using molecular replacement with PHASER [89]. The B chain model of viral substrate product 4A-4B (3M5M) [80] was used as the starting model for all structure solutions. Initial refinement was carried out in the absence of modeled ligand, which was subsequently built in during later stages of refinement. Subsequent crystallographic refinement was carried out within the Phenix program suite, with iterative rounds of TLS or restrained refinement until convergence was achieved [117]. The protein crystals of the A156T protease in complex with 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ grew as pseudo-merohedral twins. X-Triage was used to determine the twin operator, -k,-h,-l thus refinement was carried out using twin law= -k,-h,-l. This structure has 4 molecules in the asymmetric unit with the active sites of molecules A and C superposed onto molecules B and D respectively. Molecule

B was used in all of the analysis, as it possessed the lowest B factors representing the most stable and consistent conformer. The final structures were evaluated with MolProbity [118] prior to deposition in the Protein Data Bank. To limit the possibility of model bias throughout the refinement process, 5% of the data were reserved for the free R-value calculation [119]. Interactive model building and electron density viewing was carried out using the program COOT [91].

3.5.5 Inhibitor Complex Analysis

Superpositions were performed in PyMOL [120] using the C-alpha atoms of the active site protease residues 137139 and 154160. The A chain of WT–MK-5172 complex was used as the reference structure for each alignment. Two conformations are observed in the A156T–5172-linear structure, and the one with the lowest B factors within the P_{21} unit cell was used in the structural analysis. The van der Waals contact energies between protease residues and inhibitors were computed using a simplified LennardJones potential as described previously [121].

3.5.6 Molecular Dynamics Simulations

Molecular dynamics simulations were carried out, in triplicate, following previously published protocols [94] using Desmond [122, 123] with the OPLS2005 force field [124, 125]. After equilibration, each trajectory was run for 100 ns at $300\,^{\circ}\mathrm{K}$ and the coordinates recorded every 5 ps. The dynamic inhibitor envelope was calculated using the van der Waals volume of inhibitor conformers from the MD trajectories mapped onto the three-dimensional grid placed on the binding site of the enzyme and normalized by the total number of conformers to obtain a probability distribution, as detailed previously [94, 126].

3.5.7 Isothermal Titration Calorimetry

ITC experiments were performed as previously described [127, 128]. Briefly, NS3/4A protease was purified via gel filtration in 50 mM HEPES, 300 mM NaCl, 10% glycerol, 1 mM tris(2-carboxyethyl)phosphine at pH 7.4 and concentrated to 10–60 µM. Inhibitor stock was prepared in 100% Dimethyl sulfoxide and diluted to working conditions in the gel filtration buffer with final Dimethyl sulfoxide concentration not exceeding 3%. Experiments were performed with a highly sensitive Microcal ITC200 (Malvern) at $25\,^{\circ}\mathrm{C}$ with the protein in the sample cell and inhibitor solution in the syringe. The data was analyzed using Origin 7.0 and the change in enthalpy (ΔH) and corresponding dissociation constant (K_d) were determined via non-linear regression, with a one-binding site model. Due to the potency of MK-5172, the inhibition constant from enzyme kinetics assay (K_i) of WT–MK-5172 from our previous work [81] was used to calculate the Gibbs Free Energy (ΔG), as the K_d is too low to be reliably determined by ITC without a competing ligand [129] which we were unable to successfully attain. All other ΔG calculations used ITC derived K_d values.

3.6 Supplemental Information

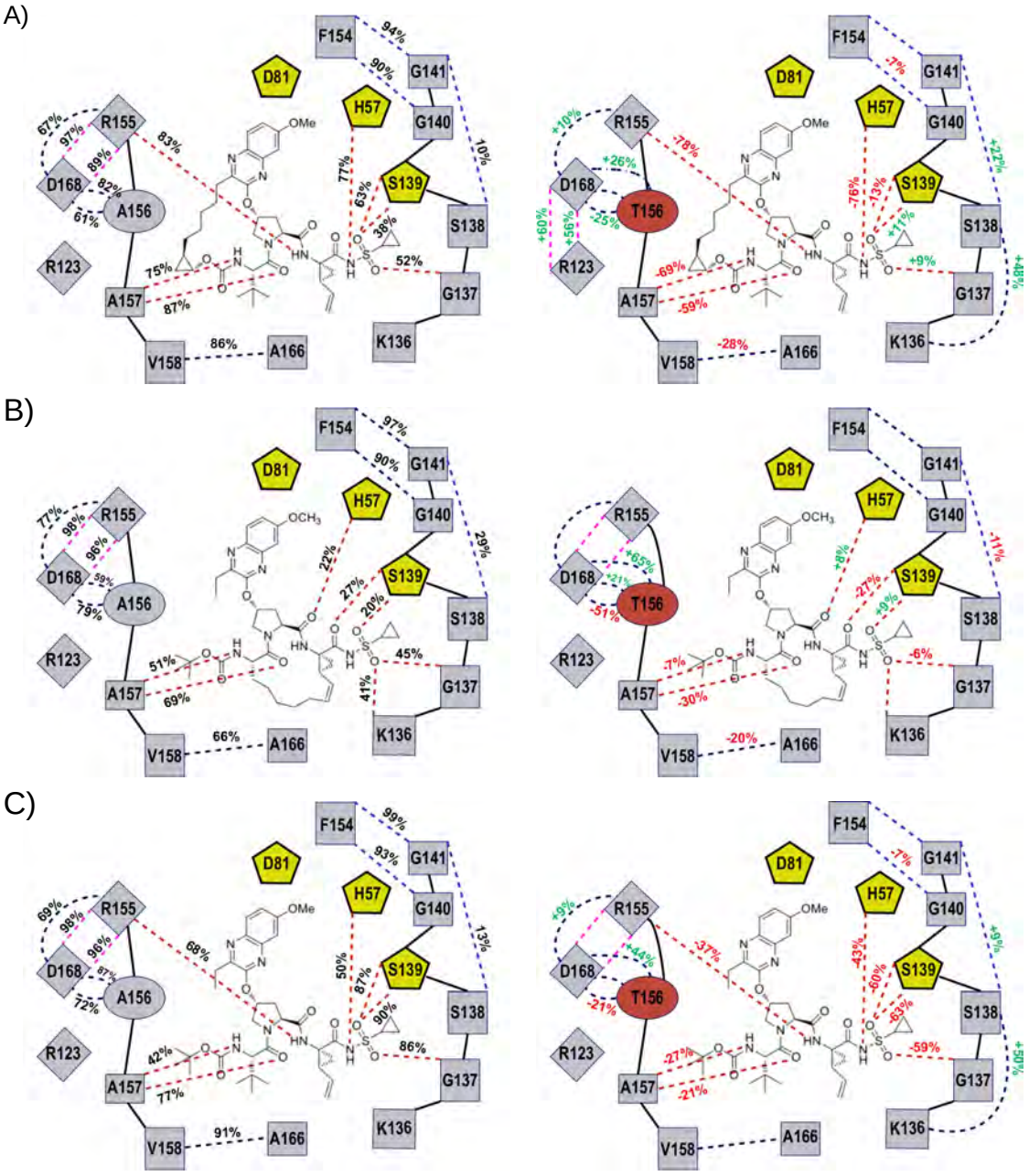


Figure 3.8: Changes in hydrogen bonding patterns across MD simulations. Schematic representation of inter-molecular (blue dashed lines) and intra-molecular (red dashed lines) hydrogen bonding fingerprints. (**A**) Hydrogen bonding interactions in MK-5172, (**B**) 5172-mc$P_1 - P_3$ and (**C**) 5172-linear. The percentage of time the bond is present during MD simulations are listed in the WT complexes (left), with the percent changes greater than 5% shown in the A156T complexes (right).

		MK-5172		5721-mcP1P3		5172-linear	
Protease	Inhibitor	WT	A156T	WT	A156T	WT	A156T
S139 HG	P1' SO1	1.7 (63%)	1.9 (49%)	2.1 (<20%)	>3.5 (<20%)	1.7 (87%)	2.0 (27%)
S139 H	P1' SO1	2.8 (38%)	2.9 (49%)	3.0 (20%)	2.8 (29%)	2.8 (90%)	2.8 (27%)
K136 HZ1	P1' SO2	2.6 (<20%)	>3.5 (<20%)	>3.5 (41%)	>3.5 (41%)	>3.5 (<20%)	>3.5 (<20%)
G137 H	P1' SO2	2.1 (52%)	2.3 (61%)	2.4 (45%)	2.2 (40%)	2.3 (86%)	2.1 (27%)
H57 NH	P1' N	2.0 (77%)	1.9 (1%)	2.3 (<20%)	2.2 (<20%)	2.1 (50%)	2.1 (7%)
S139 H	P1 CO	2.7 (<20%)	2.4 (<20%)	2.6 (27%)	2.5 (<20%)	2.5 (<20%)	2.2 (<20%)
S138 H	P1 CO	2.7 (<20%)	>3.5 (<20%)	2.7 (<20%)	2.6 (<20%)	2.7 (<20%)	2.8 (<20%)
G137 H	P1 CO	2.1 (<20%)	2.3 (<20%)	2.0 (<20%)	2.1 (<20%)	2.1 (<20%)	2.4 (<20%)
R155 CO	P1 NH	2.1 (83%)	2.0 (5%)	2.0 (<20%)	1.9 (<20%)	1.9 (68%)	1.9 (41%)
A157 NH	P3 CO	2.3 (87%)	2.5 (29%)	1.8 (69%)	1.9 (39%)	2.0 (77%)	1.7 (56%)
A157 CO	P3 NH	2.2 (75%)	2.1 (5%)	2.0 (51%)	1.9 (43%)	1.9 (42%)	1.7 (5%)

Table 3.2: Intermolecular hydrogen bonding of MK-5172 and analog complexes.
The bond lengths in crystal structures (in angstroms) and the percentage of time the bond is present during MD simulations (in parentheses) are listed.

		MK-5172		5172-mcP1P3		5172-linear	
		WT	A156T	WT	A156T	WT	A156T
PHE154-Main-N	GLY141-Main-O	1.7 (94%)	1.7 (96%)	1.7 (97%)	1.7 (97%)	1.8 (99%)	1.8 (97%)
GLY141-Main-N	SER138-Main-O	2.2 (10%)	2.0 (32%)	2.1 (29%)	2.1 (18%)	>3.5 (13%)	2.3 (23%)
GLY140-Main-N	PHE154-Main-O	2.1 (90%)	1.9 (84%)	1.8 (90%)	2.0 (93%)	1.9 (93%)	2.0 (86%)
VAL158-Main-N	ALA166-Main-O	1.8 (86%)	2.0 (58%)	1.9 (66%)	1.9 (46%)	1.9 (91%)	2.1 (88%)
ASP168-Main-N	ALA156-Main-O	1.5 (82%)	1.7 (86%)	1.7 (59%)	1.8 (80%)	1.8 (87%)	2.1 (82%)
ARG155-Main-N	ASP168-Main-O	1.8 (67%)	1.9 (77%)	2.0 (77%)	2.0 (81%)	2.0 (69%)	1.9 (77%)
A/T156-Main-N	ASP168-Main-O	2.0 (61%)	2.1 (36%)	2.0 (79%)	2.1 (28%)	2.1 (72%)	2.4 (51%)
SER138-Side-OG	LYS136-Main-O	>3.5 (<20%)	3.5 (48%)	>3.5 (<20%)	>3.5 (<20%)	>3.5 (<20%)	>3.5 (50%)
ARG123-Side-NE	ASP168-Side-OD	1.8 (<20%)	1.5 (60%)	>3.5(<20%)	>3.5 (<20%)	3.1 (<20%)	>3.5 (<20%)
ARG123-Side-NH2	ASP168-Side-OD	2.8 (<20%)	2.5 (56%)	>3.5 (<20%)	>3.5 (<20%)	>3.5 (<20%)	>3.5 (<20%)
ARG155-Side-NE	ASP168-Side-OD	2.5 (97%)	>3.5 (<20%)	1.5 (98%)	>3.5 (97%)	2.0 (98%)	1.5 (99%)
ARG155-Side-NH1	ASP168-Side-OD	1.9 (89%)	>3.5 (<20%)	2.1 (96%)	3.3 (95%)	1.9 (95%)	2.3 (96%)
THR156-Side-OG1	ASP168-Side-OD	>3.5 (<20%)	>3.5 (26%)	>3.5(<20%)	>3.5 (65%)	>3.5 (<20%)	>3.5 (44%)

Table 3.3: Intramolecular hydrogen bonding patterns of MK-5172 and analogs.
Intramolecular hydrogen bonds at the active site of HCV NS3/4A protease forming an electrostatic network in the crystal structures of inhibitor complexes. The bond lengths in crystal structures (in angstroms) and the percentage of time the bond is present during MD simulations (in parentheses) are listed.

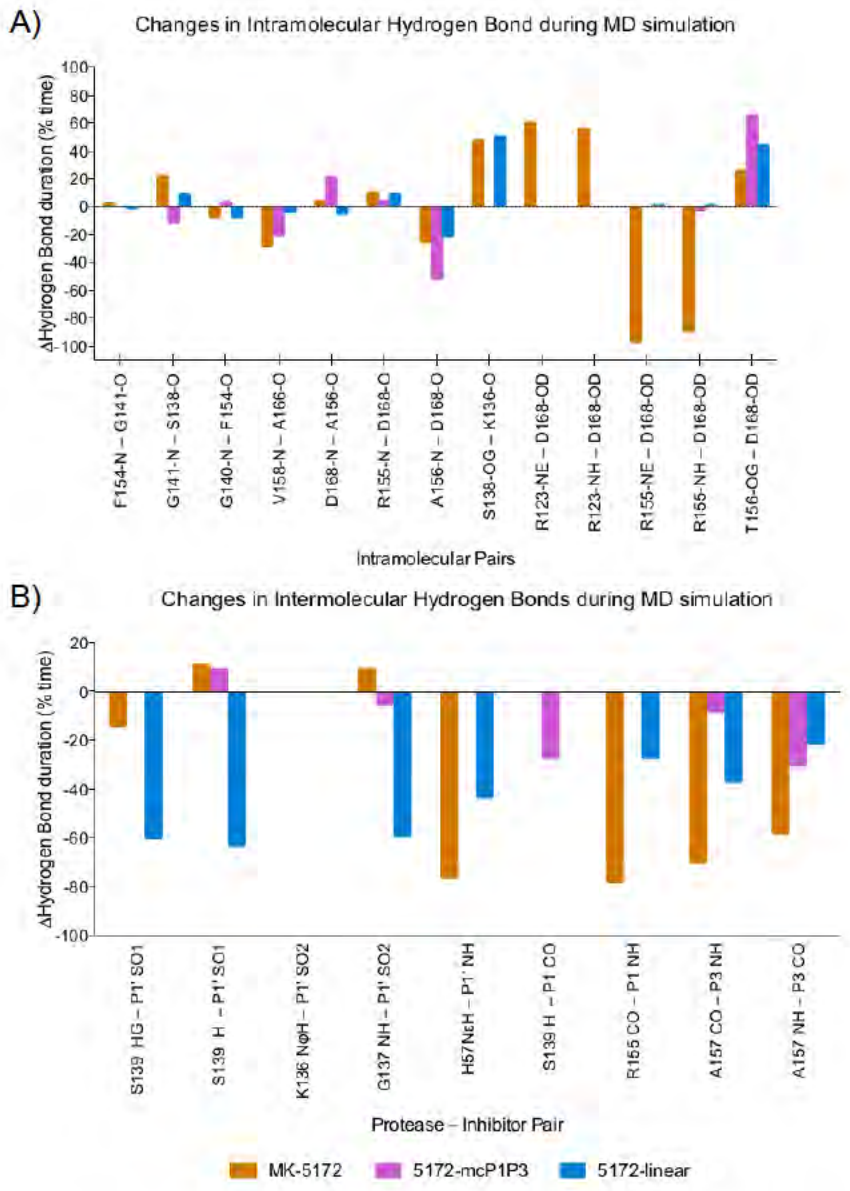


Figure 3.9: Hydrogen bonding changes during MD simulations.
Changes in intra- (**A**) and inter-molecular **B**) hydrogen bonding interactions in MK-5172 (orange), 5172-mc$P_1 - P_3$ (magenta) and 5172-Linear (blue) during 100ns MD simulations.

Protease–Inhibitor variant	Replicon IC50 (nM)	K_i (nM)	K_d (nM)	ΔH (kcal/mol)	ΔTΔS (kcal/mol)		ΔG (kcal/mol)	
					K_i	K_d	K_i	K_d
WT–MK-5172	0.29	0.14 (0.02)	1.5 (0.4)	-6.8 (1)	-6.6	N/A	-13.4	N/A
WT–5172mcP1P3	0.26	1.96 (0.13)	2.2 (0.56)	-5.8 (0.7)	-6.0	-4.6 (0.5)	-11.8	-10.4 (0.2)
WT–5172-Linear	4.74	9.5 (0.75)	6.7 (2.34)	_8.0 (0.58)	-2.9	_3.1 (0.76)	-10.9	_11.1 (0.2)
A156T–MK-5172	46.6	537 (74)	359.7 (0.09)	-4.1 (0.082)	-4.4	-4.5 (0.44)	-8.5	-8.8 (0.36)
A156T–5172mcP1P3	3.96	250 (29)	39.9 (5)	_6.3 (1)	-2.7	_3.9 (1.6)	-9	_10.2 (0.54)
A156T–5172-linear	23.1	1708 (230)	154.3 (4)	_6.9 (1)	-0.92	_2.4 (1.2)	-7.8	_9.3 (0.16)

Table 3.4: Binding thermodynamics of Mk-5172 and analog complexes
Actual dissociation constant (K_d) of WT–MK-5172 is below the sensitivity of ITC via direct binding experiments. The inhibition constant from enzyme kinetics assay 1 (K_i) was used as the K_d for Gibbs free energy (ΔG) calculation of MK-5172. ΔG and TΔS values are given with both K_i and K_d for comparison when appropriate. Values in parentheses () represent the standard deviation.

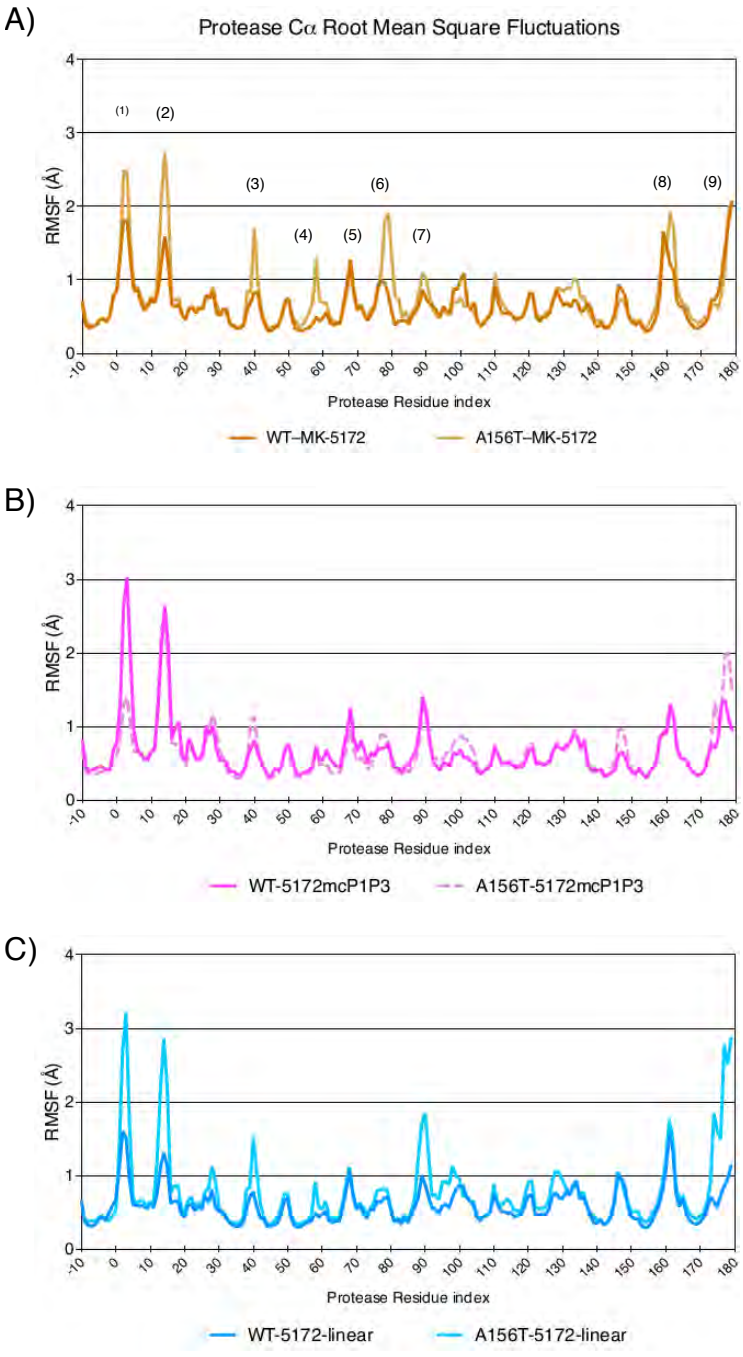


Figure 3.10: NS3/4A backbone dynamics during 100 ns MD simulations. Cα root mean square fluctuations (RMSF) of MK-5172 (**A**), 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ (**B**) and 5172-linear (**C**). Regions in the protease displaying high C-alpha rmsf are numbered.

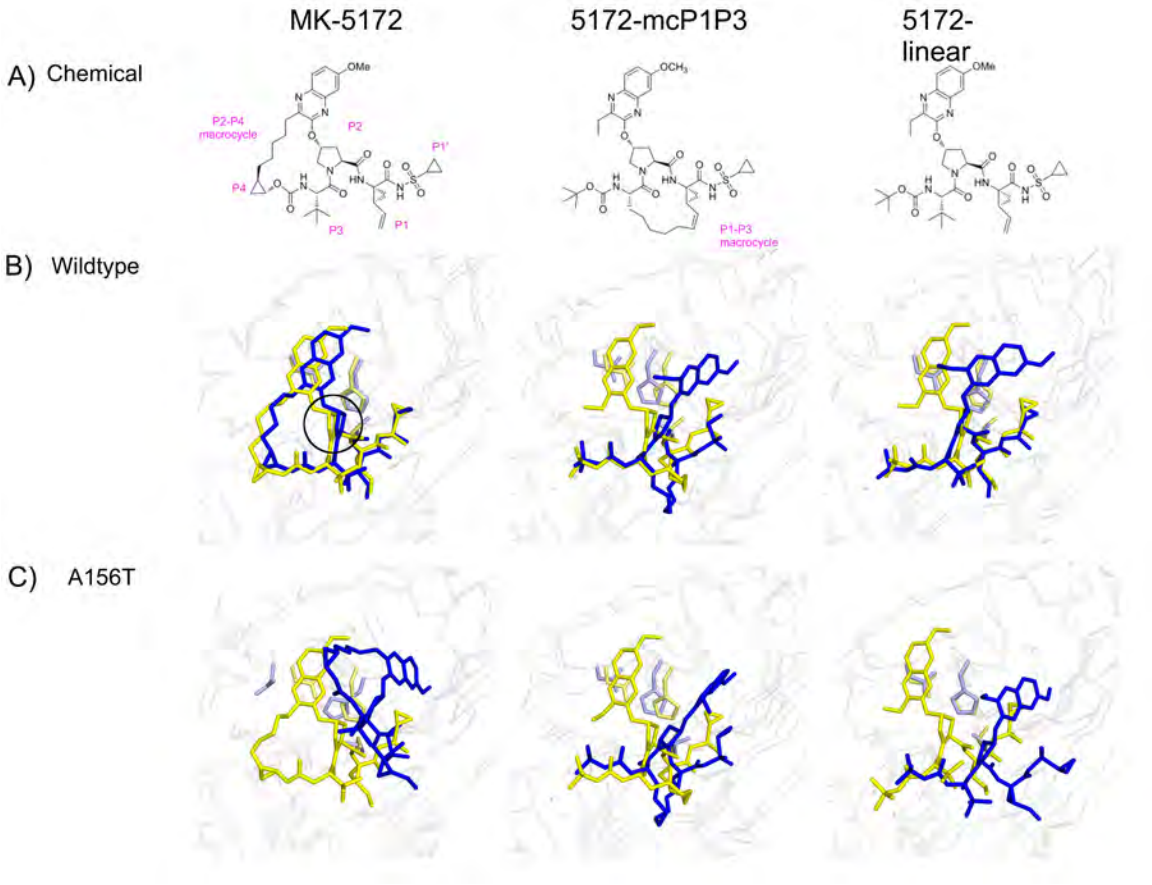


Figure 3.11: Snapshots of NS3/4A inhibitor complexes during 100 ns MD simulations. The Structural Alignments of NS3/4A complexes (white ribbons), the first (yellow) and last (blue) frames of MD simulations, illustrate the various inhibitor conformations. In WT complexes (**B**), both aligned frames superpose as judged by the stacking of the P_2 prolines (red circle). In A156T variant complexes (**C**), MK-5172 (bottom left) displays the highest degree of inhibitor misalignment as its P_2 proline rotates 90 degrees towards the $\mathrm{S}_{1'}$ binding site. 5172-mc$\mathrm{P}_1 - \mathrm{P}_3$ (bottom center) is able to maintain a similar conformation as in the WT complex (top center). The P_2 proline in A156T–5172-Linear (bottom right) shifts 4 Å away from the active site and gets locked in an unfavorable conformation.

PREFACE TO : CHAPTER IV

This chapter is being formatted for submission.

Soumana, D., Yilmaz N. K., Prachanronarong, K. P., Ali, A., and Schiffer, C. A. (2015). **Elucidating the Molecular Basis for HCV Protease Inhibitor Failure in Genotype 3**

Author contributions: DS and CAS designed the study. AA made the inhibitors. **DS** performed all the experiments (including MD simulations), analyses and figure constructions. KP provided technical assistance. **DS**, NYK and CAS interpreted the data and wrote the manuscript

chapter IV

Elucidating the Molecular Basis for HCV Protease Inhibitor Failure in Genotype 3

Chapter IV

Elucidating the Molecular Basis for HCV Protease Inhibitor Failure in Genotype 3

4.1 Abstract

Hepatitis C virus, affecting an estimated 150 million people worldwide, is the leading cause of viral hepatitis, cirrhosis and hepatocellular carcinoma. The virus exhibits high genetic diversity with over six genotypes and multiple subtypes. Genotypes 1 and 3, the most prevalent viral species, have different geographic distributions. Recent advances in direct-acting antivirals against Hepatitis C virus have led to the development of potent inhibitors of viral nonstructural proteins, including protease inhibitors several of which are approved by regulatory agencies. Protease inhibitors are potent antivirals against genotype 1, but not genotype 3. Lack of structural information and difficulties developing a genotype 3 subreplicon model have limited our ability to understand this differential efficacy of inhibitors. Here, we present structural and biophysical data elucidating the molecular basis for genotype 3 NS3/4A escape from current protease inhibitors. We determined high-resolution crystal structures of asunaprevir, danoprevir and vaniprevir in complex with a genotype 1a-3a chimeric NS3/4A protease and compared these structures and their molecular dynamics to those of native genotype 1a. We found that genotype 1 and 3 active sites only differ by three amino acids (Arg/Thr123, Ile/leu132 and Asp/Gln168); and when these polymorphisms are incorporated into a genotype 1 backbone, the resulting chimeric protease exhibits a similar response to inhibitors as the genotype 3 protease. This recapitulation occurs via a disruption of the 155–168 electrostatic interaction, key for intermolecular

dependence of quinoline and indoline based inhibitors.

4.2 Introduction

Hepatitis C virus (HCV) causes chronic liver infection that affects about 3% of the global population and is the main cause of hepatitis, cirrhosis, and hepatocellular carcinoma [69, 70, 130]. HCV is a highly evolved, highly diverse virus with six known genotypes (GT) and multiple subtypes [2, 3]. The two most prevalent viral species, genotype 1 (GT-1) and 3 (GT-3) account for roughly 46% and 30% of the global infections respectively [2, 4] and have a different geographic distribution. While GT-1 is most endemic in Northern America and Europe, GT-3 is most prevalent in South Asia. The remaining GTs 4, 5 and 6 span the Middle East to Central Africa, South Africa, and Southeast Asia [2, 4].

Before the recent availability of direct-acting antivirals (DAAs), the standard of care consisted of pegylated-interferonα (Peg-INF) and ribavirin (RBV), however, this treatment had moderate to low rates of cure across genotypes and was poorly tolerated [96, 130]. Current therapeutic efforts aim to determine the best-in-class DAAs that target several viral proteins including the viral entry protein, the NS3/4A protease, the NS5A and NS5B non structural proteins [97] and host microRNAs [98] either individually or in combination. Sofosbuvir (SOF), the HCV blockbuster drug, is a prodrug nucleotide analog targeting the highly conserved palm region of the viral NS5B RNA dependent RNA polymerase [97, 131]. In a combination therapy with or without Peg-INF/RBV, more than 90% of GT-1 patients who received the SOF-based therapy experienced undetectable viremia, sustained virological response (SVR24) for 24 weeks post treatment follow-up [132]. Thanks to its high resistance barrier [133] and potency [131], SOF is progressively becoming the backbone of the standard of care as it is currently approved for use in combination with other DAAs, ledipasvir [134], simeprevir [112], and daclatasvir [135], resulting in greater

than 95% cure rates at 12 weeks post treatment follow-up, (SVR12) [71, 136].

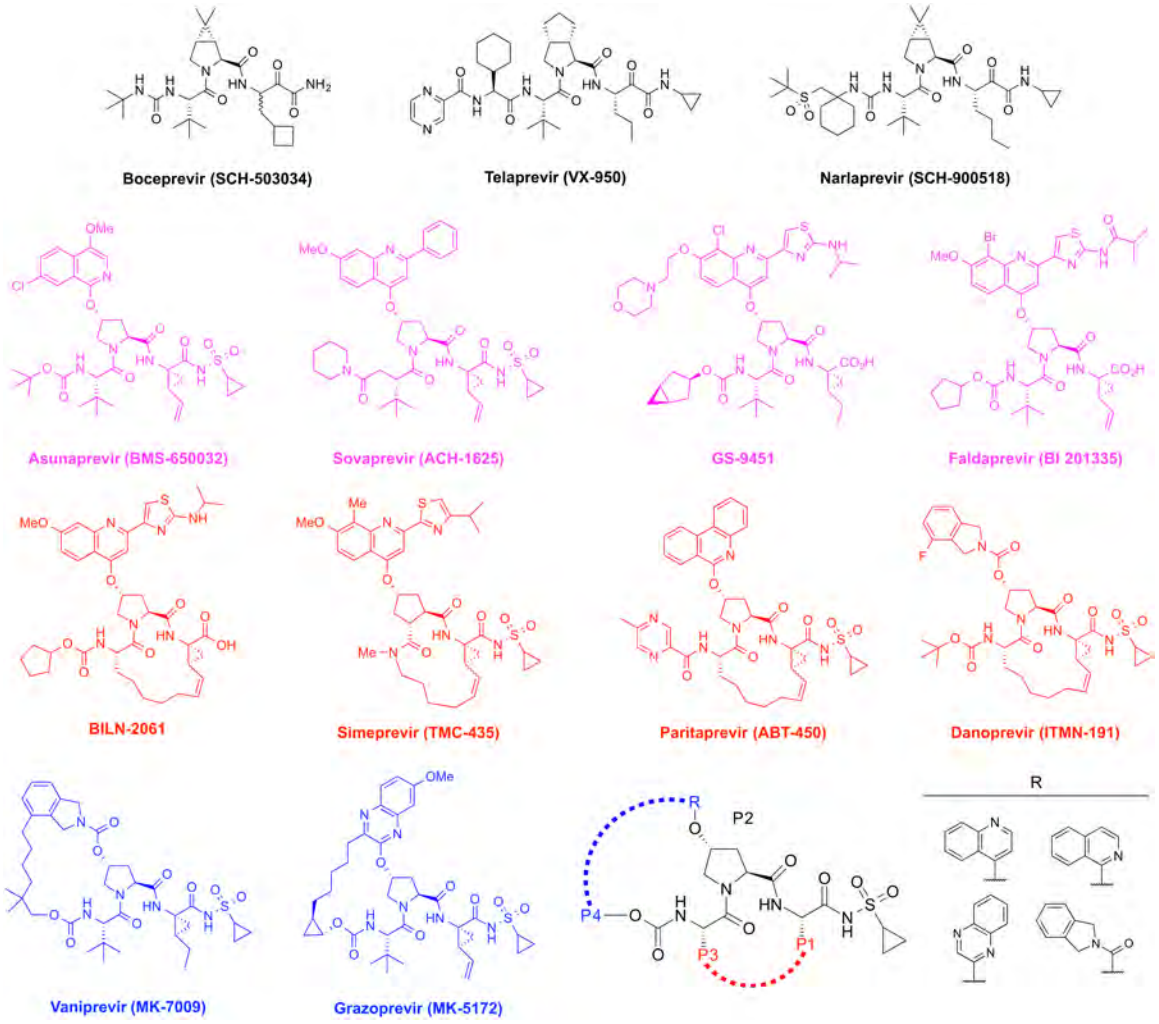


Figure 4.1: Schematic representation of NS3/4A protease inhibitors
The three broad classes of HCV PIs are shown: linear αketoamide covalent (top-black); linear non covalent (second row-pink); and macrocyclic (P_1-P_3 and P_2-P_4 in red and blue respectively.

Four FDA-approved HCV inhibitors (telaprevir [31,32], boceprevir [35], simeprevir [39], and most recently, paritaprevir [40]) target the NS3/4A protease (Fig. 4.1).

The NS3/4A protein is a bifunctional enzyme containing an N-terminal serine protease domain (amino acids 1-180) with the classic catalytic triad (S139, H57, D81) of the chymotrypsin superfamily, and a C-terminal DExH/D-box helicase of superfamily II with NTPase activity [99, 100, 102]. The NS3/4A protease is responsible for cleaving the viral polyprotein and host factors involved in the innate immune response, including TRIF and mitochondrial antiviral signaling (MAVS) [26–28]. Serving as a two-pronged attack on the virus by preventing viral maturation and restoring the immune response, targeting the NS3/4A protease has historically been an industry priority. To date, NS3/4A protease inhibitors (PIs) represent the largest class of DAAs. However, despite serving as the first proof of principle for targeting an HCV viral protein [46, 137] and being extensively studied [111, 138–144], PIs suffer from a low genetic barrier to resistance [145] and lack of cross-genotypic breadth. In fact, all PIs are ineffective against genotype 3 (GT-3) NS3/4A protease.

GT-3 infected patients suffer from the lowest SVR rates. Only 60% GT-3 patients treated with a Sofosbuvir monotherapy achieve SVR24 [146], whereas a combination therapy of ribavirin and Sofosbuvir experienced SVR12 rates of 57% [147]. The recently FDA approved combination of daclatasvir (a NS5A inhibitor) + sofosburvir results in roughly 96% SVR12 rates in GT-3 [148]; however, this therapy is only available in the US and the standard of care for GT-3 patients outside of the U.S is still PEG-INF + RBV.

To elucidate the structural basis for differential efficacy of protease inhibitors in GT-3 relative to GT-1, we engineered a chimeric GT-1a3a (1a3a) protease to map the minimal active site mutations necessary and sufficient for recapitulating GT-3 behavior. We then determined high-resolution inhibitor co-crystal structures of GT-1a and 1a3a in complex with asunaprevir (ASV), danoprevir (DAN) and vaniprevir (VAN). Through a detailed bio-

physical, structural and dynamic analysis of the complexes, we find that, while GT-1 and GT-3 share a 76% sequence identity, their active sites only differ by three polymorphism R123T, I132L and D168Q. Incorporating the GT-3 polymorphism into the GT-1 protease allows a recapitulation of GT-3 activity. The loss of inhibitor efficacy in the 1a3a and GT-3 proteases is a consequence of disrupted electrostatic interactions between amino acids 168 and 155, which is critical for potent binding of quinoline and isoindoline based PIs. We reveal details of molecular and structural basis for the lack of PI efficacy against GT-3, which are needed for design of pan-genotypic inhibitors.

4.3 Results and Discussion

4.3.1 Protease inhibitors are significantly less potent in GT-3 than in GT-1

To assess the efficacy of PIs in GT-1 and 3, we performed a comprehensive analysis of enzyme inhibition constants (K_i values) for 7 PIs (telaprevir, boceprevir, asunaprevir, danoprevir, vaniprevir, grazoprevir and MK-6325) representing the three general classes of inhibitor (Fig. 4.2). Consistent with previous reports, the GT-3 NS3/4A protease is less responsive to all PIs tested [51,149]. Among the first generation PIs, the reversible covalent inhibitors telaprevir and boceprevir have a similar K_i against GT-1 (33 nM); however, their potencies are differentially decreased in GT-3. While telaprevir's potency decreases 8 fold (266 nM), boceprevirs K_i changes only by 2.2 fold (71 nM). Second wave inhibitors (asunaprevir, danoprevir, and vaniprevir) have low to subnanomolar range (2.7, 1.0, 0.7 nM respectively) potencies in GT-1, yet low to submicromolar K_is in GT-3. Relative the GT-1 the linear asunaprevir experiences a 975 fold change (FC) in K_i, while the macrocyclic DAN and VAN experience a 878 and 533 FC respectively.

While the nature of the P_2 moiety enhances potency, macrocyclization of PIs modulates the extent of the potency enhancement [81], through improved entropic binding energetics relative to their linear counterparts. The position of the macrocycle affects potency enhancement as is evident when comparing the profiles of DAN and VAN, two inhibitors containing the same isoindoline P_2, but harboring different macrocyclic strategies. While GT-1 WT protease is insensitive to the macrocyclization position, GT-3 favors the P_2-P_4 macrocycle. Of the panel, next generation PIs, grazoprevir (GRZ) and MK-6325, are the most efficacious inhibitors against GT-1 (33 and 31 nM respectively). They are characterized by (1) a P_2 quinoxaline and (2) a P_2-P_4 or P_2-P_4/P_1-P_3 macrocyclization for GRZ and

MK-6325 respectively. Here the improvement in inhibition is due to the synergistic effects of the quinoxaline moiety, shown to favorably stack against the invariant catalytic residues, and macrocyclization [68].

The differential efficacy of PIs across GT-1 and 3 is in agreement with the documented failures of inhibitors against resistant variants caused by single or multi-amino acid mutations [59, 80, 150]. Thus, a careful analysis of the protease amino acid sequence across genotypes could provide insight into the molecular basis for PI failure in GT-3.

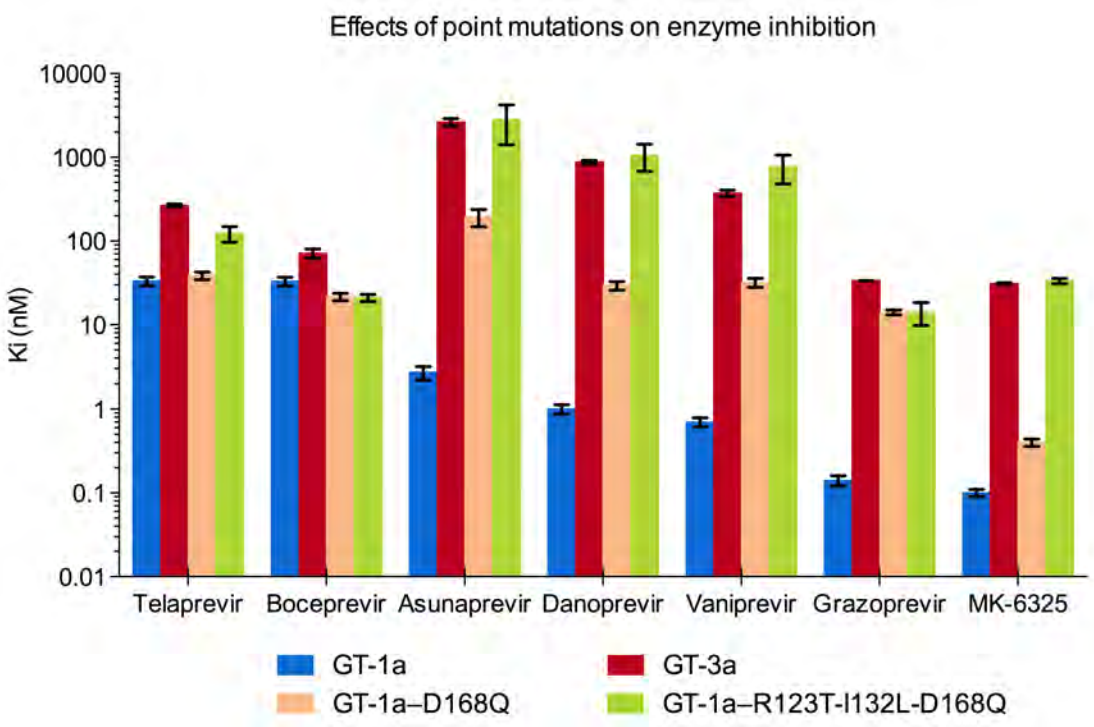


Figure 4.2: Effects of active site point mutations on NS3/4A inhibition
Enzyme inhibition constants (K_is) for 7 PI were determined in four enzyme constructs: WT–GT-1a (blue), WT–GT-3a (red), D168Q–GT-1a (orange), R123T-I132L-D168Q-GT-1a (green).

4.3.2 Of all genotypes, only GT-3 protease harbors active site polymorphisms

To analyze the diversity of NS3/4A across genotypes, sequences from the 6 genotypes were aligned and percent identities were calculated. Across genotypes, the viral protease is highly conserved with sequence identities and similarities ranging from 71-84% to 82-92% respectively (Fig. 4.3). While the GT-3 NS3/4A has the second most divergent sequence (76% identity vs. 71% in GT-2) relative to GT-1, all PIs are less potent against GT-3 protease than GT-2; with simeprevir, danoprevir, and vaniprevir more potent in GT-5 and 6 (both 82% identical to GT-1) than in GT-2. The various polymorphic sites are spatially distributed throughout the protease in all six GTs (Fig. 4.4), with a distinct added feature in GT-3.

PERCENT IDENTITY MATRIX

HCV.1a.M62321	100.0%					
HCV.2b.D10988	71.1%	100 0%				
HCV.3a.D17763	76.1%	70.6%	100 0%			
HCV.4a.DQ418788	80 0%	73.3%	75.0%	100.0%		
HCV.5a.AF064490	82 8%	74.4%	77.8%	80.0%	100 0%	
HCV.6b.D84262	82 2%	73.3%	79.4%	81.7%	84.4%	100.0%
	HCV.1a.M62321	HCV.2b.D10988	HCV.3a.D17763	HCV.4a.DQ418788	HCV.5a.AF064490	HCV.6b.D84262

PERCENT S MILARITY MATRIX

HCV.1a.M62321	100.0%					
HCV.2b.D10988	83 3%	100 0%				
HCV.3a.D17763	87 8%	82.2%	100 0%			
HCV.4a.DQ418788	90 0%	86.1%	86.7%	100.0%		
HCV.5a.AF064490	93 3%	87.8%	87.8%	90.6%	100 0%	
HCV.6b.D84262	92 2%	86.7%	90.6%	92.8%	92.2%	100.0%
	HCV.1a.M62321	HCV.2b.D10988	HCV.3a.D17763	HCV.4a.DQ418788	HCV.5a.AF064490	HCV.6b.D84262

GLOBAL S MILARITY (BLOSUM62) RESULTS

Reference protein	HCV.1a.M62321	HCV.2b.D10988	HCV.3a.D17763	HCV.4a.DQ418788	HCV.5a.AF064490	HCV.6b.D84262
HCV.1a.M62321	1					
HCV.2b.D10988	0.74	1				
HCV.3a.D17763	0.78	0.74	1			
HCV.4a.DQ418788	0.82	0.79	0.78	1		
HCV.5a.AF064490	0.87	0.79	0.81	0.83	1	
HCV.6b.D84262	0.84	0.77	0.81	0.83	0.86	1

Figure 4.3: HCV NS3/4A sequence simularity matrix
NS3/4A amino acid sequence simularity is shown for GT-1 through 6.

Active site differences between GT-1 and GT-3 were mapped onto the protease structure (Fig. 4.4B). The NS3/4A active site is highly conserved. Of the 36 residues in direct contact with inhibitors (within 5 Å of the ligand), GT-3 differs in only 3 amino acids: Thr123, Leu132 and Gln168 (Fig. 4.4C). While these changes represent an 8% divergence, the nature of the change is functionally significant. The GT-1 positively charged Arg123 is substituted by a polar uncharged threonine in GT-3, while the negatively charged GT-1 Asp168 is mutated into a polar uncharged glutamine. These two changes effectively alter the active site electrostatics and hydrogen bonding pattern. Amino acids 123 and 168 were previously reported for their contribution to an active site salt bridge formation, inhibitor binding, and resistance in the case of residue 168 [68, 94, 106]. Hence, polymorphisms at the GT-3 protease active site may lead to electrostatic changes with implications for inhibitor efficacy.

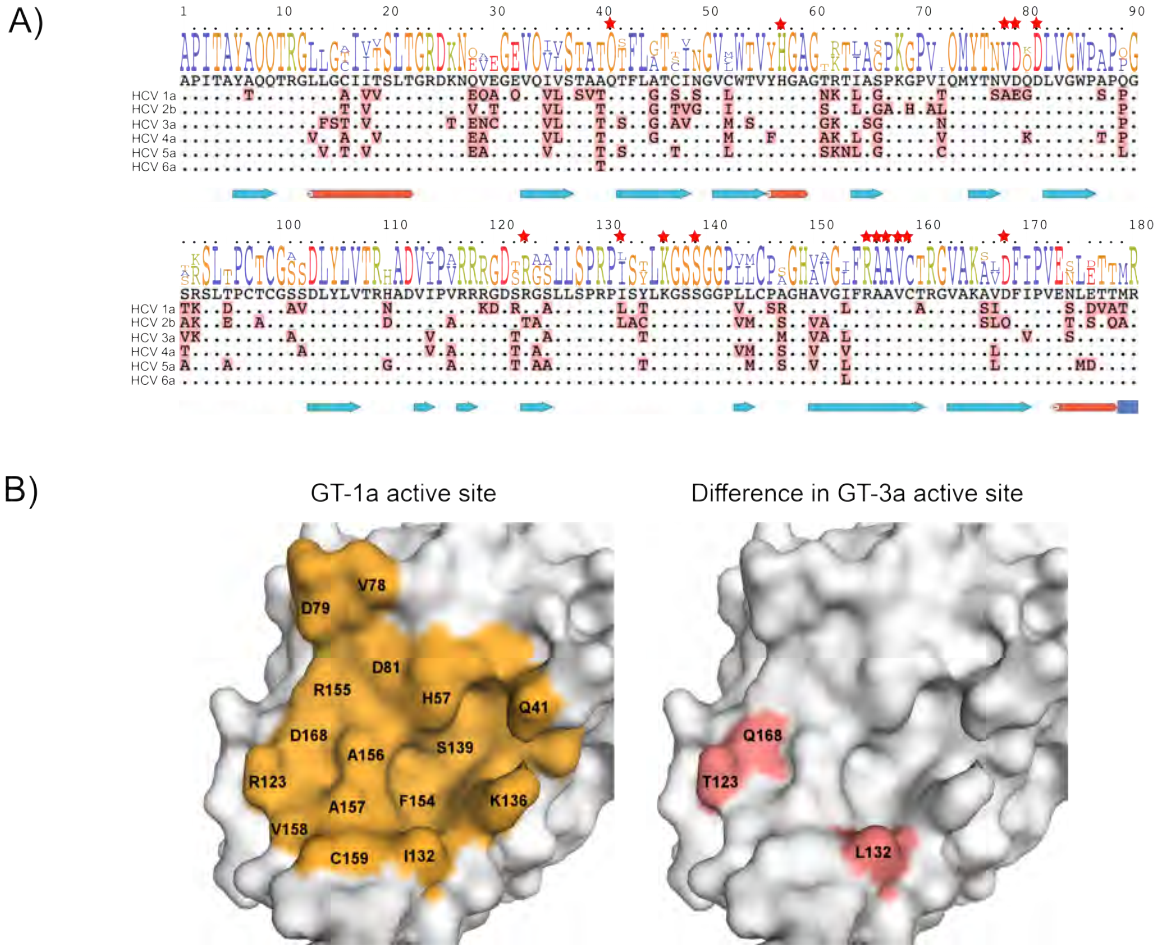


Figure 4.4: Strcuturally mapping the HCV NS3/4A polymorphism **(A)**Alignment of NS3/4A protease domain in GT 1 through 6 with consenesus and weblogo displayed. Red stars identify amino acids located in the protease active site. **(B)** Mapping of polymorphic differences between GT-1 and 3 active sites. GT-1 active site residues involved in inhibitor-protease contacts (orange) are shown (left panel); GT-3 active site residues with differences relative to GT-1 shown (red).

4.3.3 Polymorphisms at amino acids 123, 132, and 168 are responsible for GT-3 specific decreased inhibitor potency.

To analyze the role of the Thr123, Leu132 and Gln168 polymorphisms on inhibitor efficacy, we engineered a single D168Q and triple R123T/I132L/D168Q (1a3a) mutant protease in the GT-1 background and characterized their enzyme inhibition profiles with the 7 PIs listed above (Fig. 4.2). In the linear covalent inhibitors telaprevir and boceprevir, the D168Q single mutation does not significantly affect inhibitor potency relative to the GT-1 WT (1.2 and 0.7 FC respectively), likely due to a lack of an extended P_2 moiety reliant on interactions with the S_2 subpocket. In contrast, this single mutation causes detrimental effects on the non-covalent inhibitors. In ASV, DAN, VAN, and GRZ, the GT-1 D168Q variant results in a 72, 30, 44 and 102 fold loss in potency relative to GT-1 WT. However, with only a 4-fold loss in potency, MK-6325 is able to accommodate the Asp to Gln change and maintain a subnanomolar inhibition constant. The synergistic effects of MK-6325's bis-macrocycle and cyclopentyl P_4 capping may underlie this sustained efficacy in the presence of D168Q mutation (AppendixA). In the NS3/4A protease, amino acid 168 is a multidrug resistance mutation hotspot with G/A/V/E/N/Y/K substitutions observed both experimentally and clinically in GT-1 infected patients [59]. Understanding MK-6325's resistance profile and binding mode against D168 variants might provide insight into optimizing PIs for better GT-3 targeting.

Introducing the R123T/I132L/D168Q triple mutation in GT-1 recapitulates the GT-3 inhibitor profile for most of the drugs in the panel. ASV, DAN, VAN and MK-6325 experience a three order of magnitude loss in potency relative to GT-1 WT, essentially mirroring the observed loss of potency against GT-3. MK-6325 loses its advantage over D168Q when I132L and R123T are present. Finally, the triple mutant has the same sensitivity as the D168Q variant to BOC and GRZ, indicating this mutation is the main reason for the

loss of potency against GT-3.

4.3.4 Crystal structures and molecular dynamics simulations of inhibitor complexes

To gain insights into the structural basis for GT-3 escape from PIs, we attempted to structurally characterize native GT-3 protease. However, extensive crystallization efforts were unsuccessful due to expression and purification challenges. Thus, we resorted to our 1a3a chimeric NS3/4A, which includes three key polymorphisms at the GT-3 active site and fully recapitulates the inhibition profile. We determined high-resolution crystal structures of ASV (5EQS), DAN (5EGR) and VAN (5ESB) in complex with 1a3a chimeric NS3/4A protease (Table A1) and compared them to their corresponding GT-1 structures [85, 106]. To supplement this analysis, fully solvated 100 ns molecular dynamic (MD) simulations were performed in triplicates for each of the complexes. The structures and simulations permitted an in depth structural analysis of these complexes providing insights into the molecular mechanisms which differentiate GT-3 activity.

Table 4.1: Crystallographic Statistics of 1a3a in complex with asunaprevir, danoprevir and vaniprevir

	Asunaprevir (ASV)	Danoprevir (DAN)	Vaniprevir (VAN)
Complexes	1a/3a	1a/3a	1a/3a
PDB Code	5EQS	5EGR	5ESB
Resolution (Å)	1.8	1.9	2.2
Space Group	P_{212121}	P_{212121}	P_{212121}
Molecules in AS	1	1	1
Cell Dimensions			
a	39.1	54.9	55.1
b	60.3	58.4	58.3
c	79.6	59.9	59.9
β (°)	90	90	90
Completeness (%)	92	92.8	94
Measured Reflections	69630	60292	33585
Unique Reflections	15339	14041	7436
Average I/(σ)	10.2	7.7	19.5
Redundency	4.4	4.3	4.5
R-merge (%)	9	9	6.2
RMSD in			
Bonds (Å)	0.015	0.012	0.042
Angles (°)	1.6	1.3	1.8
R_{factor}	19.3	16.5	18.2
R_{free}	24.5	20	23.2

4.3.5 Differences in GT-3 active site alter inhibitor packing

To analyze the details of inhibitor packing at the active site, van der Waals (vdW) interaction energies between the inhibitor and protease residues in the crystal structures were calculated (Fig. 4.5). In line with the inhibitor potency losses (Fig. 4.2), vdW contacts are moderately lost in the 1a3a complexes relative to GT-1 (ASV: -44.4 vs. -45.3 kcal/mol; DAN: -41.7 vs. -43.7 kcal/mol; and VAN: -41.1 vs. -41.5 kcal/mol). However these changes in contacts are not sufficient to fully justify the extent of potency loss in GT-3.

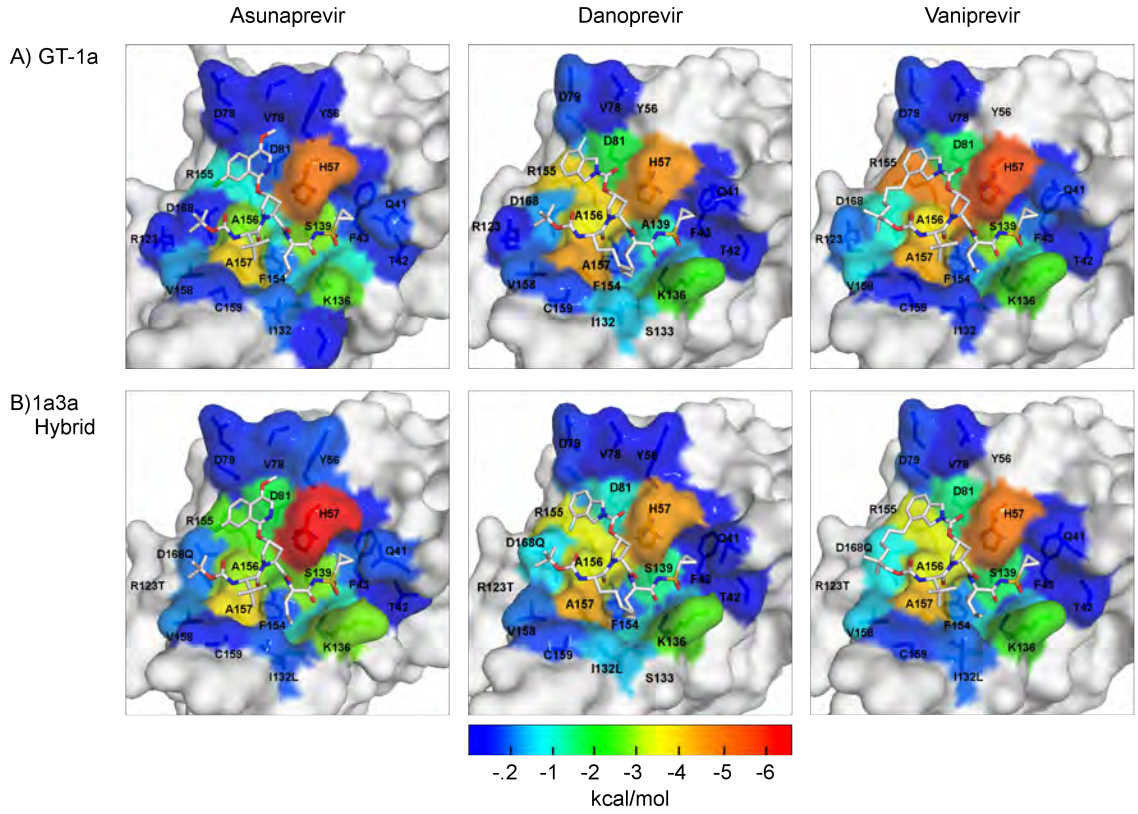


Figure 4.5: Surface representation of van der Waals potential energy of active site residues. The vdW contact energy with the inhibitor mapped onto the protease surface in **(A)** WT–GT-1 and **(B)** 1a3a chimera. The warmer (red) and cooler (blue) colors indicate more and less contacts with the inhibitor, respectively.

4.3.6 Disruption of the active site electrostatic network correlates with inhibitor efficacy

In HCV protease, high-affinity ligand binding is characterized by formation of an electrostatic network spanning residues D81, R155, D168, and R123 [68, 85]. Within this network, R155–D168 interaction is particularly key for maintaining an electrostatic surface that accommodates high-affinity ligand binding [106]. This interaction is present in all GT-1 complex crystal structures (Fig. 4.6A), with both carbonyl oxygens of Asp168 oriented towards Arg155's Nϵ and Nη to form two hydrogen bonds (H-bonds) and Arg132's Nϵ oriented towards D168's Oδ for one H-bond formation. However, in the 1a3a complexes, the D168Q substitution results in simultaneous loss of the R155-Nη interaction and maintenance of the 168-CO–155-Nϵ interaction (Fig. 4.6B).

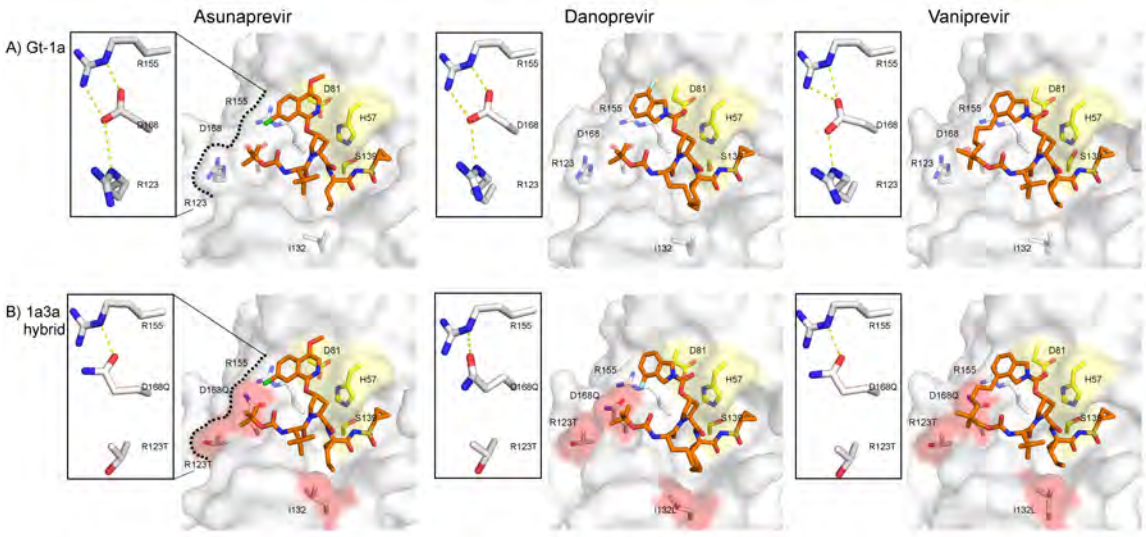


Figure 4.6: Mutation induced change in active site electrostatic interactions
Surface representation of protease-inhibitor complexes with Asunaprevir (left), Danoprevir (center) and Vaniprevir (right) shown in orange sticks in the context of the wild type GT-1 **(A)** and 1a3a chimeric protease **(B)**. Catalytic residues and mutation sites are highlighted in yellow and red respectively, while the observed electrostatic interaction and hydrogen bonding (yellow dotted line) are shown in the inset.

Rearrangements of this electrostatic interaction underlie the resistance mechanism of many PIs, including asunaprevir and danoprevir, to the R155K multidrug resistant mutation [68, 106], where Asp168 side chain shifts towards Arg123 for electrostatic interaction. With Gln168 only engaging in one H-bond with Arg155, the possibility of a rotation of Gln168 to Arg123 for electrostatic interaction is feasible. Thus, to assess the stability of the active site electrostatic interaction, we computed the average time the interactions were observed during MD simulations (Table 4.2). The simulations revealed that, in all three GT-1 complexes, both of D168s carbonyl oxygens were oriented towards Arg155's Nϵ and Nη over 93% of the simulation duration. In agreement with recent reports [107], the Arg123–D168 interaction was only observed in the ASV and VAN complex roughly 30% of the time, suggesting that Arg123 might not be important for the formation of the inhibitor binding surface.

With the loss of a terminal carboxyl group at amino acid 168, all three 1a3a complexes experience weaker Gln168–Arg155 electrostatic interaction relative to their GT-1 counterparts. In ASV, this interaction is observed only during 47% of the simulation. This 52% decrease relative to GT-1 is the most severe of the three complex structures and correlates with ASVs weaker antiviral activity in GT-3. On the other hand, with DAN and VAN, the 168–155 interaction drops only by roughly 21% relative to their 1a counterparts suggesting that the added rigidity of the ligand may help in maintaining the electrostatic interaction and the electrostatics of the surface. Furthermore, there appears to be a macrocycle–location dependent effect on the formation of the electrostatic surface. Comparing the three 1a3a complexes, the Q168–R155 interaction is most persistent in DAN (78%) where the macrocycle is at the P_1-P_3 versus VAN (64%) and ASV (47%) representing respectively the P_2-P_4 macrocyclic and linear inhibitors. The substitutions in GT-3 active site relative to GT-1 do not significantly change protease–inhibitor intramolecular H-bonds, except for the loss of at least one H-bond at the $P_{1'}$ moiety for all three PIs. Overall, the molecular rear-

rangements in ASV lead to the most changes in H-bond ranging from -10 to -61%, while both DAN and VAN are better able to preserve their H-bonding interactions (Table 4.3). Hence, the active site substitutions profoundly disrupt the active site electrostatic network and consequently affect inhibitor binding and intermolecular H-bond signatures.

4.3.7 Loss in enzyme–inhibitor atomic fluctuation correlations underlies inhibitor efficacy characteristics.

To assess the overall consequences of the active site substitutions on protease–inhibitor dynamics, we computed the root square mean fluctuations (RMSF) of both inhibitor and protease Cα atoms during the MD simulations. The 1a3a triple mutation had little influence on the backbone Cα RMSF values relative to the GT-1, suggesting that the overall protein dynamics remained intact (Fig. 4.9). However, the inhibitors atomic fluctuations changed both depending on the inhibitor and when bound to the chimeric protein relative to GT-1. While the inhibitor backbone appears relatively rigid, the P_2 moieties sample different degrees of mobility (Fig. 4.7).

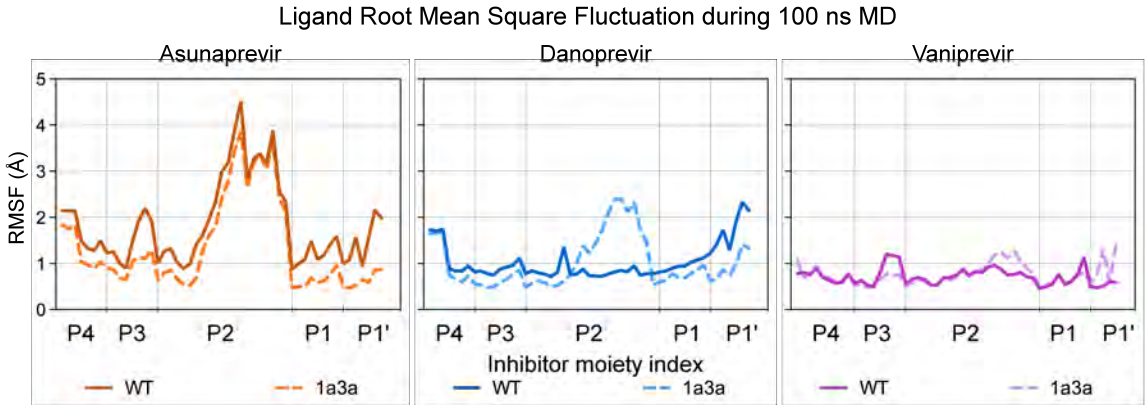


Figure 4.7: Root square mean fluctuations of inhibitor atoms during MD simulations Atomic RMSF of asunaprevir (orange), danoprevir (blue) and vaniprevir (magenta) in complex with WT–GT-1 (solid lines) and 1a3a (dashed lines) during 100 ns MD simulation are shown.

In ASV, the isoquinoline undergoes the highest RMSF fluctuations of the complexes (up to 4 Å RMSF) regardless of the protein construct. DAN's isoindoline is rigid compared to ASV (1 Å RMSF) but becomes more dynamic in the 1a3a complex (2 Å RMSF). VAN, with its P_2-P_4 linked isoindoline, has the most rigid P_2 of the three inhibitors and undergoes no change in fluctuations due to the 1a3a mutations. Intriguingly, the differences in inhibitor P_2 atomic fluctuations is inversely correlated with intermolecular vdW contacts and binding potency (Fig, 4.2 and 4.5A). ASV's isoquinoline is highly dynamic in GT-1, yet, among the three inhibitors, it makes the least contacts with the protease S_2 residues resulting in the weakest-binding inhibitor. The isoindoline in both DAN and VAN display very low fluctuations, but make strong contacts with the S_2 residues, producing tighter-binding drugs. Hence, there is an interdependent relationship between inhibitor P_2 flexibility, protease-inhibitor contacts and potency.

In principle, tight binding inhibitors are characterized by strong intermolecular interactions, which persist over the dynamics of the enzyme [94]. To draw a stronger link between the interdependence of the electrostatic interaction and the binding modes of inhibitor P_2 isoquinoline and isoindoline, we determined cross-correlation coefficients of atomic fluctuations between protease active site residues and inhibitor atoms. During the MD simulations, atom pairs either fluctuate in the same direction (positive correlation, $CO_{i.j} = 1$), opposite direction (negative correlation, $CO_{i.j} = -1$) or have random movements with respect to one other (no correlation, $CO_{i.j} = 0$).

In the 1a complexes, ASV and VAN display strong positive $(CO_{i.j} > 0.6)$ correlations, specifically with respect to protease residues 132 through 157 and including 168 (Fig. 4.10), whereas DAN has both positive and negative correlations with respect to residues 132 through 168 and 41 through 43 respectively. In addition to the conserved R155–D168 interaction revealed by conserved H-bond interactions, the cross correlation analysis showed high cooperativity between the P_2 moieties of ASV and DAN with R155–D168 residues.

VAN's isoindoline has positive correlations with R155 but not with D168. It is possible that the D168 aids in stabilizing the R155's conformation for parallel stacking with the heterocyclic rings.

In the 1a3a complexes, the cross-correlations exhibit significant alterations ranging from partial losses to complete disruption compared to GT-1. In ASV, the P_4-P_3 moieties experience complete loss of correlations with eight residues, 132 through 154 (Fig.4.8), constituting the oxyanion hole, S_1 pocket and the 3_{10}-helix located between β strands C2 and D2. This disruption in correlations at the S_1 subpocket is in agreement with the observed loss in $P_{1'}$ intermolecular interactions in this inhibitor. Similar to ASV, DAN suffers complete disruptions in correlations with respect to residues 132 through 138 (Fig. 4.11). While protease–P_4-P_3 moiety correlations are somewhat maintained in the 1a3a complexes, the intermolecular cooperativity relative to the inhibitor P_2 moieties of ASV and DAN is completely abolished due to the mutations (Fig. 4.8B, 4.11B) but slightly maintained for VAN (Fig. 4.12B).

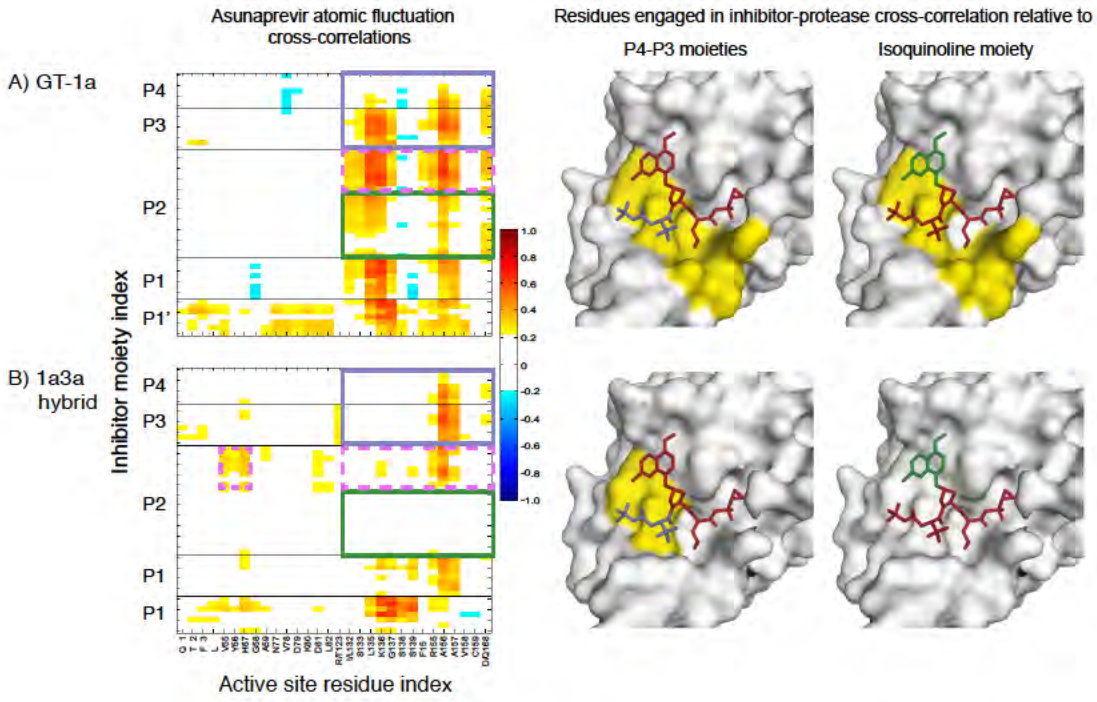


Figure 4.8: Effects of 1a3a mutations on intrinsic Protease–ASV dynamic cooperativity **(Left)** Atomic fluctuation cross-correlation matrix of active site residues from GT-1 (top) and 1a3a (bottom) in complex with asunaprevir. Sections representing the P_4-P_3 (purple), P_2 isoquinoline (green) and P_2 proline (pink dashes) are highlighted in boxes. Correlation coefficients are represented using a red–blue color scale where positive and negative correlations are represented by red and blue shades respectively. **(Right)** Surface representation of cross correlation changes relative to the P_4-P_3 (purple) and P_2 (green) moieties and boxed sections in (A).

4.3.8 Conclusions

HCV therapeutic efforts have greatly benefited from structure based drug design (SBDD), specifically protease inhibitors. The elucidation of the structural mechanism for substrate binding, coupled with rational medicinal chemistry efforts, has spawned the current landscape of PIs with subnanomolar potency and remarkable antiviral activity. Sadly, this activity is genotype dependent and new more effective therapies are needed for GT-3 infected patients. The same SBDD efforts that have facilitated these advancements are hindered by the lack of GT-3 protease structural information and robust design tools. In this study, we reveal the major contributors to PI failure in GT-3, and demonstrate that three active site polymorphisms, R123T-I132L-D168Q, are responsible for efficacy loss of current PIs against GT-3. Our crystal structures and dynamic analysis indicate that the active site electrostatics, which we have reported previously to be key for tight ligand binding [68, 106, 107], are disrupted by GT-3 active site polymorphism.

The molecular mechanisms underlying loss of PI efficacy against GT-3 relative to GT-1 is reminiscent of drug resistance due to active site mutations in GT-1 protease [68, 107], understanding drug resistance can serve as a template for understanding inhibitor efficacy in GT-3 protease. The ability of single active site mutations to confer multidrug resistance has been extensively studied, specifically in the context of D168 resistance associated variants. These mutations unfavorably alter or completely disrupt the electrostatic network required for ligand binding. Inhibitors that retain potency against D168 variants of GT-1 protease may be better inhibitors for GT-3 protease, and the chimeric 1a3a protease that we have constructed here can be used to test such designed inhibitors for GT-3 activity in both structural studies and potentially replicon assays.

In viruses like HCV, where 10^{12} new virions are produced daily, it is estimated that every possible nonsense mutation is introduced in the viral genome on a daily basis [62].

Consequently, a heterogeneous viral population exists in every patient. However, outside of the native GT-3 context, neither the single D168Q nor the simultaneous presence of the triple R123T/I132L/D168Q mutations has been reported. The absence of these mutations in GT-1 suggests other secondary GT-3 polymorphisms outside the active site may be required for maintaining protease fitness. In the context of GT-3, the R123T and D168Q mutations disrupt the R123-D168-R155 electrostatic network to an only Q168–R155 interaction.

A thorough understanding ligand-receptor interactions is key for a successful drug discovery project. For instance, an appreciation for the molecular basis for drug resistance in HIV-1 protease was not possible until a complete understanding of the atomic basis for substrate recognition and inhibitor binding was established. The resulting findings, known as the substrate envelope hypothesis (SEH), demonstrated that (1) active site amino acids prone to resistance associated variants correspond to sites where an inhibitor protrudes outside of the consensus substrate vdW volume [151]. (2) When used as an added constraint in drug development, the SEH has led to the design robust inhibitors with enhanced resistance barriers and improved antiviral activity [152]. SEH was successfully implemented to design HIV PIs more robust against resistance in SBDD, and was fundamental in our understanding of the structural basis for HCV drug resistance [80,85]. In addition to SEH, the study of atomic cross-correlations we present here for the 1a3a chimera could greatly enhance PI design against GT-3. For instance, inhibitor moieties could be first modeled in 1a3a and tested for their ability to maintain or increase correlations relative to GT-1. As both the P_4-P_3 and P_2 moieties of ASV, DAN, and VAN were shown to lose correlations in the chimeric complex, SAR studies to identify moieties with stronger cross correlations could yield more robust inhibitors. This hypothesis is supported by a recent report detailing a P_4 capping SAR in a faldaprevir derivative [153].

In conclusion, we present the alterations in the active site electrostatics as the molecular mechanism underlying loss of PI efficacy against HCV GT-3, and a chimeric 1a3a protease

amenable for structural studies for the study and future development of more robust GT-3 NS3/4A inhibitors. By mapping the major amino acids responsible for the differential activity of PIs across GT-1 and GT-3, our chimera enables understanding the lack of potency of PIs against GT-3 protease, allows for a direct comparison of ligand binding modes against the extensive repertoire of GT-1 inhibitor complexes, and serves as an added tool in the computational drug discovery toolkit. The commonality between mechanisms underlying drug resistance due to D168 active site mutations in GT-1 and loss of potency against GT-3 promises the possibility of designing inhibitors robust against both GT-1 mutations and GT-3 polymorphisms. Revealing molecular mechanisms of how polymorphisms alter inhibitor binding across genotypes should not be limited to HCV or other Flaviviruses, but be applied to other rapidly evolving diseases where viral heterogeneity alters inhibitor efficacy across genotypes.

4.4 Materials and Methods

4.4.1 Protein Constructs

The HCV genotype 1a NS3/4A protease domain gene [68] was synthesized by GenScript and cloned into the pET28a expression vector (Novagen). The highly soluble single-chain construct consists of NS3/4A protease domain (residues 4–181) fused to a fragment of the cofactor NS4A (residues 12–23) via an SD linkage. A similar protease construct exhibited catalytic activity comparable to that of the authentic full-length protein [115]. All protease variants were generated using the QuikChange Site-Directed Mutagenesis Kit from Stratagene.

4.4.2 Protein Expression and Purification

Protein expression and purification were carried out as previously described [68]. Briefly, transformed BL21(DE3) E. coli cells were grown at $37\,^{\circ}\mathrm{C}$ and induced at an optical density of 0.6 by adding 1 mM IPTG. Cells were harvested after 5 hours of expression, pelleted, and frozen at -$80\,^{\circ}\mathrm{C}$ for storage. Cell pellets were thawed, resuspended in 5 mL/g of resuspension buffer (50 mM phosphate buffer, 500 mM NaCl, 10% glycerol, 2 mM β-ME, pH 7.5) and lysed with a cell disruptor. The soluble fraction was retained, applied to a nickel column (Qiagen), washed with resuspension buffer, and eluted with resuspension buffer supplemented with 200 mM imidazole. The eluent was dialyzed overnight (MWCO 10 kD) to remove the imidazole, and the His-tag was simultaneously removed with thrombin treatment. The nickel-purified protein was then flash frozen and stored at -$80\,^{\circ}\mathrm{C}$.

4.4.3 Crystallization

The above-mentioned protein solution was thawed, concentrated to 3 mg/mL and loaded on a HiLoad Superdex75 16/60 column equilibrated with gel filtration buffer, 25 mM 2-(N-Morpholino)ethanesulfonic (MES), 500 mM NaCl, 10% glycerol, 30 mM zinc chloride, and 2 mM 1,4-dithiothreitol, pH 6.5. The protease fractions were pooled and concentrated to 20–25 mg/mL with an Amicon Ultra-15 10 kD device (Millipore). The concentrated samples were incubated for 1 hour with 1–3 molar excess of inhibitor. Diffraction-quality crystals were obtained overnight by mixing equal volume of concentrated protein solution with precipitant solution (20–26% PEG-3350, 0.1 M sodium MES buffer, 4% ammonium sulfate, pH 6.5) in 24-well VDX hanging drop trays.

4.4.4 Data Collection and Structure Solution

X-ray diffraction data were collected on an in-house Rigaku X-ray system with a Saturn 944 CCD detector. Diffraction intensities were indexed, integrated and scaled using the program HKL2000 [116]. All structure solutions were generated using molecular replacement with PHASER [89]. The B chain model of viral substrate product 4A-4B (PDB ID: 3M5M) [80] was used as the starting model for all structure solutions. Initial refinement was carried out in the absence of modeled ligand, which was subsequently built in during later stages of refinement. Subsequent crystallographic refinement was carried out within the Phenix program suite, with iterative rounds of TLS or restrained refinement until convergence was achieved [117]. 1a3a–VAN complex was refined using the CCP4 program suite [154]. The final structures were evaluated with MolProbity [118] prior to deposition in the Protein Data Bank. To limit the possibility of model bias throughout the refinement process, 5% of the data were reserved for the free R-value calculation [119]. Interactive model building and electron density viewing was carried out using the program COOT [91].

4.4.5 Enzyme Inhibition Assays

All enzyme inhibition assays were performed in nonbinding surface 96-well black half-area plates (Corning) in a reaction volume of 60 μL. The NS3/4A protease (2 nM) was preincubated with increasing concentration of drugs in 50 mM Tris, 2.5% glycerol, 0.1% OβG, 5 mM Tris(2-carboxyethyl)phosphine, 1% Dimethyl sulfoxide, pH 7.5 for an hour. The reaction was initiated by the rapid injection of 5 μL of HCV NS3/4A protease substrate, Ac-DE-Dap(QXL-520)-EE-Abu-ψ-[COO]AS-C(5-FAMsp)- NH2 (Anaspec), to a final concentration of 200 nM and kinetically monitored using a Perkin-Elmer EnVision plate reader (excitation at 485 nm; emission at 530 nm). At least four independent data sets were collected for each inhibitor with each protease construct. Each inhibitor titration included at least 12 inhibitor concentration points, which were globally fit to the Morrison equation to obtain the K_i value [81].

4.4.6 Structural Analysis

Superpositions were performed in PyMOL [122] Cα atoms of the active site protease residues 137–139 and 154–160. The A chain of WT–asunaprevir complex was used as the reference structure for each alignment. Hydrogen bonding analysis was carried out in Maestro, Schrodinger suite, with bond distance and donor-acceptor angles cutoff of 3.5 Å and 120°, respectively. VAN der Waals contact energies between protease–inhibitor residues were computed using a simplified Lennard–Jones potential, as described previously [121].

4.4.7 Molecular Dynamics Simulations

Molecular dynamics simulations were carried out in triplicate, following previously published protocols [94] using Desmond [122] with the OPLS2005 force field [124] [125]. After equilibration, each trajectory was run for 100 ns at $300\,^{\circ}\mathrm{K}$ and the coordinates recorded

every 5 ps. The percentage of time a hydrogen bond existed between the protease and an inhibitor was calculated using VMD [155]. A hydrogen bond was defined by a distance between the donor and acceptor of less than 3.5 Å and a hydrogen-donor-acceptor angle of less than 30°. Only the hydrogen bonds that existed more than 20% of the time were considered in the analyses. Salt bridges were defined as an interaction between a side-chain oxygen atom of Asp or Glu within 4.0 Å of a nitrogen atom of Arg or Lys [156].

4.4.8 Atomic Fluctuation Dynamics

The normalized cross-correlations of residue pairs were defined as:

$$CO_(i.j) = \frac{\langle \overrightarrow{\Delta R_i} \cdot \overrightarrow{\Delta R_j} \rangle}{(\langle \Delta {R_i}^2 \rangle \langle \Delta {R_j}^2 \rangle)^{1/2}}$$

where Ri is the fluctuation in the position vector R of site i and Rj is the fluctuation in the position vector R of site j. The brackets represent time averages over recorded snapshots. The cross-correlations vary in the range [-1, 1], with the lower and upper limits indicating fully anticorrelated and correlated atomic fluctuations, respectively. $CO_{i,j} = 0$ reflects uncorrelated atomic fluctuations.

4.5 Acknowledgments

We thank W. Royer, F. Massi and D. Lambright for helpful discussions. The National Institute of Allergy and Infectious Disease (R01-AI085051) supported this work. DIS was also supported by National Institute of General Medical Sciences of the NIH (F31-GM103259) as well as the HOPE scholarship sponsored by the Biomedical Science Career Program (BSCP).

4.6 Supplemental Information

Coordinates of crystal structures were deposited in the RCSB http://www.rcsb.org. PDB codes are: 5EQS for 1a3a–ASV, 5EGR for 1a3a–DAN, and 5ESB for 1a3a–VAN. The authors declare no competing financial interest.

		Asunaprevir		Danoprevir		Vaniprevir	
		WT	1a3a	WT	1a3a	WT	1a3a
SER1139-Main-N	P1' SO1	3.1(48.1%)	2.9(89.4%)	2.7(<20%)	2.7(<20%)	2.7(86.6%)	2.8(<20%)
SER139 H	P1' SO1	1.9(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)
SER1139-Side-Oγ	P1' SO1	2.1(17.4%)	>3.5(38.7%)	>3.5(<20%)	2.0(19.8%)	>3.5(92.6%)	2.1(20.2%)
GLY1137-Main-N	P1' SO2	2.2(70.1%)	2.2(81.6%)	2.3(83.2%)	2.3(85.2%)	2.1(84.6%)	2.2(88.2%)
HIP1057-Side-Nϵ2	P1' NH	2.1(<20%)	2.0(62.1%)	2.1(34.9%)	2.2(59.3%)	2.0(61.3%)	2.1(58.1%)
GLY1137-Main-N	P1 CO	2.0(<20%)	2.1(<20%)	2.0(<20%)	2.1(<20%)	2.3(<20%)	2.2(<20%)
SER1138-Main-N	P1 CO	2.9(<20%)	3.0(<20%)	2.7(<20%)	2.5(<20%)	2.8(<20%)	2.6(<20%)
SER1139-Side-Oγ	P1 CO	2.6(<20%)	2.8(<20%)	2.6(<20%)	2.5(<20%)	2.7(<20%)	2.5(<20%)
ARG1155-Main-O	P1 NH	1.9(57.5%)	2.0(84.4%)	1.9(66.7%)	2.0(84.4%)	2.0(92.6%)	1.9(83.6%)
ALA1157-Main-O	P3 NH	1.9(43.3%)	1.9(41.9%)	1.9(82.0%)	1.9(79.2%)	2.0(88.2%)	2.0(82.0%)
ALA1157-Main-N	P3 CO	1.9(75.1%)	1.9(60.7%)	1.9(85.4%)	2.0(80.4%)	1.9(83.0%)	1.9(83.0%)

Table 4.2: 1a3a Intermolecular H-bonding network of protease inhibitor complexes. Hydrogen bond network observed in both crystal structures (distance in Å) and their averaged persistence during 100ns simulations (percent time).

		Asunaprevir		Danoprevir		Vaniprevir	
		WT	1a3a	WT	1a3a	WT	1a3a
TYR56-Main-N	ASP81-Main-O	2.1(80.4%)	2.0(18.0%)	2.2(81.6%)	>3.5(50.1%)	2.2(80.4%)	>3.5(23.0%)
TYR56-Main-N	ASP81-Side-Oδ2	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)
HIS57-Main-N	ASP81-Side-Oδ	1.8(72.5%)	1.9(93.8%)	1.9(72.1%)	1.8(82.6%)	1.9(72.5%)	1.8(85.1%)
HIS57-Side-Nδ1	ASP81-Side-Oδ	1.6(96.8%)	1.6(89.2%)	1.6(98.4%)	1.6(96.4%)	1.6(96.8%)	1.8(97.0%)
ARG123-Side-Nϵ	ASP168-Side-Oδ1	2.7(<20%)	>3.5(<20%)	2.4(<20%)	>3.5(<20%)	1.9(<20%)	>3.5(<20%)
ARG123-Side-NH2	ASP168-Side-Oδ1	>3.5(31.3%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	3.4(31.3%)	>3.5(<20%)
LYS136-Main-N	ILE132-Main-O	2.3(28.1%)	2.0(25.8%)	2.2(<20%)	2.3(<20%)	2.1(28.1%)	2.5(<20%)
SER138-Side-Oγ	LYS136-Main-O	>3.5(10.0%)	>3.5(<20%)	>3.5(33.5%)	>3.5(21.2%)	>3.5(<20%)	>3.5(<20%)
SER138-Side-Oγ	GLY137-Main-O	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)
GLY140-Main-N	PHE154-Main-O	1.8(90.0%)	1.9(92.2%)	1.9(85.4%)	1.8(80.4%)	1.9(90.0%)	1.9(88.4%)
GLY141-Main-N	SER138-Main-O	>3.5(<20%)	2.2(<20%)	2.1(<20%)	2.2(19.2%)	2.1(<20%)	2.2(<20%)
PHE154-Main-N	GLY141-Main-O	1.8(96.0%)	1.8(96.2%)	1.8(96.0%)	1.8(94.2%)	1.8(96.0%)	1.7(94.8%)
ARG155-Main-N	ASP168-Main-O	1.9(81.8%)	1.9(67.7%)	2.0(76.1%)	1.9(60.1%)	1.9(81.8%)	1.9(69.9%)
ARG155-Side-Nϵ	ASP168-Side-Oδ2	1.6(99.2%)	>3.5(25.0%)	1.7(99.0%)	>3.5(<20%)	1.9(99.2%)	>3.5(<20%)
ARG155-Side-Nϵ	GLN168-Side-Oϵ1	>3.5(<20%)	1.8(47.7%)	>3.5(<20%)	2.1(78.2%)	>3.5(<20%)	1.7(63.7%)
ARG155-Side-NH1	GLN168-Side-Oϵ1	>3.5(<20%)	>3.5(<20%)	>3.5(<20%)	2.6(<20%)	>3.5(<20%)	>3.5(<20%)
ARG155-Side-NH1	ASP168-Side-Oδ1	2.3(94.2%)	>3.5(<20%)	2.4(54.3%)	>3.5(<20%)	>3.5(94.2%)	>3.5(<20%)
ARG155-Side-NH1	ASP168-Side-Oδ2	>3.5(<20%)	>3.5(<20%)	>3.5(39.3%)	>3.5(<20%)	1.8(<20%)	>3.5(<20%)
ARG155-Side-NH2	LYS80-Main-O	1.9(76.1%)	1.7(46.1%)	2.0(81.6%)	2.2(78.0%)	2.0(76.1%)	2.1(67.7%)
ALA156-Main-N	ASP168-Main-O	2.1(50.1%)	2.1(86.6%)	2.1(71.5%)	2.2(90.2%)	2.1(50.1%)	1.9(88.4%)
VAL158-Main-N	ALA166-Main-O	1.9(63.3%)	1.9(60.1%)	1.9(82.8%)	1.8(75.3%)	1.9(63.3%)	2.2(69.9%)
LYS165-Main-N	VAL158-Main-O	2.6(<20%)	2.2(19.0%)	2.3(46.9%)	2.0(<20%)	2.3(<20%)	2.4(<20%)
ASP168-Main-N	ALA156-Main-O	1.8(71.3%)	1.8(77.1%)	1.8(85.2%)	1.6(70.1%)	1.8(71.3%)	1.6(66.1%)

Table 4.3: 1a3a Intramolecular hydrogen bonding network of protease inhibitor complexes Hydrogen bond network observed in both crystal structures (distance in Å) and their averaged persistence during 100ns simulations (percent time).

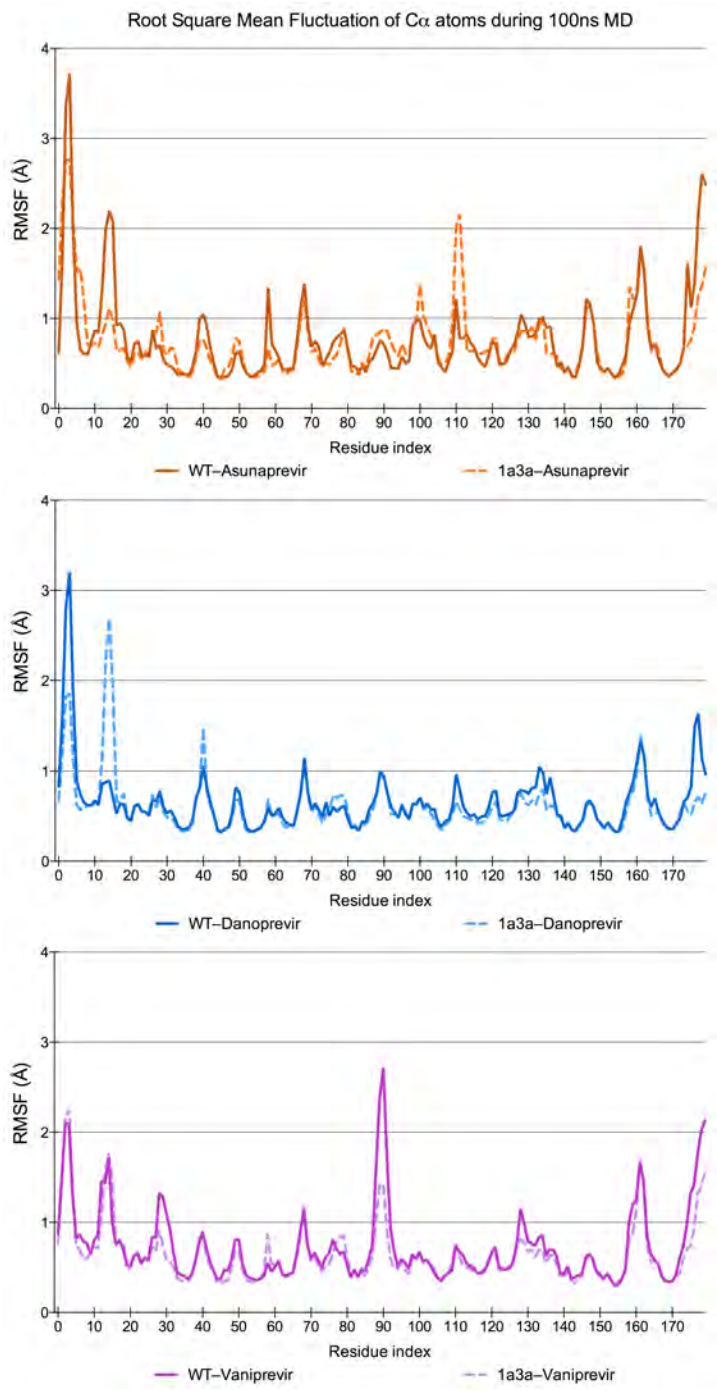


Figure 4.9: Root square mean fluctuations of protease C atoms during MD simulations. Atomic RMSF of asunaprevir (orange), danoprevir (blue) and vaniprevir (magenta) in complex with WT–GT-1 (solid lines) and 1a3a (dashed lines) during 100 ns MD simulation are shown.

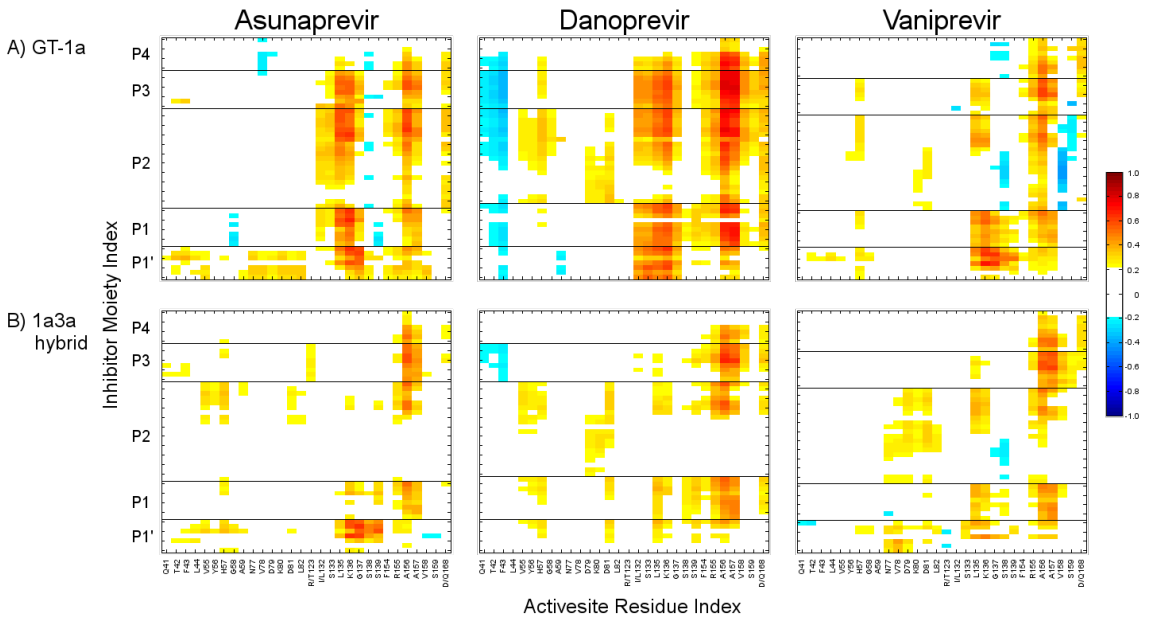


Figure 4.10: Effect of 1a3a mutations on intrinsic Protease–inhibitor dynamic cooperativity.

Atomic fluctuation cross-correlation matrix of active site residues from GT-1 **(A)** and 1a3a **(B)** in complex with asunaprevir (left), danoprevir (center) and vaniprevir (right). Correlation coefficients are represented using a red–blue color scale where positive and negative correlations are represented by red and blue shades respectively.

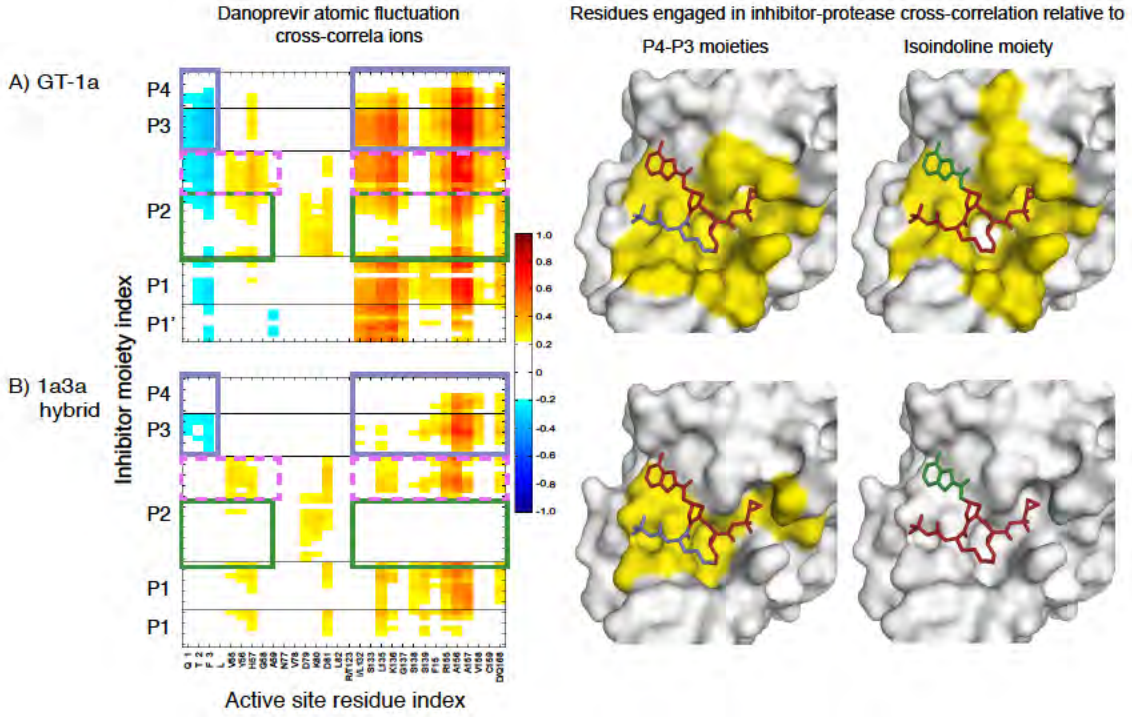


Figure 4.11: Effects of 1a3a mutations on intrinsic Protease–DAN dynamic cooperativity. **(Left)** Atomic fluctuation cross-correlation matrix of active site residues from GT-1 (top) and 1a3a (bottom) in complex with danoprevir. Sections representing the P_4-P_3 (purple), P_2 isoindoline (green) and P_2 proline (pink dashes) are highlighted in boxes. Correlation coefficients are represented using a red–blue color scale where positive and negative correlations are represented by red and blue shades respectively. **(Right)** Surface representation of cross correlation changes relative to the P_4-P_3 (purple) and P_2 (green) moieties and boxed sections in (A).

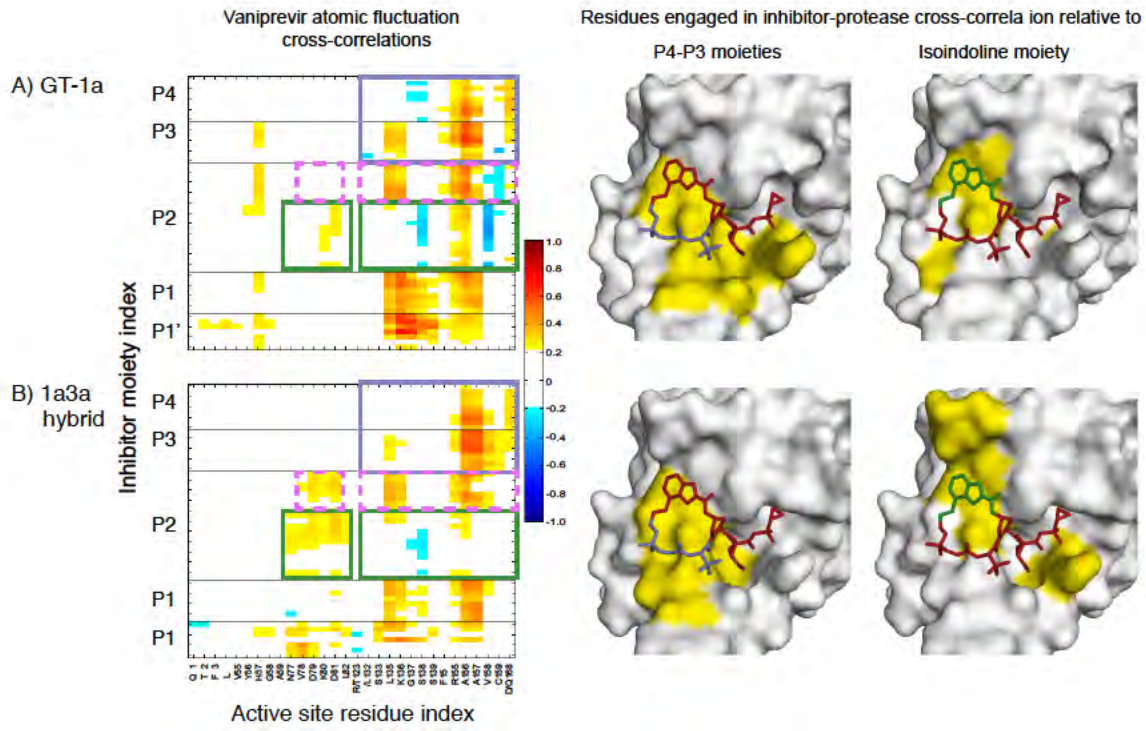


Figure 4.12: Effects of 1a3a mutations on intrinsic Protease–VAN dynamic cooperativity. **(Left)** Atomic fluctuation cross-correlation matrix of active site residues from GT-1 (top) and 1a3a (bottom) in complex with vaniprevir. Sections representing the P_4-P_3 (purple), P_2 isoindoline (green) and P_2 proline (pink dashes) are highlighted in boxes. Correlation coefficients are represented using a red–blue color scale where positive and negative correlations are represented by red and blue shades respectively. **(Right)** Surface representation of cross correlation changes relative to the P_4-P_3 (purple) and P_2 (green) moieties and boxed sections in (A).

Chapter V

Discussion

5.1 Rational Drug Design

The history of HCV antiviral development is a success story. The available HCV protease inhibitor (PI) discovery literature chronicles medicinal chemistry efforts and structure activity relationship (SAR) studies focused on identifying lead compounds with high antiviral activity and a favorable pharmacokinetic profile against GT-1 protease [46, 66, 92, 157]. While these goals were met, the rapid rise of drug resistance variants severely undermined these achievements. Sadly, incorporation of resistance associated variants (RAVs) into the lead candidate optimization process did not occur until the discovery of grazoprevir (GRZ) [79] in 2012, 8 years after the discovery of the first PI [48] and the first report of drug resistance in PIs [158]. Consequently, protease inhibitors suffer severely from drug resistance, and any treatment based on them must be supplemented with a different class of antivirals [112, 150, 159] to achieve high sustained virological response (SVR) rates.

Resistance is a serious problem that all drugs targeting rapidly evolving diseases face. A thorough characterization of the genomic and structural basis for resistance is critical for robust antiviral design and should be a priority in the current rational drug design paradigm. However, a full understanding of resistance cannot be established until elucidation of the targets substrate recognition has been completed [152, 160]. However, once this critical phase is finalized, a clean picture of the structural determinants for viral escape can be painted while incorporating inhibitor binding information.

In 2010, when I joined the Schiffer laboratory, our group had just established the molecular basis for substrate recognition in HCV NS3/4A protease, effectively defining the HCV

substrate envelope [80]. My contributions to the advancement of our field start with an extensive structural characterization of inhibitor binding modes and how they change with respect to resistance variants [68, 106] (Fig. 5.1).

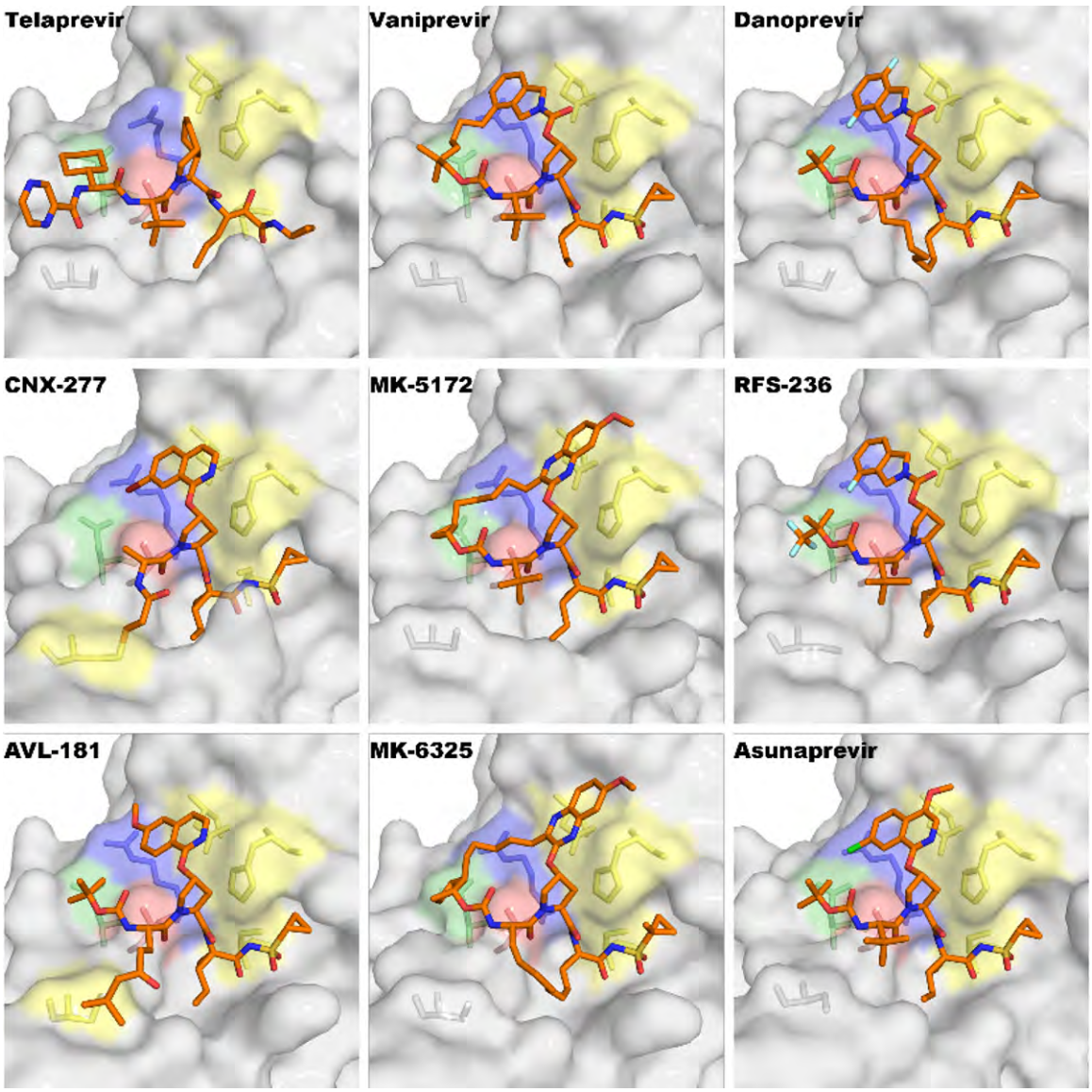


Figure 5.1: Structural characterization of HCV NS3/4A protease–inhibitor complexes. Surface representation of the HCV NS3/4A protease (white) in complex with the various inhibitor (orange sticks) studied in this thesis. Active site amino acid prone to resistance are highlighted in magenta (R155), red (A156) and green (D168); catalytic residues and Cys159 are shown in yellow.

5.2 Picking the right medicinal chemistry strategy

In HCV PI development, while structure based drug design (SBDD) has greatly influenced design efforts, strong and creative medicinal chemistry programs were the true driving force for the therapeutic advancements. Unfortunately, these two methods have not always worked hand in hand. Each novel discovery served as a building block for future generations of inhibitors. For example, the P_2 quinoline derivative was first synthesized in BILN-2061 and modified versions were synthesized in asunaprevir, sovaprevir, GS-9451, faldaprevir, and simeprevir. The main handicap with this approach is low PI diversification and possible generation of a class of ineffective compounds. When we evaluated the effect of P_2 identity on inhibitor potency by comparing three compounds, which only differed by their P_2 moiety (quinoline, isoquinoline and quinoxaline), we demonstrated that the P_2 quinoline had the worst antiviral profile of the three moieties with respect to drug resistance variants and the quinoxaline performed the best [81] (Fig. 5.2).

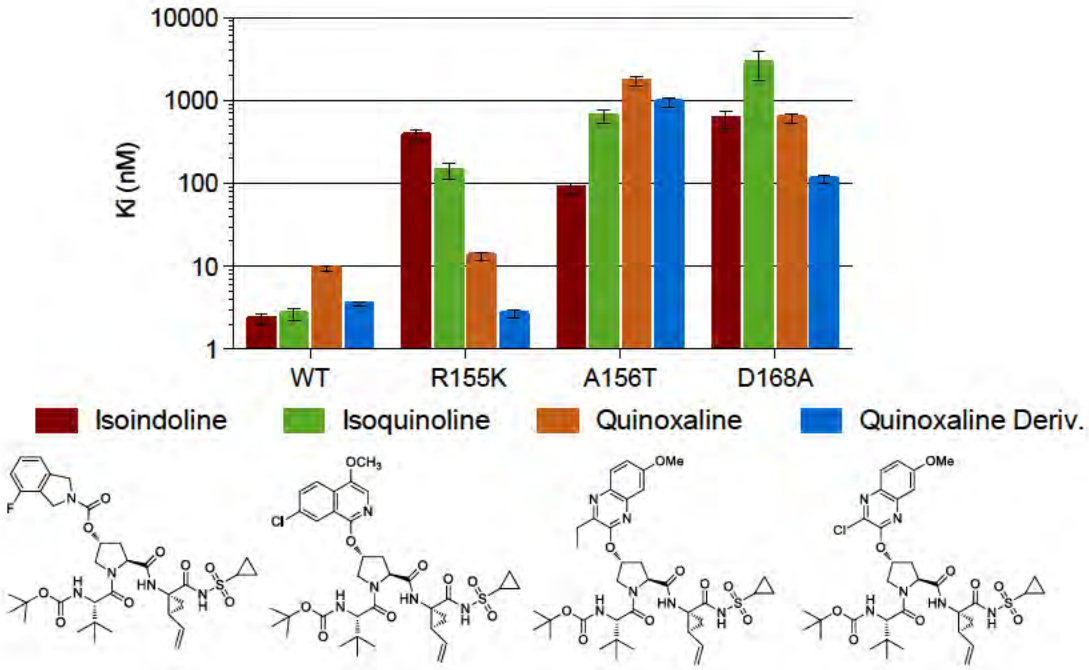


Figure 5.2: Effect of P_2 moiety on inhibitor potency and resistance profile. Structural activity relationship evaluation of the role of an inhibitor's P_2 moiety in potency and resistance profile. Isoindoline (red) isoquinoline (green) and quinoxaline based compounds were tested against the WT and major multi-drug resistance variants.

By designing quinoline based compounds with no significant alterations, various research groups have essentially created a subclass of PIs with the same drug resistance profile. Going forward, rational drug design of PIs must incorporate SBDD to create inhibitors that are designed to perform well against drug resistant variants. Our findings propose the use of a quinoxaline moiety at P_2 as a starting point for future PI development. Two main findings prompt this suggestion: (1) the quinoxaline favors a unique intermolecular interaction with catalytic residues, forming the foundation for GRZ's improved resistance barrier relative to other PIs; and (2) our crystallographic studies have shown potential chemical space for P_2 extension and enhancement. The former rationale is supported by the presence of ordered water molecules at two sites in the structures of GRZ in complex with the protease: near the P_2 6-position methoxy group on one hand and the 1-position nitrogen atom of the quinoxaline ring system on the other. One commonly used SBDD strategy is to extend an inhibitor towards the location of ordered water molecules with a functional group, which preserves water mediated intermolecular hydrogen bonding interactions [161, 162]. The recent discovery of MK-2748 illustrates this principle [163] and supports my recommendation. In that work, a SAR study was used to explore the favorability of a hydrogen bond acceptor external to GRZ's quinoxaline ring at the N-1 position. The study demonstrated that use of a carbonyl group at this N-1 position led to improved antiviral activity against known multi-drug resistance variants (MDRs) and GT-3 regardless of the P_2 heterocyclic identity.

Future PI developments efforts should explore the use of a 3-chloro-6-methoxyquinoxaline moiety at P_2 with SAR at positions 1 and 6 toward the ordered waters.

5.3 Macrocyclization and thermodynamic optimization

Macrocyclic HCV PIs have significantly improved potency and resistance profiles relative to their linear counterparts [81]. A factor contributing to these enhancements is the stabilization of the inhibitor in a binding competent conformation while allowing for sufficient flexibility to accommodate resistance mutations [164]. In GRZ, the $P_2 - P_4$ is flexible enough to accommodate amino acid changes at R155 and D168, but not A156. Conversely, the $P_1 - P_3$ macrocycle can accommodate all three amino acid changes because it is located away from the RAVs hotspot in a buried subpocket. While the choice between the two strategies might seem simple, picking the right macrocyclization strategy, one that will lead to a highly potent compound with suitable absorption, distribution, metabolism, and excretion (or ADME) profile, is not a trivial task as overall antiviral activity is influenced by the synergistic effect of all of the inhibitor moieties.

The thermodynamic forces (enthalpy and entropy) governing ligand binding affinity can be experimentally determined via isothermal titration calorimetry (ITC). In HIV, a thorough characterization of the thermodynamic signature of protease inhibitors revealed that binding of the newer and more potent HIV PIs is predominantly enthalpically driven, favoring intermolecular hydrogen bonding and van der Waals contacts over entropically driven forces (hydrophobic interactions, desolvation, and rigidity) [165]. Knowledge of a ligand's thermodynamic footprints can serve as a platform for potency enhancement through enthalpic or entropic optimization [166]. My thermodynamic characterization of GRZ's acyclic and macrocyclic analogues showed that the macrocyclic compounds were significantly more entropically driven than their acyclic counterparts. Additional thermodynamic characterization of HCV PIs needs to occur before we can know the feasibility of a thermodynamic optimization. Nonetheless, an enthalpic optimization can serve as a guide for SAR studies.

Taken together, my recommendation for future drug design efforts is to include the $P_1 - P_3$ macrocycle strategy coupled with the aforementioned 3-chloro-6-methoxyquinoxaline moiety at P_2. A SAR study guided by enthalpy optimization could yield robust inhibitors.

5.4 Targeting genotype 3 protease

A recent study, aimed at developing a pan-genotypic PI, demonstrated that enhancing the potency and pharmacokinetic profile of GRZ relative to GT-3 could be achieved [167]. In this study, the GRZ's quinoxaline based $P_2 - P_4$ moieties were substituted for a quinoline based $P_2 - P_4$. The resulting quinoline was extended at position 4 toward a conserved Asp at position 79 at the interface with the helicase domain. The hypothesis behind this strategy was that a charge complementarity to Asp79 could improve potency. The use of electrostatic complementarity as a basis for inhibitor development is an established method [168] with many successful applications, such as with human des-Gla-coagulation factor Xa [169]. The lead compound of the pan-genotypic inhibitor study contained a bicyclic piperidine-ether at the 4-position of the quinoline based P_2-P_4, a methylcyclopropyl sulfonamide at $P_{1'}$, and a cyclohexyl P_3 moiety. Relative to GRZ, these changes increased antiviral activity against GT-3 by 50 fold (0.26 vs. 13.0 nM EC_{50}) and improved pharmacokinetics in rats (0.7 vs. 1.2 AUC, uM/h). An additional significant discovery from this study is that the lead inhibitor, compound 37, is less susceptible to the A156T MDR variant than GRZ (0.42 vs. 5.3 nM). We have shown that $P_2 - P_4$ macrocyclic compounds are susceptible to the A156T mutation because the threonine clashes with the macrocycle, resulting in unfavorable binding [68]. The improved potency in compound 37 would suggest that the macrocycle in this specific instance has sufficient flexibility to accommodate mutations at 156. Future studies should attempt to structurally characterize this novel quinoline based $P_2 - P_4$ strategy.

An alternative pan-genotypic strategy worth exploring is the use of a bis-macrocycle and P_4 cyclopentyl capping similar to MK-6325 [163]. This compound is identical to GRZ with the exception of a cyclopentyl P_4, a methylcyclopropyl sulfonamide at $P_{1'}$, and the addition of a $P_1 - P_3$ macrocycle. Relative to GRZ, these modifications in MK-6325 resulted in a two-fold and four-fold improvement in potency against GT-3 (0.26 vs. 0.7 nM IC_{50}), and GT-1 WT (0.005 vs. 0.02 nM) respectively. With this antiviral activity profile, MK-6325 has reached a PI milestone, being the first single digit picomolar inhibitor in GT-1 and outcompeting GRZ with respect to major MDR variants R155K (0.07 vs. 0.013 nM), A156T (5.3 vs. 0.42 nM) and D168Y (0.14 vs. 0.036 nM) [163, 167]. My structural characterization of MK-6325 (appendix A) elucidated the mechanism by which the potency improvement occurs. While the cyclopentyl P_4 allows the $P_2 - P_4$ macrocycle sufficient flexibility to accommodate amino acid changes at residue 156, unlike the cyclopropyl ring in GRZ, the $P_1 - P_3$ macrocycle behaves like an electronic resistor and reduces the susceptibility against D168 RAVs. As mutations at D168 confer resistance to GRZ in the clinic [170] and a D168Q polymorphism accounts for a major part of the loss in potency of PIs relative to GT-3, strategies that improve antiviral activity against D168 RAV could serve a dual purpose of increasing the overall resistance barrier and augmenting pan-genotypic breadth. The 1a3a chimera cross-correlation analysis of atomic fluctuations that I describe in chapter IV could provide valuable insights for a rapid and streamlined in silico evaluation of candidate moieties.

5.5 Implications for other evolving diseases

Outside of HCV in the Flaviviridae family, flaviviruses such as Dengue virus (DENV) and West Nile Virus (WNV) infect millions of people and cause widespread morbidity and mortality [171, 172] Like HCV, DENV and WNV are positive strand RNA viruses composed of similar structural and non structural machinery [173, 174]. Their viral NS2B/3 protein is a serine protease, and like the HCV NS3/4A, it is responsible for polyprotein processing and viral maturation [174]. To date, no approved DAA against either DENV or WNV exists, underscoring the urgency for robust drug development. The DENV and WNV proteases are attractive drug targets and can benefit from the strategies developed for HCV PIs. Several reports describe the binding modes of small peptides to the DENV NS2B/3 protease [172,175,176] and an examination of DENV serotype-3 protease in complex with the Bz-nKRR-H peptide shows the C terminal Arginine 4.1 Å away form the N-terminal benzoyl group (Fig. 5.3). This proximity within the $S_1 - S_4$ subpockets favors the incorporation of a macrocyclic linkage between the $P_1 - P_4$ inhibitor moieties. Recently, the first nanomolar DENV-WNV peptide inhibitor was described (12 nM IC_{50}), however, it lacks in vivo efficacy (3.4 μM EC_{50}) [177]. Incorporating macrocycles in inhibitor design should not only improve potency but also enhance pharmacokinetic and pharmacodynamic properties [104, 110]. As the inhibitors become more advanced and are less peptide-like, design efforts should have a strong drug resistance analysis and prevention focus. Lessons learned from both HIV and HCV proteases could accelerate antiviral discovery efforts.

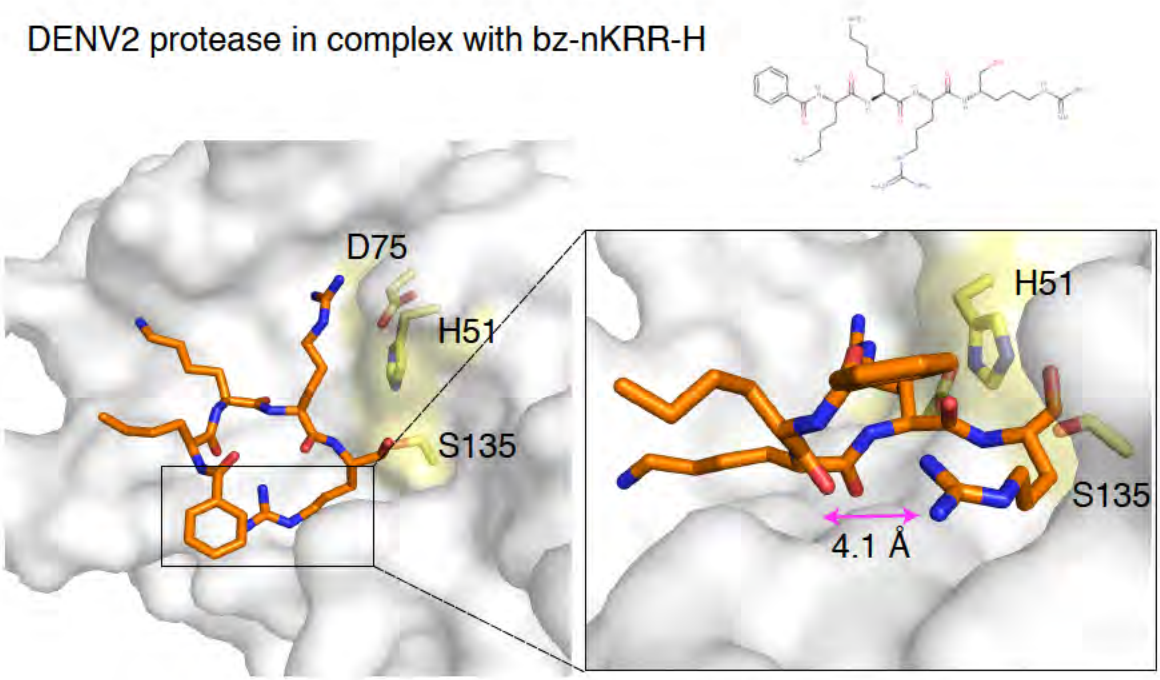


Figure 5.3: Structural representation of peptide bound DEN NS2B/3 protease. Surface representation of the DEN NS2B/3 protease (white) in complex with a bz-nKRR-H peptide inhibitor (orange sticks). The proximity of the peptide side chains within the S_1 and S_4 subpockets favors the incorporation of a macrocyclic linkage (pink) between the P_1 and P_4.

5.6 Concluding remarks

When this thesis research started, HCV infected over 180 million individuals worldwide and had no approved direct acting antiviral. Thanks to relentless efforts, close to a dozen therapeutic options currently exist for patients and their doctors. These therapies, reliant on combinations of NS5B polymerase, NS3/4A protease and/or NS5A inhibitors, have raised SVR above 95% while reducing the time of therapy to 12 weeks instead of the initial 48. Despite these significant improvements in disease treatment, efforts to develop novel robust antivirals must not slow down, as the efficacy of DAAs and access to cure rates are greatly diminished in genotype 3 infected patients. It is my sincere hope my work will inspire future scientists in their efforts to develop robust inhibitors.

Appendix A

Structural Characterization of MK-6325, a Bis-Macrocyclic Next-Generation Pangenotypic HCV NS3/4A inhibitor

1.1 Preface

Chemically, MK-6325 and GRZ share three key features the same peptidomimetic backbone, P_2 methoxy-quinoxaline, and lastly, $P_2 - P_4$ macrocycle. However, MK-6325 is different from GRZ in that it carries a methylcyclopropyl sulfonamide at $P_{1'}$, a larger P_4 capping and has $P_1 - P_3$ macrocycle. the MK-6325 is not only is first NS3/4A PI to use a bis-macrocyclic strategy but its also the first PI to achieve a single digit picomolar antiviral activity in GT-1.

To gain insight into the atomic basis for MK-6325's enhanced antiviral activity, I solved its structure in complex with WT and A156T and compared those with previously solved MK-5172, 5172-mc$P_1 - P_3$ structure. All experiments and data collection have been completed and a manuscript for submission as a communication is in preparation. The following figures highlight my findings:

Complexes/crystal	WT–MK-6325	A156T–MK-6325
Resolution (Å)	1.7	1.85
Space Group	P_{212121}	P_{212121}
Molecules in AS	1	1
Cell Dimensions		
a	34.12	36.59
b	58.7	58.12
c	81.88	79.32
β (°)	90	90
Completeness (%)	99.7	99.8
Measured Reflections	243253	60489
Unique Reflections	46302	15038
Average I/σ	16.8	20.6
Redundency	5.4	4
R-merge (%)	5.9	3.7
RMSD in		
Bonds (Å)	0.008	0.008
Angles (°)	1.614	1.64
R_{factor}	17.74	18.75
R_{free}	22.14	21.93

Table A1: Crystallographic statistics of MK-6325 in complex with Wild type and A156T drug resistant variant.

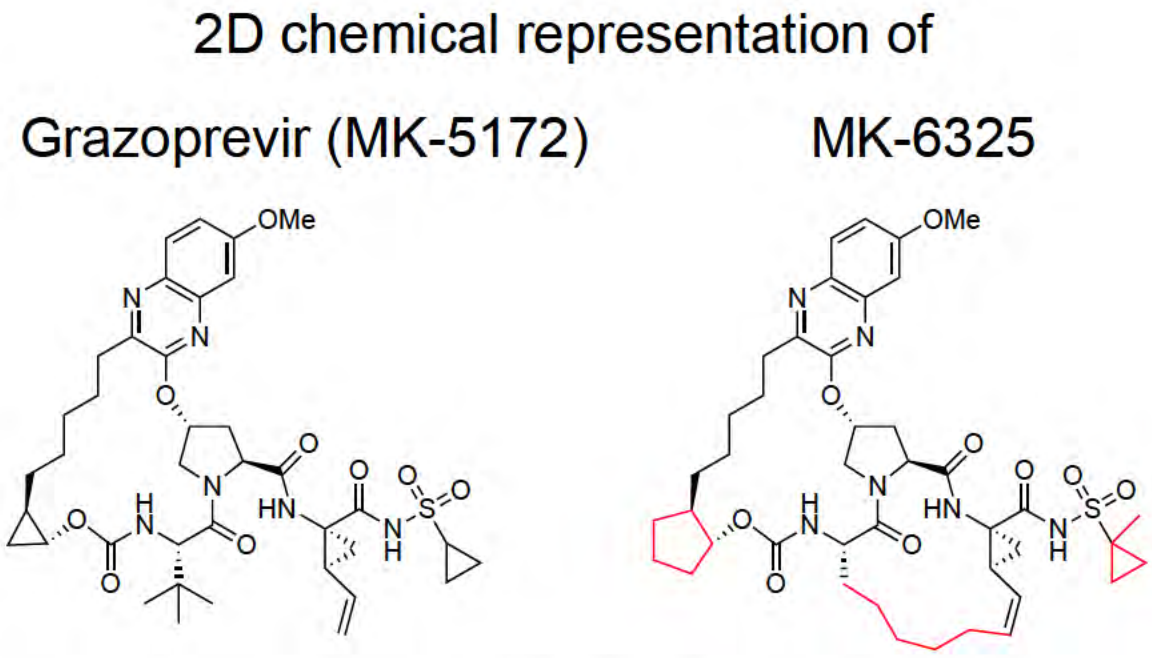


Figure A1: 2D structure of MK-6325 and Grazoprevir (MK-5172)
Similarities between MK-5172 (left) and MK-6325 (right) are shown in black, whereas dissimilarities in MK-6325 are highlighted in red.

		MK-5172		MK-6325		MK-2748	
GENOTYPE	Variant	IC50 (nM)	Fold Increase	IC50 (nM)	Fold Increase	IC50 (nM)	Fold Increase
1A	WT	0.015	1.0	0.032	1.0	0.038	1.0
1A	(R155K)	0.084	5.5	0.086	2.7	0.10	2.7
1B	WT	0.0067	1.0	0.017	1.000	0.024	1.0
1B	(A156S)	0.051	7.6	0.072	4.4	0.060	2.5
1B	(A156T)	4.3	642	0.72	44	3.7	151
1B	(A156V)	6.0	892	1.2	70	3.6	149
1B	(D168A)	0.35	52	0.086	5.2	0.093	3.8
1B	(D168E)	0.038	5.6	0.048	2.9	0.043	1.8
1B	(D168V)	0.14	21	0.089	5.4	0.057	2.3
1B	(D168Y)	0.15	23	0.088	5.3	0.11	4.5
1B	(R155K)	0.030	4.6	0.030	1.8	0.059	2.4
3A	WT	0.98		0.57		2.2	

Table A2: Antiviral activity of GRZ, MK-6325 and MK-2748

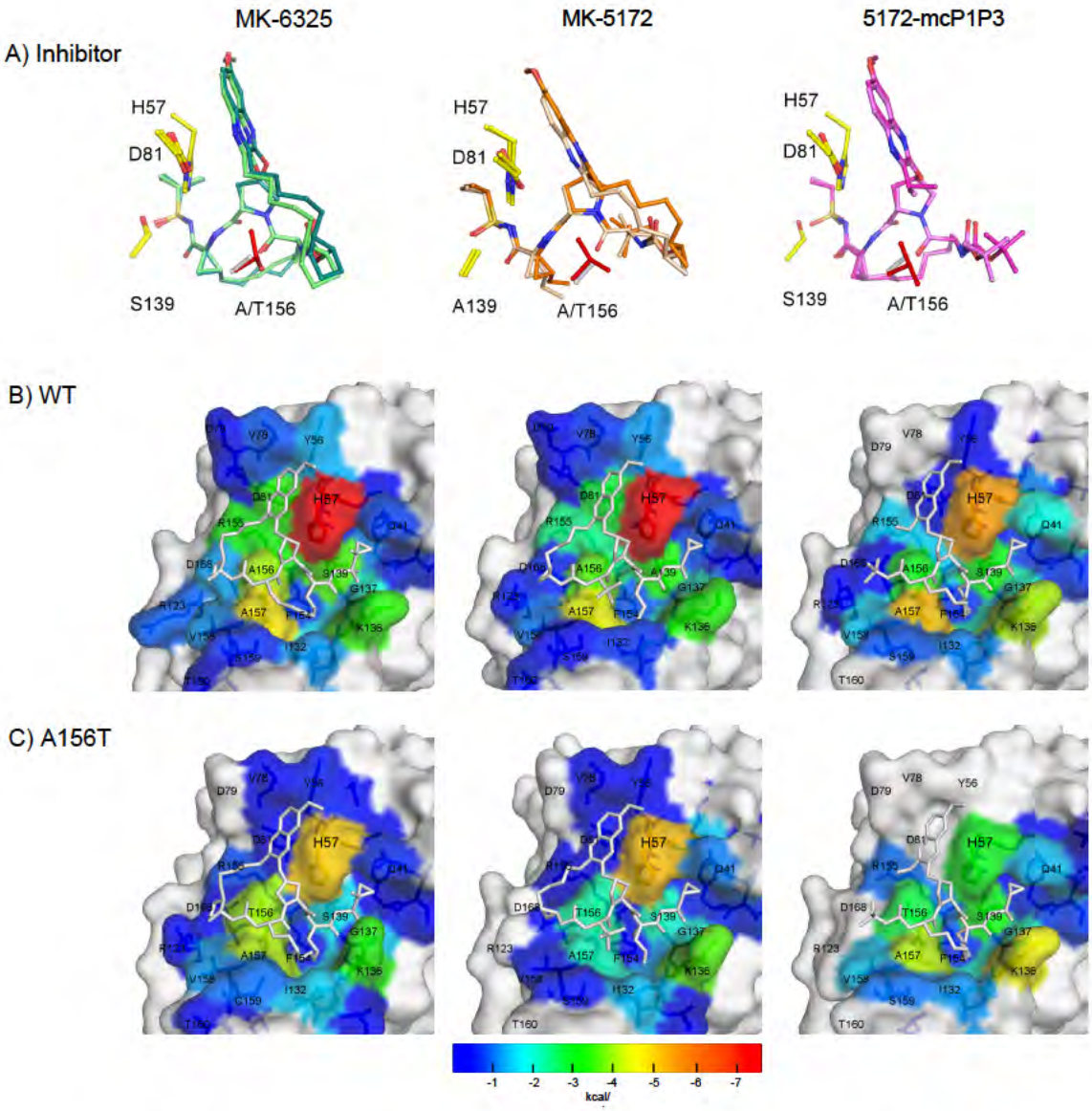


Figure A2: Surface representation of van der Waals contact energy in MK-6325 Comparative changes in the packing of the inhibitors due to A156T at the HCV NS3/4A active site evaluated by the vdW contact energies. (**A**) Superposition of the inhibitor bound to WT and A156T protease. The vdW contact energy with the inhibitor mapped onto the protease surface in (**B**) WT and (**C**) A156T variants. The warmer (red) and cooler (blue) colors indicate more and less contacts with the inhibitor, respectively.

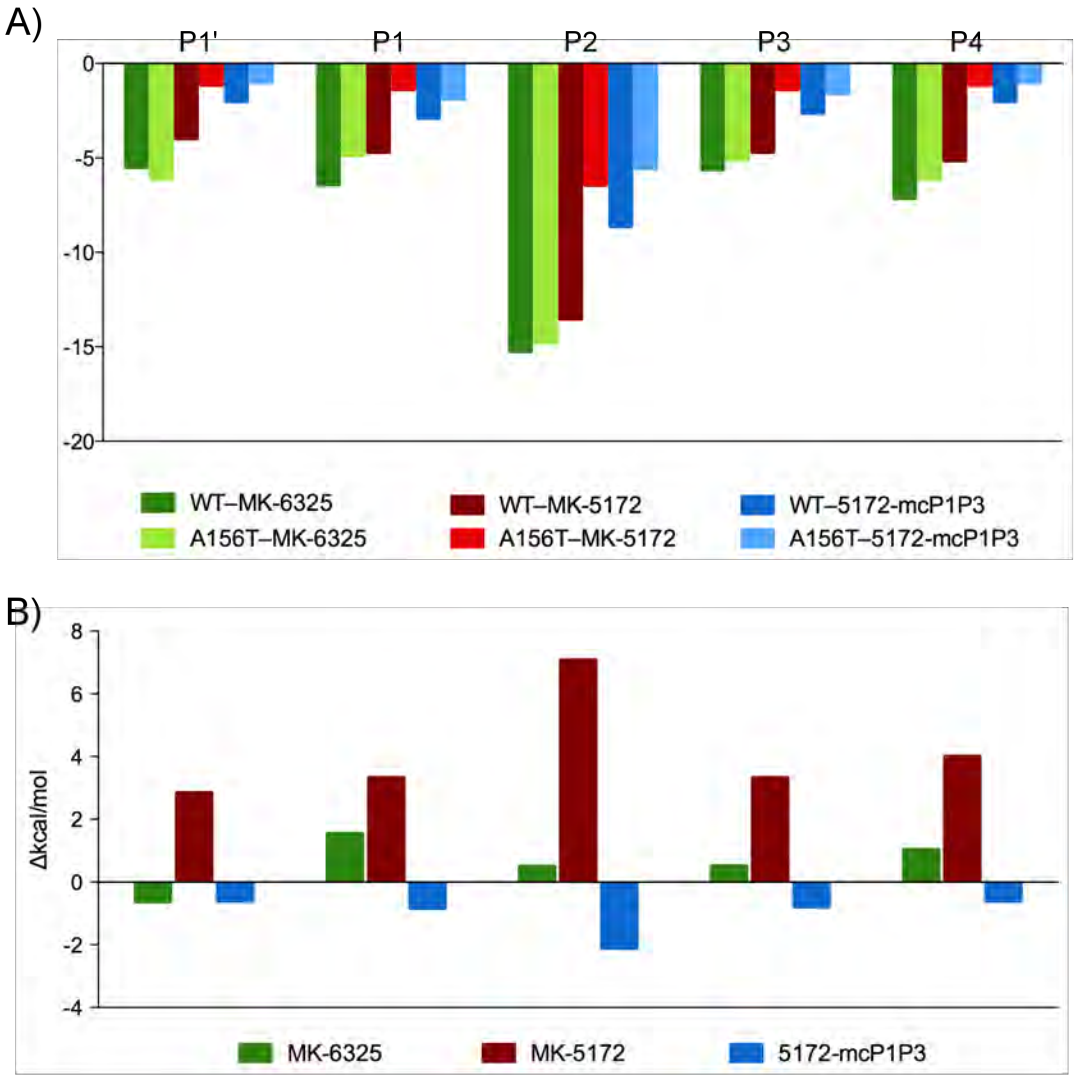


Figure A3: MK-6325's van der Waals contact energy by inhibitor moiety
(**A**) Relative van der Waals interaction of MK-6325 (green), MK-5172 (red) and 5172-mc$P_1 - P_3$ (blue) across $P_{1'} - P_4$ moieties are shown with dark and light shades representing the WT and A156T variant respectively. (**B**) changes in vdW contact potentials the A156T relative to WT.

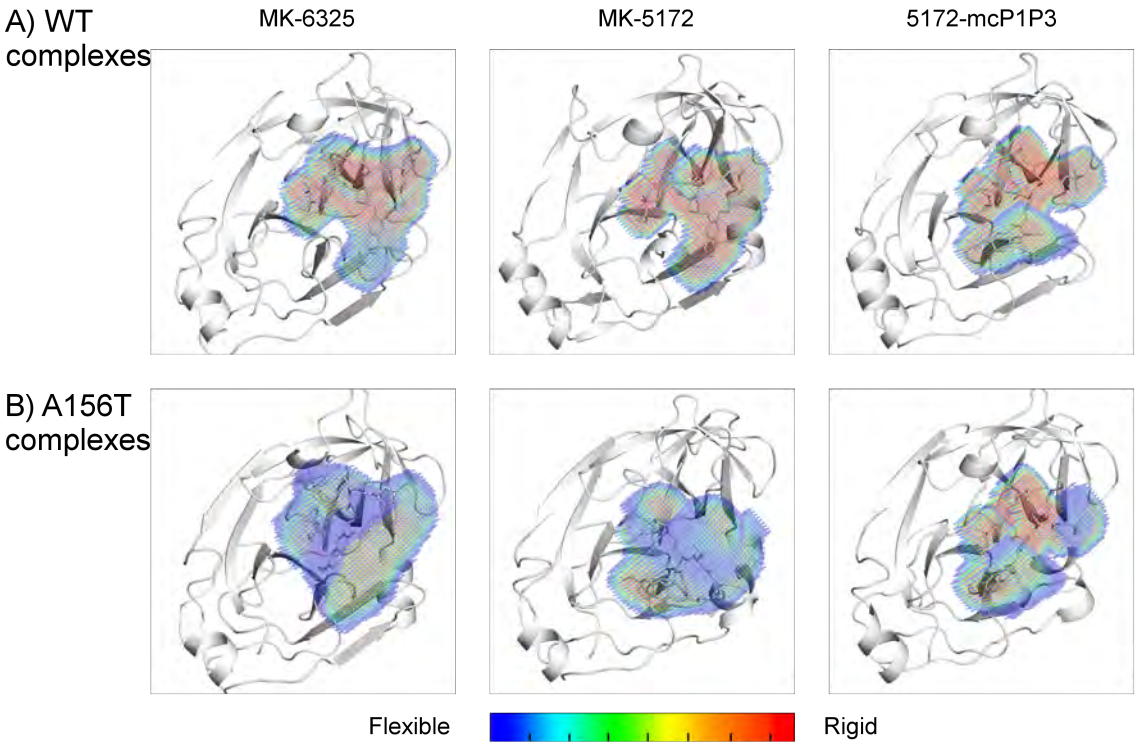


Figure A4: dynamic effects of the bis-macrocyclization on inhibitor's flexibility
The dynamic inhibitor envelope when bound to (**A**) WT and (**B**) A156T protease. Red and blue indicate less and more flexible regions of the inhibitor, respectively.

1.2 Summary of findings

- P_4 cyclopentyl ring engages in more hydrophobic interactions that the cyclopentyl capping potentially underlying potency enhancements in D168 resistance associated variants.

- P_4 cyclopentyl ring increase the inhibitor's dynamic flexibility, allowing it to accommodate the A156T mutation.

Appendix B

Structural Characterization of RFS-256

2.1 Preface

This project was a collaboration between the Schiffer laboratory at UMass Medical School and Schinazi Laboratory at RFS Pharma.

The goal was to investigate the structural binding mode of RFS-256, a linear analog of Danoprevir but with trifluoromethyl at P_4.

2.2 Methods and Results

Complexes	WT-RFS-256
Resolution (Å)	2
Space Group	P_{21}
Molecules in AU	2
Cell Dimensions	
a	55.2
b	58.5
c	60.1
β (°)	90
Completeness (%)	98
Measured Reflections	99451
Unique Reflections	2389
Average I/σ	11.8
Redundency	4.1
R-sym (%)	5
RMSD in	
Bonds (Å)	0.013
Angles (°)	1.7
R_{factor}	14.5
R_{free}	19.9

Table B1: Crystallographic statistics of RFS-256.

2.2.1 Conserved active site packing

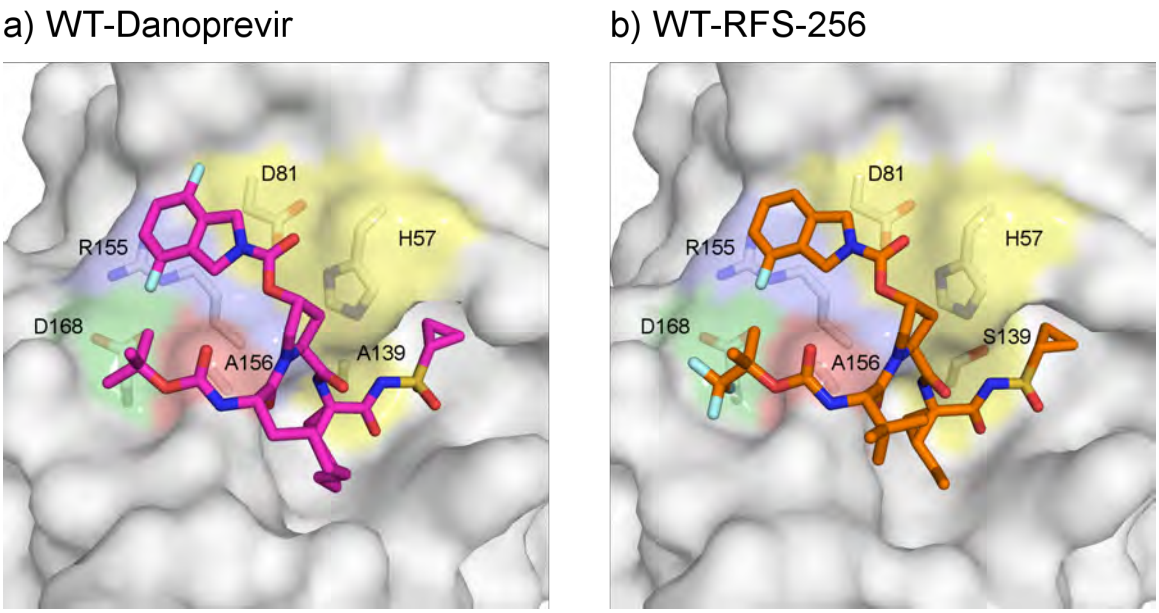


Figure B1: Surface representation of RFS-256 and Danoprevir
Surface representation of WT-NS3/4A in complex with **(A)** Danoprevir (magenta) and **(B)** RFS-256 (orange) with key active site residues highlighted: catalytic triad (H57, D81 and S139) in yellow and sites prone to multidrug resistance mutations R155, D168 and A156 shown in blue, green and red, respectively.

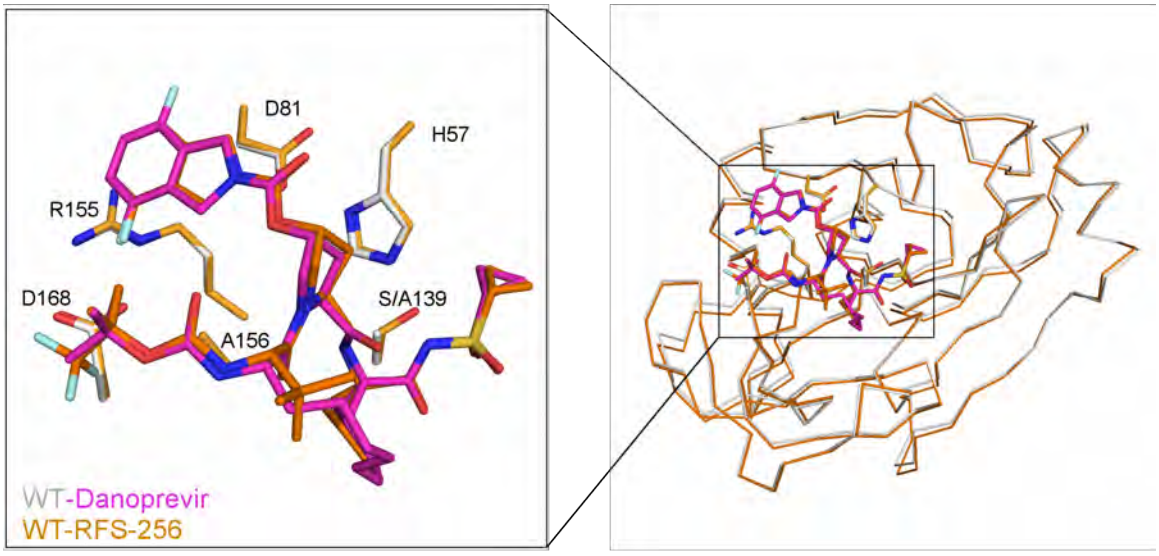


Figure B2: Structural overlay of RFS-256 and Danoprevir
Cα structural superposition of WT–Danoprevir (magenta) and WT–RFS-256 (orange) reveals a 0.2 Å root mean square deviation (RMSD).

2.2.2 RFS-256 is characterized by similar vdW contact landscape as Danoprevir

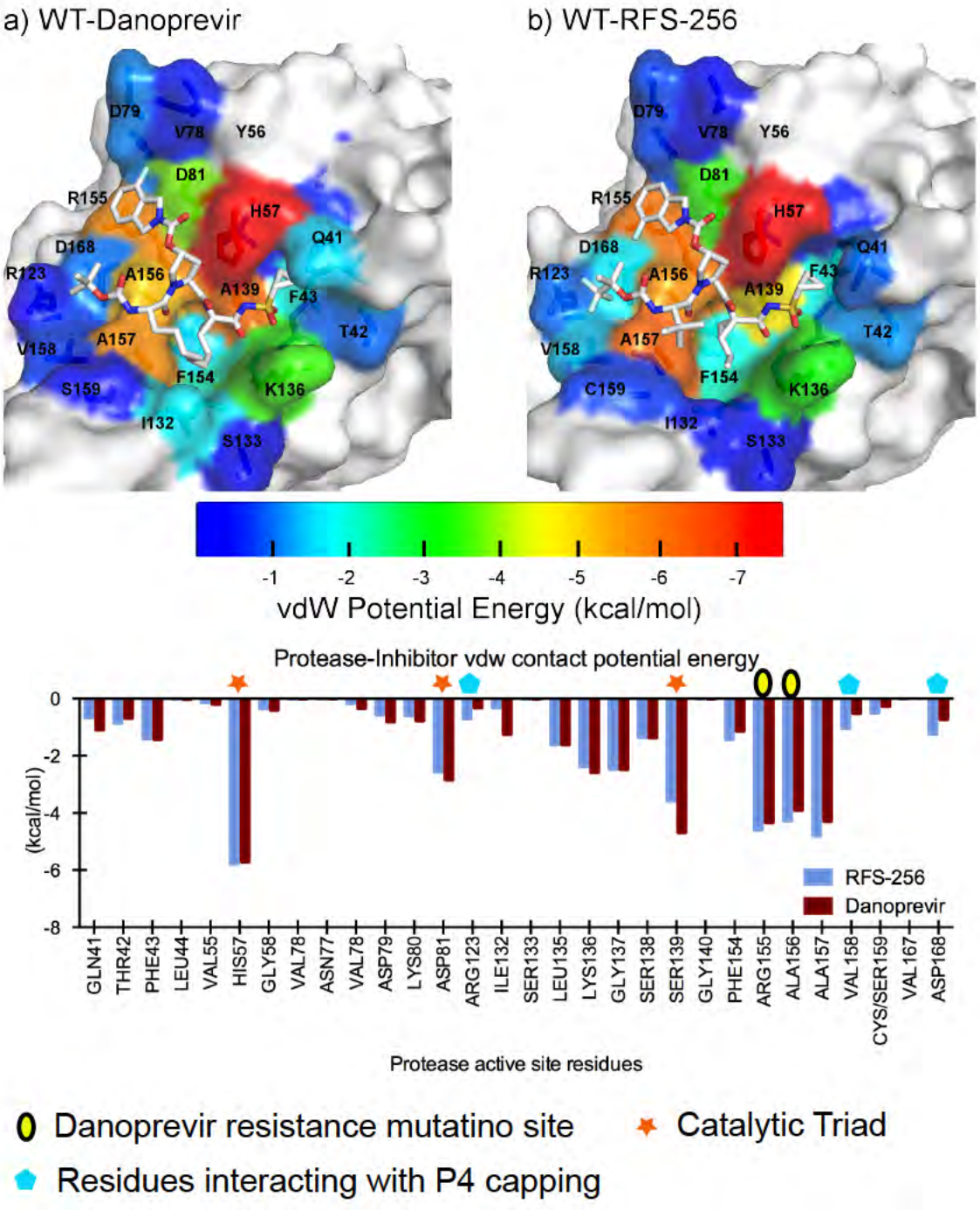


Figure B3: RFS-256 van der Waals contact energy landscape
Surface representation of protease–inhibitor van der Waals contact energy (top) for **(A)** WT–Danoprevir and **(B)** WT–RFS-256, with potential energies of individual residues shown (bottom).

Key findings:

1. Catalytic residues D81 and S139 in WT-RFS-256 lose contacts relative to WT-Danoprevir;
2. Danoprevir associated resistance sites, 155 and 168 gain contacts in WT-RFS-256;
3. D168, R123 and V158 exhibit altered vdW potential due to RFS-256's P_4 capping group.

Appendix C

Secondary Contribution 1: Molecular Basis for Drug Resistance in HCV

3.1 Secondary Contribution

Author contributions: **DS** expressed, purified, crystallized various protease–inhibitor complexes in this study. Additionally, **DS** participated in data collection, analysis as well as manuscript preparation.

OPEN ACCESS Freely available online PLoS PATHOGENS

The Molecular Basis of Drug Resistance against Hepatitis C Virus NS3/4A Protease Inhibitors

Keith P. Romano[1], Akbar Ali[1], Cihan Aydin[1], Djade Soumana[1], Ayşegül Özen[1], Laura M. Deveau[1], Casey Silver[1], Hong Cao[1], Alicia Newton[2], Christos J. Petropoulos[2], Wei Huang[2], Celia A. Schiffer[1]*

1 Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, United States of America, **2** Monogram Biosciences, San Francisco, California, United States of America

Abstract

Hepatitis C virus (HCV) infects over 170 million people worldwide and is the leading cause of chronic liver diseases, including cirrhosis, liver failure, and liver cancer. Available antiviral therapies cause severe side effects and are effective only for a subset of patients, though treatment outcomes have recently been improved by the combination therapy now including boceprevir and telaprevir, which inhibit the viral NS3/4A protease. Despite extensive efforts to develop more potent next generation protease inhibitors, however, the long term efficacy of this drug class is challenged by the rapid emergence of resistance. Single site mutations at protease residues R155, A156 and D168 confer resistance to nearly all inhibitors in clinical development. Thus, developing the next generation of drugs that retain activity against a broader spectrum of resistant viral variants requires a comprehensive understanding of the molecular basis of drug resistance. In this study, 16 high resolution crystal structures of four representative protease inhibitors – telaprevir, danoprevir, vaniprevir and MK 5172 – in complex with the wild type protease and three major drug resistant variants R155K, A156T and D168A, reveal unique molecular underpinnings of resistance to each drug. The drugs exhibit differential susceptibilities to these protease variants in both enzymatic and antiviral assays. Telaprevir, danoprevir and vaniprevir interact directly with sites that confer resistance upon mutation, while MK 5172 interacts in a unique conformation with the catalytic triad. This novel mode of MK 5172 binding explains its retained potency against two multi drug resistant variants, R155K and D168A. These findings define the molecular basis of HCV N3/4A protease inhibitor resistance and provide potential strategies for designing robust therapies against this rapidly evolving virus.

Citation: Romano KP, Ali A, Aydin C, Soumana D, Ozen A, et al. (2012) The Molecular Basis of Drug Resistance against Hepatitis C Virus NS3/4A Protease Inhibitors. PLoS Pathog 8(7): e1002832. doi:10.1371/journal.ppat.1002832

Editor: Andrea Gamarnik, Fundacion Instituto Leloir CONICET, Argentina

Received April 9, 2012; **Accepted** June 13, 2012; **Published** July 26, 2012

Funding: This work was supported by the National Institute of Health grants R01 GM65347 and R01 AI085051 The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Alicia Newton, Christos J. Petropoulos, and Wei Huang are employees of Monogram Biosciences. From roughly 2008 August 2010 Celia Schiffer served as a consultant to Monogram Biosciences, on topics unrelated to this research and thus has no competing interest. Since the completion of this research project Celia Schiffer has received small "Sponsored Research Grants" from Avila (CelGene) and Merck. This does not alter our adherence to all PLoS Pathogens policies on sharing data and materials.

* E mail: Celia.Schiffer@umassmed.edu

Introduction

Hepatitis C virus (HCV) is a genetically diverse positive-stranded RNA virus of the *Flaviviridae* family infecting an estimated 170 million people worldwide [1,2]. Based on genetic diversity, HCV is divided into six major genotypes (genotypes 1 6) and numerous subtypes with different geographic distributions; genotypes 1 and 3 are the most prevalent worldwide [3]. HCV infection is the leading cause of chronic liver disease that persists for decades and eventually progresses to cirrhosis, liver failure, or liver cancer [4]. The current anti-HCV standard of care is a combination of pegylated interferon (Peg-IFN), ribavirin (RBV), and boceprevir or telaprevir, two recently approved antiviral agents targeting the viral NS3/4A protease [5]. Sustained virologic response (SVR) which is tantamount to cure is achieved only in a subset of treated patients, depending on a combination of viral and host-cell genetic factors [6 10]. For example, a human polymorphism at the IL28B gene is associated with poor interferon response [11]. Most patients undergoing interferon-based therapies also experience significant adverse effects, including flu-like symptoms, anemia, and depression [12]. Thus, current anti-HCV therapies are often not tolerated and ineffective for many patients, and novel direct-acting antiviral drugs are required for safer, more efficacious treatment.

Direct-acting antiviral agents have the potential to improve SVR rates and minimize treatment duration. The HCV NS3/4A protease a chymotrypsin-like serine protease is a prime therapeutic target that cleaves four known sites along the virally encoded polyprotein [13]. The NS3/4A protease also hydrolyzes two human proteins, TRIF and MAVS, which are part of the innate immune system, thereby confounding the innate immune response to viral infection [14,15]. Pharmaceutical companies have invested significant effort in developing NS3/4A protease inhibitors. Proof-of-concept of antiviral efficacy was first demonstrated in 2002 with the macrocyclic inhibitor BILN-2061 (ciluprevir) [16,17], which was later discontinued due to concerns about its cardiotoxicity [18]. As noted above, boceprevir [19] and telaprevir [20,21] are two NS3/4A protease inhibitors recently approved by the Food and Drug Administration, marking an important milestone in anti-HCV research and drug development

Author Summary

Hepatitis C virus (HCV) infects over 170 million people worldwide and is the leading cause of chronic liver diseases, including cirrhosis, liver failure, and liver cancer. New classes of directly acting antiviral agents that target various HCV enzymes are being developed. Two such drugs that target the essential HCV NS3/4A protease are approved by the FDA and several others are at various stages of clinical development. These drugs, when used in combination with pegylated interferon and ribavirin, significantly improve treatment outcomes. However HCV evolves very quickly and drug resistance develops against directly acting antivi ral agents. Thus, despite the therapeutic success of NS3/4A protease inhibitors, their long term effectiveness is chal lenged by drug resistance. Our study explains in atomic detail how and why drug resistance occurs for four chemically representative protease inhibitors –telaprevir, danoprevir, vaniprevir and MK 5172. Potentially with this knowledge, new drugs could be developed that are less susceptible to drug resistance. More generally, understand ing the underlying mechanisms by which drug resistance occurs can be incorporated in drug development to many quickly evolving diseases.

over the past two decades. Both boceprevir and telaprevir are linear ketoamide compounds that form a reversible, covalent bond with the catalytic serine of NS3/4A protease. Several non-covalent xprotease inhibitors have also advanced into human clinical trials; these inhibitors include both linear (BMS-650032 [22], BI 201335 [23]) and macrocyclic compounds, containing either a P1 P3 (danoprevir [24], TMC435 [25]) or a P2 P4 (vaniprevir [26], MK-5172 [27]) macrocycle (Figure 1).

The NS3/4A protease inhibitors rapidly reduce HCV RNA titers when administered as monotherapy [17,28 31] and substantially improve SVR rates when given in combination with Peg-IFN and RBV [6 10,32 34]. However, the high rate of HCV replication and poor fidelity of HCV's RNA-dependent RNA polymerase lead to heterogeneous virus populations in infected patients [35,36]. These viral quasispecies exist at low levels in untreated patients, and resistant populations emerge under the selective pressure of direct-acting antiviral agents [36 38]. In the majority of patients undergoing protease inhibitor therapy, resistance develops rapidly due to overlapping but distinct sets of NS3/4A mutations [37]. In patients with genotype 1a, the R155K mutation causes resistance against nearly all inhibitors, but rarely occurs in genotype 1b patients [29,30,32,37 42]. Instead, distinct resistance mutations arise in genotype 1b patients depending on the class of protease inhibitor used; A156 mutates in response to treatment with linear ketoamide protease inhibitors [39 41], while macrocyclic inhibitors more commonly select for D168A and R155K variants [29,30,32,42]. Mutations at V36, T54, and V36+A155 are also associated with resistance to ketoamide inhibitors [39 41]. Variations in the patterns of resistant mutations arise from the complex interplay between genotype, replication rates, mutation rates, and the resulting effect of mutations on viral fitness and drug potency. Clearly, despite the benefits of combination therapy in improving SVR rates, the emergence of resistance challenges the long-term efficacy of NS3/4A protease inhibitors.

Most primary drug-resistance mutations in NS3/4A protease occur around the active site in regions where drugs protrude from the substrate binding space, defined as the substrate envelope, because these changes can preferentially disrupt drug binding with minimal effect on substrate binding and viral fitness [43]. The protease inhibitors danoprevir, TMC435, and boceprevir protrude from the substrate envelope in regions that correlate with known sites of resistance mutations. Notably, the large P2 moieties of danoprevir and TMC435 bind in the S2 subsite and extensively interact with residues R155, D168, and A156 [43], which mutate to confer multi-drug resistance [37,38,44]. These and other inhibitors with large P2 moieties derive much of their potency from binding in the S2 subsite [45], but how molecular changes at these residues selectively weaken inhibitor binding without compromising the binding of viral substrates is not clear. Elucidating the underlying molecular mechanisms of NS3/4A protease inhibitor resistance is therefore essential for developing new drugs that are less susceptible to resistance.

How single-site mutations at residues R155, A156 and D168 confer resistance against most protease inhibitors has not been elucidated in atomic detail. In this study, we report that four chemically representative protease inhibitors telaprevir, danoprevir, vaniprevir and MK-5172 exhibit distinct susceptibilities to the protease variants R155K, A156T and D168A (Table 1). Sixteen high-resolution crystal structures of inhibitors in complex-with the wild-type protease and three drug resistant variants reveal the molecular basis underlying the unique resistance profiles of these inhibitors (Table 2). The P2 quinoxaline moiety of MK-5172 stacks against the protease catalytic triad in a novel conformation, explaining its retained potency against R155K and D168A. The flexible P2 isoindoline moiety of danoprevir containing a P1 P3 macrocycle packs against the mutated surfaces of A156T and D168A variants, explaining its relatively higher activity against both protease variants. However, the isoindoline moiety in vaniprevir is constrained due to the P2 P4 macrocycle, resulting in significantly lower activity against all three variants. Thus, incorporating either quinoxaline or flexible substituents at the P2 proline confers clear advantages. Taken together, these data highlight potential strategies for designing novel drugs that retain potency against a broader spectrum of resistant viral variants.

Results

Drug susceptibility assays

Drug activities were determined for telaprevir, danoprevir, vaniprevir and MK-5172 against wild-type genotype 1a HCV and resistant variants R155K, D168A, and A156T using viral replicon-based inhibition assays. The antiviral activities against the resistant variants trended with changes in binding affinities measured in enzyme inhibition assays (Table 1). Against wild-type protease, macrocyclic inhibitors danoprevir, vaniprevir and MK-5172 exhibited antiviral potencies in the sub nM range (IC_{50} 0.24, 0.34 and 0.11 nM, respectively), while telaprevir potency was significantly lower (IC_{50} 1030 nM), consistent with previous reports [46,47]. Relative to the wild type, R155K caused large reductions in potency for danoprevir and vaniprevir, but MK-5172 remained highly active (R155K IC_{50} 0.55 nM). Telaprevir potency was slightly better against D168A relative to the wild type, while danoprevir, vaniprevir and MK-5172 lost 100- to 1000-fold potency against D168A. However, both danoprevir and MK-5172 still were significantly more potent than telaprevir against D168A. Among the macrocyclic drugs, danoprevir and MK-5172 retained higher activities against D168A (D168A IC_{50} 48 nM and 13 nM, respectively) relative to vaniprevir (D168A IC_{50}>400 nM). Danoprevir also retained significantly higher potency against A156T (A156T IC_{50} 5.7 nM), while the other three drugs incurred large-fold

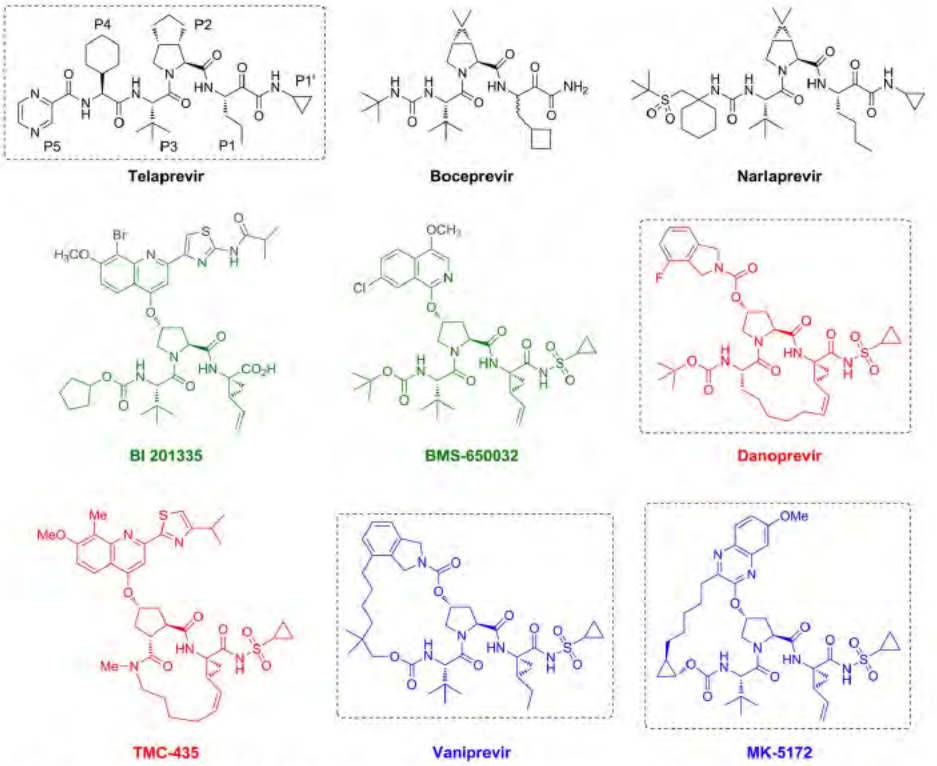


Figure 1. The chemical structures of NS3/4A protease inhibitors. The canonical nomenclature for drug moiety positioning is indicated using telaprevir. Telaprevir (black), danoprevir (red), vaniprevir and MK-5172 (blue) are representative of many other protease inhibitors in development. Telaprevir, recently approved for clinical use, is an acyclic ketoamide inhibitor that forms a reversible, covalent bond with the protease. Danoprevir, currently in phase II clinical trials, is a non-covalent acylsulfonamide inhibitor with a P1–P3 macrocycle. Vaniprevir and MK-5172 are also non-covalent acylsulfonamide inhibitors, but contain P2–P4 macrocycles. Vaniprevir and MK-5172 differ in the construction of their P2 moieties: vaniprevir contains a carbamate linkage between the P2 proline and the isoindoline moiety, whereas MK-5172 contains a shorter ether linkage between its P2 proline and the quinoxaline moiety.
doi:10.1371/journal.ppat.1002832.g001

Table 1. Drug susceptibilities against wild type and resistant HCV clones and inhibitory activities against NS3/4A proteases.

	Replicon - IC_{50} (nM)[a]				Binding - K_i (nM)[a]			
Drug	**WT**	**R155K**	**D168A**	**A156T**	**WT**	**R155K**	**D168A**	**A156T**
Telaprevir (VX 950)	1030	5300 **(5.1)**	420 **(0.4)**	>50,000 **(>49)**	34.4±3.0	823±60 **(24)**	12.2±0.9 **(0.35)**	>10000 **(>291)**
Danoprevir (ITMN 191)	0.24	>100 **(>416)**	48 **(200)**	5.7 **(24)**	1.0±0.1	162±16 **(162)**	208±66 **(208)**	44.8±3.6 **(45)**
Vaniprevir (MK 7009)	0.34	>400 **(>1176)**	>400 **(>1176)**	176 **(518)**	0.74±0.07	554±64 **(749)**	2635±702 **(3561)**	958±162 **(1295)**
MK 5172	0.11	0.55 **(5)**	13 **(118)**	108 **(982)**	0.14±0.02	0.84±0.05 **(6)**	27.8±12.1 **(199)**	620±71 **(4429)**

[a]Numbers in parentheses reflect fold change relative to wild type; >indicates IC_{50} and *Ki* values higher than the maximum drug concentration tested in the assay.
doi:10.1371/journal.ppat.1002832.t001

Table 2. X-ray data collection and crystallographic refinement statistics.

Drug	Telaprevir				Danoprevir				Vaniprevir				MK-5172			
Protease variant	WT	R155K	D168A	A156T	WT[i]	R155K[i]	D168A[i]	A156T[i]	WT[i]	R155K[i]	D168A[i]	A156T[i]	WT[t]	R155K[t]	D168A[t]	A156T[i]
PDB ID	3SV6	3SV7	3SV8	3SV9	3M5L	3SU0	3SU1	3SU2	3SU3	3SU4	3SU5	3SU6	3SUD	3SUE	3SUF	3SUG
Resolution (Å)	1.40	1.55	2.50	1.60	1.25	1.16	1.40	1.50	1.30	2.25	1.55	1.10	1.96	2.20	2.30	1.80
Space group	$P2_12_12_1$	$P2_12_12_1$	$P4_12_12$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P6_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$P2_1$	$P2_1$	$P2_12_12_1$
Twin Law	-	-	-	-	-	-	-	-	-	-	-	-	-h, -k, h+l	-h, -k, h+l	-h, -k, h+l	-
Twin Fraction	-	-	-	-	-	-	-	-	-	-	-	-	0.42	0.42	0.43	-
Molecules in AU[b]	1	1	1	1	1	1	1	1	1	2	1	1	4	4	4	1
Cell dimensions:																
a (Å) =	55.1	55.3	69.5	54.8	55.2	55.3	55.0	54.9	55.1	85.8	55.1	55.0	56.1	56.3	56.0	53.9
b (Å) =	58.8	58.8	69.5	58.7	58.7	58.5	58.5	58.5	58.5	85.8	58.8	58.5	102.7	103.3	103.6	58.2
c (Å) =	60.3	60.4	79.1	60.7	61.1	60.6	60.0	60.0	60.3	97.4	60.0	59.8	73.3	73.5	73.5	62.0
β (°) =	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0*	90.0	90.0	112.5	112.6	112.0	90.0
Completeness (%)	99.9	99.6	99.8	99.8	99.7	97.9	100.0	100.0	95.7	99.9	95.8	96.6	91.6	91.6	95.9	90.1
Measured reflections	39340	183407	91636	147126	317137	394396	233176	228689	196059	152278	190711	400120	101245	77906	107542	62460
Unique reflections	218617	29054	7137	26564	52645	67429	38857	31822	46251	19213	27508	76300	50198	36194	37671	16890
Average I/σ^a	13.8	11.1	7.5	12.0	20.7	11.2	9.8	10.0	12.0	7.2	8.8	12.1	10.0	9.5	9.4	19.0
	(3.6)	(3.9)	(7.5)	(3.5)	(3.8)	(2.8)	(3.8)	(4.2)	(3.6)	(4.9)	(3.7)	(2.7)	(3.6)	(2.4)	(2.6)	(3.0)
Redundancy	5.6	6.3	12.8	5.5	6.0	5.8	6.0	7.2	4.2	7.9	6.9	5.2	2.0	2.2	2.9	3.7
R_{sym} (%)[a,c]	3.5	6.1	11.0	4.5	4.6	4.1	4.3	4.3	4.1	7.7	6.9	4.4	6.0	6.7	6.3	2.8
	(41.2)	(37.6)	(58.8)	(40.0)	(31.5)	(39.9)	(45.2)	(45.3)	(29.6)	(37.9)	(44.0)	(28.3)	(19.5)	(36.2)	(38.7)	(30.9)
RMSD[d] in:																
Bonds (Å) =	0.009	0.009	0.012	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
Angles (°) =	1.32	1.30	1.47	1.31	1.48	1.32	1.37	1.40	1.44	1.26	1.39	1.36	1.33	1.31	1.36	1.40
R_{factor} (%)[e] =	15.8	16.5	20.5	17.2	15.0	15.3	15.7	15.4	16.4	16.9	16.5	15.0	18.3	18.5	19.7	19.4
R_{free} (%)[e] =	17.1	19.9	28.3	19.7	16.8	17.2	17.8	17.7	18.2	22.3	18.3	16.5	23.8	22.7	25.6	23.0

[i]S139A protease mutant used for crystallization.
[t]pseudo-merohedral twin.
[a]values in parentheses are for the highest resolution shell.
[b]AU, asymmetric unit.
[c]$R_{sym} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I = observed intensity, $\langle I \rangle$ = average intensity over symmetry equivalent.
[d]RMSD, root mean square deviation.
[e]$R_{work} = \Sigma \; ||F_o| - |F_c|| \; / \Sigma |F_o|$. R_{free} was calculated from 5% of reflections, chosen randomly, which were omitted from the refinement process.
[*]$\gamma = 120.0$.

doi:10.1371/journal.ppat.1002832.t002

losses in potency. Notably, MK-5172, though active against the other two variants, lost significant potency against A156T (A156T IC_{50} 108 nM). Thus, the four drugs exhibited varied susceptibilities to protease inhibitor-resistant viral variants R155K, D168A and A156T.

Structure determination and analyses

To elucidate the underlying mechanism by which chemically diverse inhibitors bind to the wild-type protease and drug-resistant variants, crystal structures were determined for 16 inhibitor-protease complexes. These complexes include wild-type protease and resistant variants R155K, D168A and A156T each bound to telaprevir, danoprevir, vaniprevir and MK-5172, with resolutions ranging from 1.10 2.50 Å (Table 2); S139A protease variants were used except for telaprevir, which requires covalent bond formation with the serine 139 for efficient binding. These high-resolution data sets afforded very detailed structural interpretations of drug-protease binding.

The binding conformations of telaprevir, danoprevir, vaniprevir and MK-5172 to the wild-type protease are shown in Figure 2 and Figure S1. In all complexes, inhibitors formed three common hydrogen bonds with the protease backbone (Table S1): (1) the P1 amide nitrogen with the carbonyl oxygen of R155, (2) the P3 carbonyl oxygen with the amide nitrogen of A157, and (3) the P3 amide nitrogen with the carbonyl oxygen of A157 (Figures 3A 6A). The P5 amide nitrogen of telaprevir formed an additional hydrogen bond with the carbonyl oxygen of S159. In the telaprevir complex, the catalytic serine (S139) was covalently bound to the C-α carbon of the ketoamide warhead. The ketoamide oxygen sat in the oxyanion hole and interacted with the backbone amide nitrogens of protease residues 137 139, while the Nε nitrogen of H57 hydrogen bonded with the keto oxygen. The acylsulfonamide groups of danoprevir, vaniprevir and MK-5172 were also positioned in the oxyanion hole, hydrogen bonding with the same set of backbone amide nitrogens, as observed previously for the TMC435 and danoprevir structures [43,45]. Meanwhile the Nε nitrogen of H57 interacted with the sulfonamide nitrogen in these complexes, suggesting that the Nε atoms were deprotonated. Thus, many of these classes of inhibitors overlap in several key interactions with the protease.

In wild-type complexes involving macrocyclic inhibitors, R155 adopted a conformation distinct from those observed in telaprevir and substrate complexes to allow binding of the extended P2 moieties in the S2 subsite. This R155 conformation is stabilized by hydrogen bond interactions involving D168 and D80. The conformation has also been observed for protease in complex with TMC435 and danoprevir, where large P2 moieties of inhibitors are positioned over the guanidine side chain, making extensive cation-π stacking interactions [43,45]. Vaniprevir, with the P2 isoindoline moiety, bound in a conformation similar to danoprevir, making favorable cation-π stacking interactions with R155, despite the P2 P4 macrocycle. In contrast, MK5172 adopted a novel conformation with the ether-linked P2 quinoxaline moiety not interacting extensively with R155 and D168, but stacking instead against H57 and D81 of the catalytic triad (Figure 2). Thus, the P2 moieties of these three macrocycles pack in a variety of conformations around the active site.

To characterize binding patterns of the drugs relative to natural substrates, the wild-type drug complexes were analyzed with respect to the substrate envelope, the consensus binding volume of the substrates [43] (Figures 3A 6A). Inhibitors are generally more vulnerable to resistance where they protrude beyond the substrate envelope and contact residues less essential for substrate recognition and turnover. All four drugs protruded from the substrate envelope in the protease S2 subsite near residues R155, A156 and D168, which individually mutate to confer multi-drug resistance [37,38,44]. Telaprevir, with the small P2 cyclopentylproline moiety, made fewer van der Waals contacts with R155, A156 and D168 relative to danoprevir and vaniprevir, which contain the carbamate-linked P2 isoindoline moieties that protruded from the substrate envelope and made extensive van der Waals contacts with these residues (Figures 3A 5A). Danoprevir's isoindoline moiety bound in two conformations in the wild-type complex, but adopted a single conformation in mutant complexes. Notably MK-5172, with an ether-linked P2 quinoxaline moiety, while protruding from the substrate envelope, stacked against the catalytic triad, avoiding direct van der Waals contact with R155 and D168 (Figure 6A). Thus, although each of these drugs protruded from the substrate envelope at the S2 subsite, each formed unique interactions with R155, A156 and D168. Mutations at these residues therefore differentially affected drug binding and potency, resulting in a distinct resistance profile for each inhibitor.

Telaprevir resistance

Telaprevir lost potency against R155K compared to the wild-type protease, although the crystal structures of both complexes were very similar maintaining the covalent bond between the ketoamide moiety and the catalytic serine (Figure 3B). R155K, however, lost interactions with D168, thereby disrupting the electrostatic network spanning R123, D168, R155 and D81, which is important for telaprevir binding. These rearrangements modulated the charge landscape along the protease surface, disrupting interactions with the adjacent P2 cyclopentylproline and P4 cyclohexylalanine moieties of telaprevir, consistent with previous modeling studies [48]. Interestingly, telaprevir showed better potency against the D168A variant than the wild-type; the crystal structure revealed that the P2 moiety bent considerably and packed closer against the D168A variant. The inhibitor shifted by approximately 0.5 Å relative the position in the wild-type complex, resulting in increased interactions with both R155 and A156 (Figures 3C, 7A). However, the A156T mutation resulted in a steric clash with telaprevir's P2 moiety, causing the inhibitor to shift significantly; the inhibitor P2 moiety moved away from R155, losing van der Waals interactions with the protease (Figures 3D, 7A). Notably, in the A156T-telaprevir complex the covalent bond between the ketoamide warhead and the catalytic serine was extended to greater than 2 Å, suggesting a reduced capacity for covalent modification, consistent with the large loss in potency against A156T (Table 1). Thus, while telaprevir's flexibility allows adaptation to D168A, it cannot accommodate the disruption by R155K or A156T. The relatively weak binding affinity of telaprevir to wild-type protease results in a potentially narrow range by which resistant mutations can be tolerated.

Danoprevir and vaniprevir resistance

For both danoprevir and vaniprevir, the R155K mutation disrupted the favorable cation-π stacking interactions with the P2 isoindoline moieties (Figure 4B), causing significant reductions in drug potencies (Table 1). The D168A mutation also disrupted stacking of the P2 moieties with R155, by disrupting the electrostatic network and therefore the position of R155 for optimal cation-π stacking. In danoprevir, the P2 isoindoline moiety shifted in response to R155K and D168A mutations, making extensive interactions with the catalytic D81 (Figure 7B). For vaniprevir, the rigidity of the P2 P4 macrocycle prevented similar compensatory changes (Figure 5C). Thus, D168A caused losses in danoprevir and vaniprevir potency by disrupting cation-π stacking. However, the flexibility of the P2 moiety of danoprevir

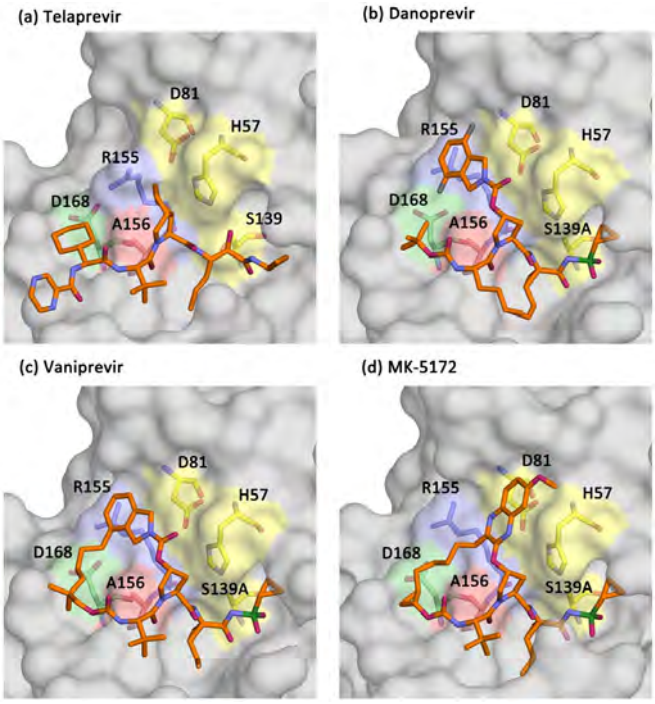


Figure 2. The binding conformations of telaprevir, danoprevir, vaniprevir and MK-5172. Surface representations of the wild-type protease in complex with (A) telaprevir, (B) danoprevir, (C) vaniprevir, and (D) MK-5172. The catalytic triad is shown in yellow and the R155, A156 and D168 side chains are highlighted in light-blue, pale-green and red, respectively.
doi:10.1371/journal.ppat.1002832.g002

compensates somewhat for this loss, explaining danoprevir's greater potency against the D168A variant relative to vaniprevir (Table 1).

The A156T mutation sterically impinges on the binding of danoprevir and vaniprevir. In both complexes with A156T, the P2 moieties shifted toward the catalytic triad and lost cation-π stacking interactions with R155. However, the flexibility of the P2 moiety of danoprevir permitted a larger shift, which allowed for more compensatory packing against the A156T variant protease surface (Figure 4D). In contrast, the P2 P4 macrocycle of vaniprevir restrained the P2 moiety and inhibitor's ability to accommodate this steric burden, more strongly compromising the activity of vaniprevir. Thus, the flexible P2 moiety of danoprevir allowed it to retain significant potency against A156T variants compared to vaniprevir.

MK 5172 resistance

Unlike in the danoprevir and vaniprevir complexes with wild type, in the MK-5172-wild-type complex the P2 quinoxaline moiety did not stack on R155 and interacted less with D168 and the electrostatic network involving these residues. Thus, the single-site mutations R155K and D168A only caused very subtle changes in the MK-5172 binding conformations (Figures 6B and 6C). This subtle effect is reflected in the small loss of potency against the R155K variant (Table 1); however, MK-5172 exhibited 100-fold lower potency against the D168A variant, likely due to less extensive interactions with D81 and K136 relative to wild-type and R155K (Figure 7). A156T, the worst of the resistance mutations for MK-5172A, sterically clashed with the P2 P4 macrocycle and caused a large shift in the binding position away from the catalytic triad relative to its wild-type structure (Figure 6D). This altered binding of MK-5172 resulted in fewer van der Waals contacts with D81 and R155, and is likely responsible for 1000-fold lower potency against the A156T variant. Overall, analysis of the four crystal structures explains MK-5172's significantly retained potency against R155K and D168A as well as its loss of potency against the A156T variant due to the rigidity of the macrocycle (Table 1).

Discussion

Despite the exciting therapeutic success of NS3/4A protease inhibitors, their long-term effectiveness is challenged by drug resistance. In this study we explain the molecular basis of this drug resistance against four NS3/4A protease inhibitors, telaprevir, danoprevir, vaniprevir and MK-5172, representing the major

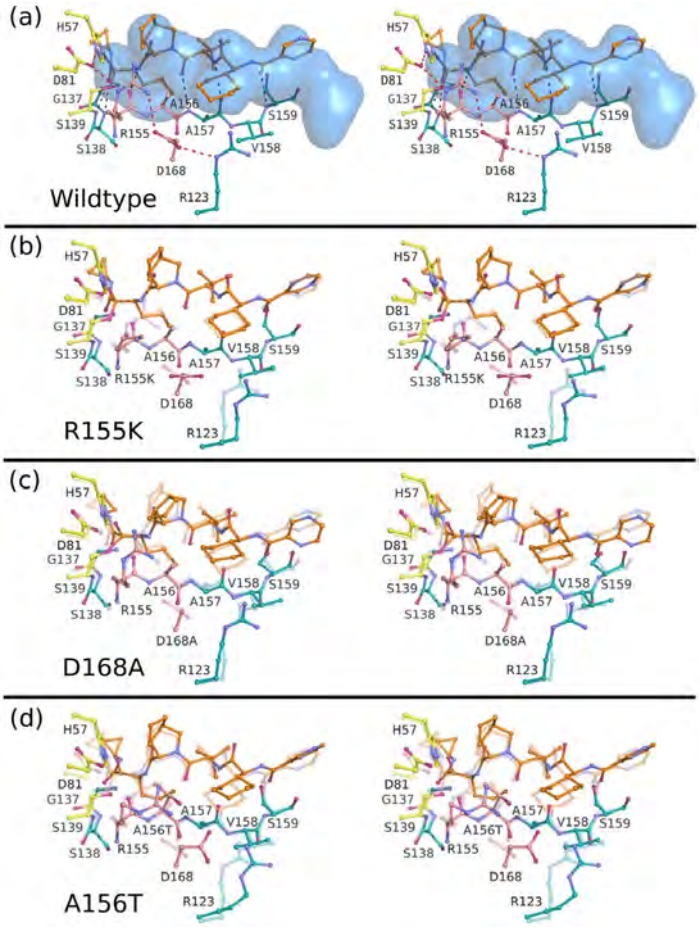


Figure 3. Stereo view of the telaprevir complexes. (A) Telaprevir bound to the wild-type protease with the substrate envelope in blue. Intra- and inter-molecular hydrogen bond interactions are marked as red and grey dashed lines. Telaprevir is also shown bound to the drug-resistant variants (B) R155K, (C) D168A and (D) A156T with the transparent coordinates representing the wild-type structure to better highlight the molecular changes of each mutation. In all cases, catalytic residues are depicted in yellow, the P2 subsite in pink, and the drug molecules in orange.
doi:10.1371/journal.ppat.1002832.g003

chemical classes of these inhibitors. Our detailed analysis of 16 high-resolution crystal structures explains the loss of inhibitor potency in the face of resistance mutations. This research supports our substrate envelope model, which stipulates that inhibitors are vulnerable to resistance where they contact protease residues beyond the substrate-binding region and therefore are not essential for substrate binding [43]. These sites can mutate with minimal effect on protease function and viral fitness. Indeed, most resistance mutations occur in regions where drugs protrude from the substrate envelope, as these changes selectively disrupt drug binding with minimal effect on substrate proteolysis.

The most potent of the NS3/4A protease inhibitors is MK-5172. We report here, for the first time, a novel binding conformation for MK-5172 in which the P2 quinoxaline moiety binds far from the S2 subsite and instead stacks against the catalytic residues H57 and D81. Unlike other inhibitors, MK5172 does not directly interact with R155 and D168, which mutate to confer multi-drug resistance. This unique binding mode of MK-5172 explains its significantly greater potency against R155K and D168A variants compared to other inhibitors. MK-5172 has a unique barrier to resistance, as neither catalytic residue (H57 or D81) can tolerate mutation. This binding conformation of MK-

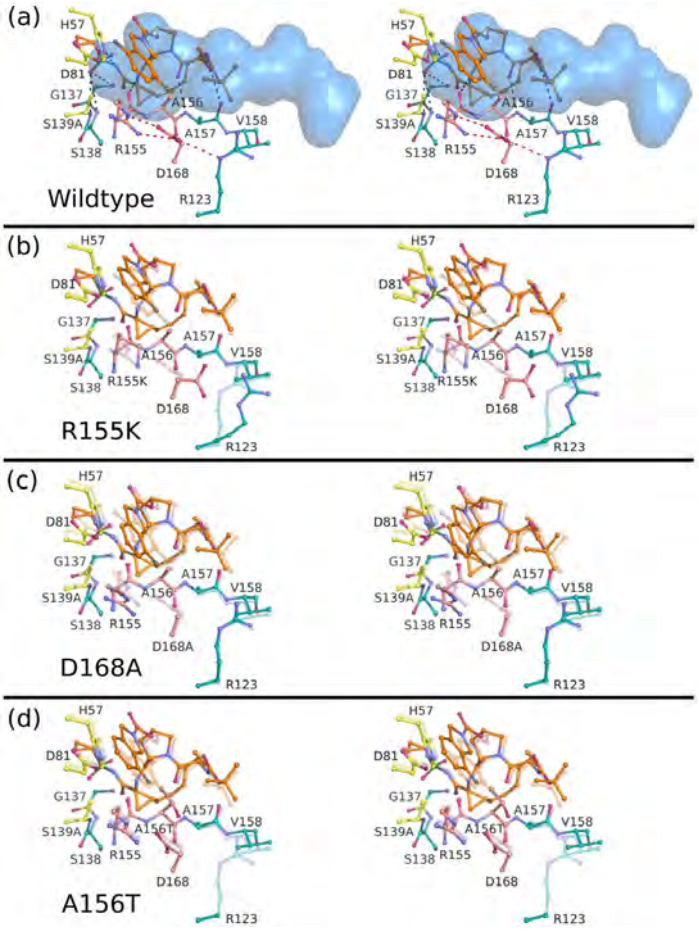


Figure 4. Stereo view of the danoprevir complexes. (A) Danoprevir bound to the wild-type protease with the substrate envelope in blue. Intra- and inter-molecular hydrogen bond interactions are marked as red and grey dashed lines. Danoprevir is also shown bound to the drug-resistant variants (B) R155K, (C) D168A and (D) A156T with the transparent coordinates representing the wild-type structure to better highlight the molecular changes of each mutation. In all cases, catalytic residues are depicted in yellow, the P2 subsite in pink, and the drug molecules in orange.
doi:10.1371/journal.ppat.1002832.g004

5172, combined with its picomolar binding affinity [27] (Table 2), will likely allow it to retain potency against a broad array of resistant viral variants and genotypes.

We define the structural basis for differential drug activities against the resistant variants R155K, D168A, and A156T for four major chemical classes of NS3/4A protease inhibitors. Telaprevir has reduced potency against R155K due to loss of van der Waals contacts but exhibits better potency against D168A as it allows tighter packing in the S2 subsite. R155K and D168A mutations confer danoprevir and vaniprevir resistance by disrupting favorable cation-π stacking interactions with R155. Interestingly, while both drugs lose considerable potency against R155K, danoprevir retains higher activity against D168A. This difference is likely due to the flexible P2 isoindoline moiety of danoprevir, which lacks P2 P4 cyclization and repacks against the D168A variant. Similarly, vaniprevir and MK-5172 exhibit significantly lower potency against the A156T variant due to direct steric clashes, while danoprevir partially accommodates this steric burden by repacking against the mutated surface. Thus, the flexibility of danoprevir's P2 isoindoline moiety allows it to retain activity

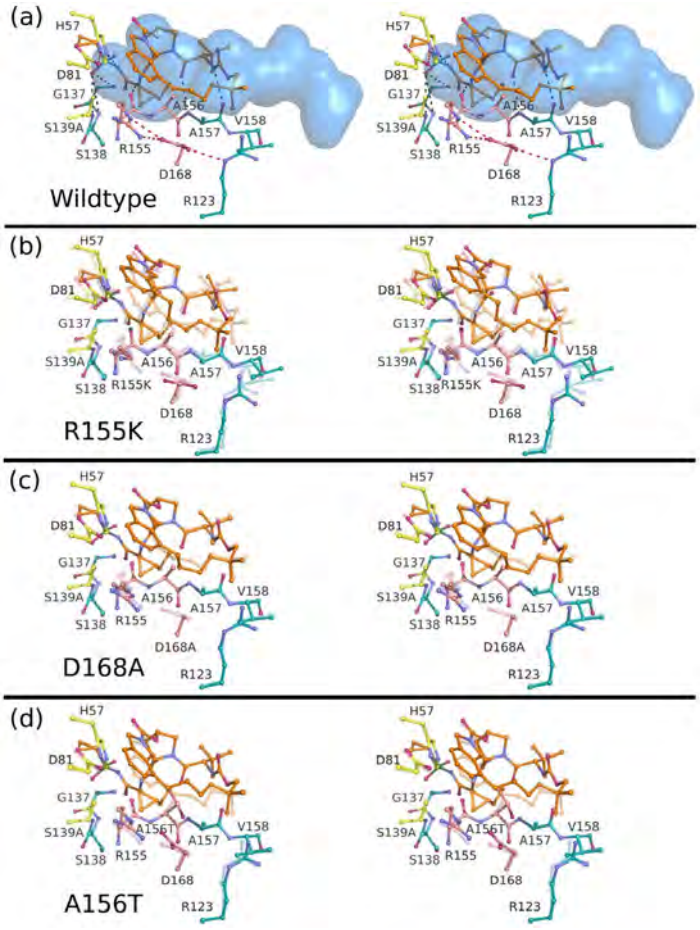


Figure 5. Stereo view of the vaniprevir complexes. (A) Vaniprevir bound to the wild-type protease with the substrate envelope in blue. Intra- and inter-molecular hydrogen bond interactions are marked as red and grey dashed lines. Vaniprevir is shown bound to the drug-resistant variants (B) R155K, (C) D168A and (D) A156T with the transparent coordinates representing the wild-type structure to better highlight the molecular changes of each mutation. In all cases, catalytic residues are depicted in yellow, the P2 subsite in pink, and the drug molecules in orange.
doi:10.1371/journal.ppat.1002832.g005

against two of the three major drug-resistant variants. Structural analysis of the 16 protease-inhibitor complexes defines the role of all three major drug-resistance mutations.

Our results also provide predictions of drug activities against other HCV genotypes and resistant strains. Interestingly, NS3/4A residues around the protease active site, including R155, A156, and D168 are highly conserved except genotype 3 viruses which contain the residues Q168 and T123, instead of D168 and R123 found in other genotypes (Figure S2). We predict that the terminal amide group of Q168 will be unable to stabilize R155 for stacking against the P2 moieties of danoprevir and vaniprevir, but may interact with T123 instead. Thus, we expect that danoprevir and vaniprevir will exhibit reduced potencies against genotype 3 viruses, while MK-5172 will remain fully active. Indeed, danoprevir and vaniprevir were recently shown to have reduced efficacy against genotype 3 viruses [49]. For genotype 1 strains, our results indicate that MK-5172 is highly active against R155K and D168A variants, while danoprevir is highly active against the A156T variant and to a lesser extent against the D168A variant. Thus, as new inhibitors are developed and HCV resistance testing becomes more available, our findings can help guide anti-HCV treatment regimens for individual patients.

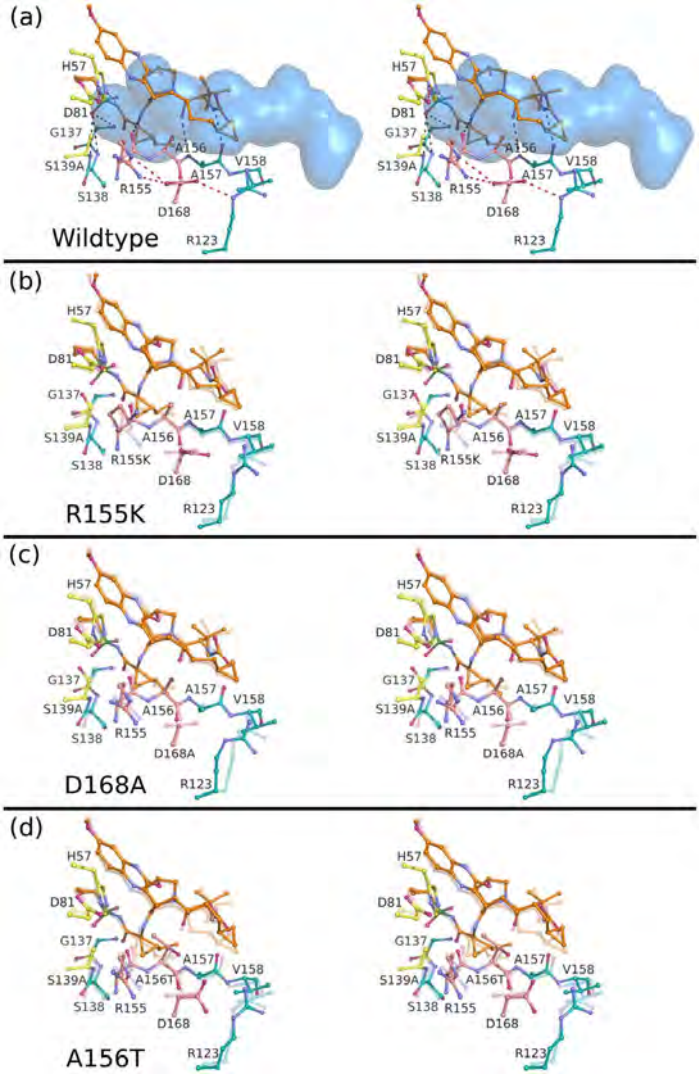


Figure 6. Stereo view of the MK-5172 complexes. (A) MK-5172 bound to the wild-type protease with the substrate envelope in blue. Intra- and inter-molecular hydrogen bond interactions are marked as red and grey dashed lines. MK-5172 is shown bound to the drug-resistant variants (B) R155K, (C) D168A and (D) A156T with the transparent coordinates representing the wild-type structure to better highlight the molecular changes of each mutation. In all cases, catalytic residues are depicted in yellow, the P2 subsite in pink, and the drug molecules in orange.
doi:10.1371/journal.ppat.1002832.g006

Overall our findings correlate with resistance profiles observed in clinical isolates. Most protease inhibitors select for R155K variants in genotype 1a patients as only one nucleotide change is required [29,30,32,37 42]. Genotype 1b patients presumably have higher barriers to R155K resistance, requiring two nucleotide substitutions; thus, mutations at A156 and D168 are more readily

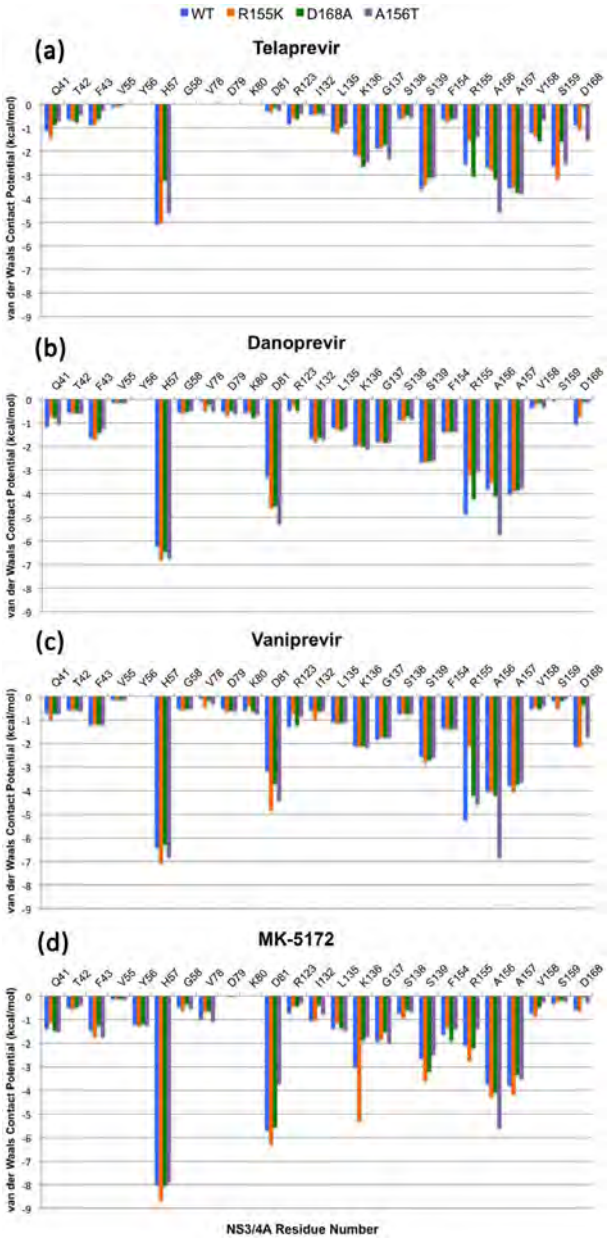

WT R155K D168A A156T
(a)
Telaprevir
(b)
Danoprevir
(c)
Vaniprevir
(d)
MK-5172
van der Waals Contact Potential (kcal/mol)
Q41 T42 F43 V55 Y56 H57 G58 V78 D79 K80 D81 R123 I132 L135 K136 G137 S138 S139 F154 R155 A156 A157 V158 S159 D168
NS3/4A Residue Number

Figure 7. Drug interactions with wild-type and mutant complexes by residue. The Van der Waals contact energy indexes for the wild-type protease and mutant variants R155K, D168A and A156T are shown by protease residue for (A) telaprevir, (B) danoprevir, (C) vaniprevir and (D) MK-5172.
doi:10.1371/journal.ppat.1002832.g007

observed in response to protease inhibitor treatment. The resistance at R155K occurs due to reduced interactions in the S2 subsite. Telaprevir and other linear ketoamide drugs select for A156T variants [39 41] by direct steric clashes, while linear (BI 201335) and macrocyclic drugs (danoprevir, vaniprevir, TMC435) with large P2 moieties select for D168A variants [29,30,32,42] by disrupting favorable stacking interactions with R155. These data also support the converse observation that D168A variants are uncommon in patients treated with telaprevir as the drug can pack tighter in the S2 subsite. Likewise, A156T variants are uncommon in patients treated with macrocyclic drugs containing flexible P2 moieties due to drug repacking against the mutated protease surface [29,30,32,42]. However, drugs such as vaniprevir and MK-5172 containing P2 P4 cyclization likely select for the A156T variant due to the rigidity of their P2 moieties. Whether A156T variants will be found in clinical isolates, however, depends on additional viral factors, such as relative differences in viral fitness between A156T variants and other competing viral variants. Our data thus provide a unique resource for preemptively predicting resistance and choosing the most appropriate protease inhibitor to treat HCV depending on the resistance profile of a particular patient viral population. Whether or not specific mutations arise in clinical isolates is ultimately determined by the complex interplay between drug potency, viral fitness, and genetic barriers to resistance. Thus, depending on the initial viral species altered pathways to resistance will exist.

The crystal structures of these NS3/4A inhibitors also provide a key resource to guide future strategies in drug design. The high potency of MK-5172, for example, derives from interactions with the essential catalytic triad residues, which cannot mutate without severely disrupting viral fitness. In addition, flexible P2 drug moieties lacking P2 P4 macrocycles mitigate losses in potency to the A156T and D168A mutations. Similar chemical features can be incorporated in future drugs to potentially evade resistance. Specifically, novel protease inhibitors that incorporate flexible P2 moieties, such as quinoxaline or similar groups, could exploit interactions with the essential catalytic residues and concurrently minimize contact with the P2 subsite, thereby reducing their sensitivities to mutations at R155, D168 and A156T. Thus, our findings suggest strategies for developing protease inhibitors that retain activity against a wider spectrum of drug-resistant HCV variants.

Materials and Methods

Inhibitor synthesis

Danoprevir, vaniprevir and MK-5172 were synthesized in house following reported methods; danoprevir was prepared using our convergent reaction sequence as described [43]; vaniprevir and MK-5172 were prepared following the synthetic methods reported by McCauley *et al.* [50] and Harper *et al.* [27], respectively, with minor modifications. Telaprevir was purchased from A ChemTek, Inc. (Worcester, MA).

Mutagenesis and gene information

The HCV genotype 1a NS3/4A protease gene described in a Bristol-Meyers Squibb patent [51] was synthesized by GenScript and cloned into the pET28a expression vector (Novagen). This highly soluble single-chain construct of the genotype 1a NS3/4A protease domain contains a fragment of the cofactor NS4A covalently linked at the N-terminus [51]. A similar protease construct exhibited catalytic activity comparable to that of the authentic full-length protein [52]. All protease variants were generated using the QuikChange Site-Directed Mutagenesis Kit from Stratagene. The codon-optimized genotype 1a helicase sequence (H77c) was cloned downstream to the protease gene to generate the full-length protease construct. Geneious [53] was used to generate the sequence alignment of the NS3/4A protease domain from HCV genotypes 1 6.

Drug susceptibility assays

Single mutations (R155K, D168A, or A156T) were introduced into the NS3 region of genotype 1a HCV Con1 luciferase reporter replicon using the mega-primer method of mutagenesis [54]. Replicon RNA of each protease variant was introduced into Huh7 cells by electroporation. Replication was then assessed in the presence of increasing concentrations of protease inhibitors (telaprevir, danoprevir, vaniprevir or MK-5172) by measuring luciferase activity (relative light units) 96 hours after electroporation. The drug concentrations required to inhibit replicon replication by 50% (IC_{50}) were calculated directly from the drug inhibition curves.

Enzyme inhibition assays

For enzyme inhibition experiments, 5 nM of the genotype 1a HCV NS3/4A protease domain was incubated with increasing drug concentrations for 15 min (90 min for telaprevir) in 50 mM Tris assay buffer (5% glycerol, 5 mM TCEP, 6 mM LDAO and 4% DMSO, pH 7.5). Proteolysis reactions were initiated by adding 100 nM HCV NS3/4A substrate [Ac-DE-Dap(QXL520)-EE-Abu-ψ-[COO]AS-C(5-FAMsp)-NH2] (AnaSpec) and monitored using the EnVision plate reader (Perkin Elmer) at excitation and emission wavelengths of 485 nm and 530 nm, respectively. The initial cleavage velocities were determined from sections of the progress curves corresponding to less than 15% substrate cleavage. Apparent inhibition constants (K_i) were obtained by nonlinear regression fitting to the Morrison equation of initial velocity versus inhibitor concentration using Prism 5 (GraphPad Software). Data for each drug were generated in triplicate and processed independently to calculate the average inhibition constant and standard deviation.

Expression and purification of NS3/4A protease constructs

Protein expression and purification were carried out as described [51,55]. Briefly, transformed BL21 (DE3) *E. coli* cells were grown at 37°C and induced at an optical density of 0.6 by adding 1 mM IPTG. Cells were harvested after 5 hours of expression, pelleted, and frozen at 80°C for storage. Cell pellets were thawed, resuspended in 5 mL/g of resuspension buffer (50 mM phosphate buffer, 500 mM NaCl, 10% glycerol, 2 mM β-ME, pH 7.5) and lysed with a cell disruptor. The soluble fraction was retained, applied to a nickel column (Qiagen), washed with resuspension buffer, and eluted with resuspension buffer supplemented with 200 mM imidazole. The eluent was dialyzed overnight (MWCO 10 kD) to remove the imidazole, and the His-tag was simultaneously removed with thrombin treatment.

The nickel-purified protein was then flash-frozen and stored at 80°C.

Crystallization of inhibitor complexes

The above-mentioned protein solution was thawed, concentrated to 3 mg/mL and loaded on a HiLoad Superdex75 16/60 column equilibrated with gel filtration buffer (25 mM MES, 500 mM NaCl, 10% glycerol, 30 µM zinc chloride, and 2 mM DTT, pH 6.5). The protease fractions were pooled and concentrated to 20 25 mg/mL with an Amicon Ultra-15 10 kD device (Millipore). The concentrated samples were incubated for 1 hour with 1 3 molar excess of inhibitor. Diffraction-quality crystals were obtained overnight by mixing equal volume of concentrated protein solution with precipitant solution (20 26% PEG-3350, 0.1 M sodium MES buffer, 4% ammonium sulfate, pH 6.5) in 24-well VDX hanging drop trays.

Crystallization, data collection and structure solution

X-ray diffraction data were collected at Advanced Photon Source LS-CAT 21-ID-F, GM/CA-CAT 23-ID-D or with the in-house RAXIS IV X-ray system. Diffraction intensities were indexed, integrated and scaled using the program HKL2000 [56]. All structure solutions were generated using simple isomorphous molecular replacement with PHASER [57]. The B chain model of viral substrate product 4A4B (3M5M) [43] was used as the starting model for all structure solutions. Initial refinement was carried out in the absence of modeled ligand, which was subsequently built in during later stages of refinement. Subsequent crystallographic refinement was carried out within the CCP4 program suite, with iterative rounds of TLS and restrained refinement until convergence was achieved [58]. The protein crystals of the wild-type protease and drug-resistant variants R155K and D168A in complex with MK-5172 grew as pseudo-merohedral twins. Amplitude-based twinned refinement was carried out during restrained refinement for all pseudo-merohedral twins. The final structures were evaluated with MolProbity [59] prior to deposition in the Protein Data Bank. To limit the possibility of model bias throughout the refinement process, 5% of the data were reserved for the free R-value calculation [60]. Interactive model building and electron density viewing was carried out using the program COOT [61]. F_{obs} F_{calc} ligand omit maps were generated with the ligand excluded from the phase calculation using the program PHENIX [62].

Inhibitor complex analysis

Superpositions were performed in PyMOL [63] using the Cα atoms of the active site protease residues 137 139 and 154 160. The wild-type-danoprevir complex was used as the reference structure for each alignment. The NS3/4A viral substrate envelope was computed as described using the full-length NS3/4A structure (1CU1) [64] and product complexes 4A4B (3M5M), 4B5A (3M5N) and 5A5B (3M5O) [43].

Van der Waals contact energy

Van der Waals contact energies between protease residues and peptide products were computed using a simplified Lennard-Jones potential as described [65]. Briefly, the Lennard-Jones potential (V_r) was calculated for each protease-drug atom pair where r, ε and σ represent the interatomic distance, vdW well depth, and atomic diameter, respectively:

$$V_r \quad 4\,\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right]$$

V_r was computed for all possible protease-drug atom pairs within 5 Å, and potentials for non-bonded pairs separated by less than the distance at the minimum potential were equated to ε.

Supporting Information

Figure S1 Ligand omit maps for protease-inhibitor complexes. The NS3/4A wild-type protease (grey cartoon) is shown with inhibitors (orange sticks): (A) telaprevir, (B) danoprevir, (C) vaniprevir and (D) MK-5172. The electron density maps (blue) depict the F_{obs} F_{calc} ligand omit maps contoured at 1σ.
(TIF)

Figure S2 Sequence alignment of the NS3/4A protease domain for HCV genotypes 1–6. Consensus sequence (1a M62321) of NS3/4A protease domain is shown in grey. Amino acid residues in disagreement are highlighted in color. Residues at positions 155 and 156 are conserved across genotypes; however, genotype 3 shows divergence from the consensus at amino acid 168.
(TIF)

Table S1 Drug hydrogen bonds and vdW contacts with wild-type protease.
(DOC)

Acknowledgments

We thank Shivender Shandilya for generating the omit maps; Nukri Sanishvili of the GM/CA CAT beamline, Markus Bohn and Andrei Korostelev for data collection of the D168A MK 5172 complex; David Smith of the LS CAT beamline for data collection of all other drug complexes; Herbert Klei, Madhavi Kolli, and William Royer for helpful discussions; Seema Mittal, and Madhavi Nalam for their computational support; and Nese Kurt Yilmaz for helpful revisions. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. DE AC02 06CH11357. Use of the LS CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri Corridor for the support of this research program (Grant 085P1000817). GM/CA CAT has been funded in whole or in part with federal funds from the National Cancer Institute (Y1 CO 1020) and the National Institute of General Medical Sciences (Y1 GM 1104).

Author Contributions

Conceived and designed the experiments: CAS KPR AA. Performed the experiments: KPR AA CA DS AO LMD CS HC AN WH. Analyzed the data: KPR AA CAS. Contributed reagents/materials/analysis tools: KPR AA CA DS AO LMD CS HC AN CJP WH. Wrote the paper: KPR AA CAS.

References

1 Shepard CW, Finelli L, Alter MJ (2005) Global epidemiology of hepatitis C virus infection Lancet Infect Dis 5 558 567
2 Lavanchy D (2009) The global burden of hepatitis C Liver Int 29 74 81
3 Simmonds P, Bukh J, Combet C, Deléage G, Enomoto N, et al (2005) Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes Hepatology 42 962 973

4 Liang TJ, Heller T (2004) Pathogenesis of hepatitis C associated hepatocellular carcinoma Gastroenterology 127 S62 S71
5 Ghany MG, Nelson DR, Strader DB, Thomas DL, Seeff LB (2011) An update on treatment of genotype 1 chronic hepatitis C virus infection 2011 practice guideline by the American Association for the Study of Liver Diseases Hepatology 54 1433 1444
6 McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, et al (2009) Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection N Engl J Med 360 1827 1838
7 Kwo PY, Lawitz EJ, McCone J, Schiff ER, Vierling JM, et al (2010) Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa 2b and ribavirin in treatment naive patients with genotype 1 hepatitis C infection (SPRINT 1) an open label, randomised, multicentre phase 2 trial Lancet 376 705 716
8 Jacobson IM, McHutchison JG, Dusheiko G, Di Bisceglie AM, Reddy KR, et al (2011) Telaprevir for previously untreated chronic hepatitis C virus infection N Engl J Med 364 2405 2416
9 Hézode C, Forestier N, Dusheiko G, Ferenci P, Pol S, et al (2009) Telaprevir and peginterferon with or without ribavirin for chronic HCV infection N Engl J Med 360 1839 1850
10 Poordad F, McCone J, Bacon BR, Bruno S, Manns MP, et al (2011) Boceprevir for untreated chronic HCV genotype 1 infection N Engl J Med 364 1195 1206
11 Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, et al (2009) Genetic variation in IL28B predicts hepatitis C treatment induced viral clearance Nature 461 399 401
12 Fried MW (2002) Side effects of therapy of hepatitis C and their management Hepatology 36 S237 S244
13 Kolykhalov AA, Mihalik K, Feinstone SM, Rice CM (2000) Hepatitis C virus encoded enzymatic activities and conserved RNA elements in the 3′ nontranslated region are essential for virus replication in vivo J Virol 74 2046 2051
14 Li K, Foy E, Ferreon JC, Nakamura M, Ferreon AC, et al (2005) Immune evasion by hepatitis C virus NS3/4A protease mediated cleavage of the Toll like receptor 3 adaptor protein TRIF Proc Natl Acad Sci U S A 102 2992 2997
15 Foy E, Li K, Wang C, Sumpter R, Jr , Ikeda M, et al (2003) Regulation of interferon regulatory factor 3 by the hepatitis C virus serine protease Science 300 1145 1148
16 Lamarre D, Anderson PC, Bailey M, Beaulieu P, Bolger G, et al (2003) An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus Nature 426 186 189
17 Hinrichsen H, Benhamou Y, Wedemeyer H, Reiser M, Sentjens RE, et al (2004) Short term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients Gastroenterology 127 1347 1355
18 Vanwolleghem T, Meuleman P, Libbrecht L, Roskams T, De Vos R, et al (2007) Ultra rapid cardiotoxicity of the hepatitis C virus protease inhibitor BILN 2061 in the urokinase type plasminogen activator mouse Gastroenterology 133 1144 1155
19 Malcolm BA, Liu R, Lahser F, Agrawal S, Belanger B, et al (2006) SCH 503034, a mechanism based inhibitor of hepatitis C virus NS3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells Antimicrob Agents Chemother 50 1013 1020
20 Perni RB, Almquist SJ, Byrn RA, Chandorkar G, Chaturvedi PR, et al (2006) Preclinical profile of VX 950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3 4A serine protease Antimicrob Agents Chemother 50 899 909
21 Kwong AD, Kauffman RS, Hurter P, Mueller P (2011) Discovery and development of telaprevir an NS3 4A protease inhibitor for treating genotype 1 chronic hepatitis C virus Nat Biotech 29 993 1003
22 McPhee F (2010) Identification and preclinical profile of the novel HCV NS3 protease inhibitor BMS 650032 J Hepatol 52 S296
23 White PW, Llinàs Brunet M, Amad Ma, Bethell RC, Bolger G, et al (2010) Preclinical characterization of BI 201335, a C terminal carboxylic acid inhibitor of the hepatitis C virus NS3 NS4A protease Antimicrob Agents Chemother 54 4611 4618
24 Seiwert SD, Andrews SW, Jiang Y, Serebryany V, Tan H, et al (2008) Preclinical characteristics of the hepatitis C virus NS3/4A protease inhibitor ITMN 191 (R7227) Antimicrob Agents Chemother 52 4432 4441
25 Lin T I, Lenz O, Fanning G, Verbinnen T, Delouvroy F, et al (2009) In vitro activity and preclinical profile of TMC435350, a potent hepatitis C virus protease inhibitor Antimicrob Agents Chemother 53 1377 1385
26 Liverton NJ, Carroll SS, DiMuzio J, Fandozzi C, Graham DJ, et al (2010) MK 7009, a potent and selective inhibitor of hepatitis C virus NS3/4A protease Antimicrob Agents Chemother 54 305 311
27 Harper S, McCauley JA, Rudd MT, Ferrara M, DiFilippo M, et al (2012) Discovery of MK 5172, a macrocyclic hepatitis C virus NS3/4a protease inhibitor ACS Med Chem Lett DOI 10 1021/ml300017p
28 Adiwijaya BS, Hare B, Caron PR, Randle JC, Neumann AU, et al (2009) Rapid decrease of wild type hepatitis C virus on telaprevir treatment Antivir Ther 14 591 595
29 Manns MP, Bourlière M, Benhamou Y, Pol S, Bonacini M, et al (2011) Potency, safety, and pharmacokinetics of the NS3/4A protease inhibitor BI201335 in patients with chronic HCV genotype 1 infection J Hepatol 54 1114 1122
30 Lim SR, Qin X, Susser S, Nicholas JB, Lange C, et al (2012) Virologic escape during danoprevir (ITMN 191/RG7227) monotherapy is hepatitis C virus subtype dependent and associated with R155K substitution Antimicrob Agents Chemother 56 271 279
31 Reesink HW, Zeuzem S, Weegink CJ, Forestier N, van Vliet A, et al (2006) Rapid decline of viral RNA in hepatitis C patients treated with VX 950 a phase Ib, placebo controlled, randomized study Gastroenterology 131 997 1002
32 Manns M, Reesink H, Berg T, Dusheiko G, Flisiak R, et al (2011) Rapid viral response of once daily TMC435 plus pegylated interferon/ribavirin in hepatitis C genotype 1 patients a randomized trial Antivir Ther 16 1021 1033
33 Bacon BR, Gordon SC, Lawitz E, Marcellin P, Vierling JM, et al (2011) Boceprevir for previously treated chronic HCV genotype 1 infection N Engl J Med 364 1207 1217
34 Zeuzem S, Andreone P, Pol S, Lawitz E, Diago M, et al (2011) Telaprevir for retreatment of HCV infection N Engl J Med 364 2417 2428
35 Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, et al (1998) Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon alpha therapy Science 282 103 107
36 Rong L, Dahari H, Ribeiro RM, Perelson AS (2010) Rapid emergence of protease inhibitor resistance in hepatitis C virus Sci Transl Med 2 30ra32
37 Sarrazin C, Zeuzem S (2010) Resistance to direct antiviral agents in patients with hepatitis C virus infection Gastroenterology 138 447 462
38 Kieffer TL, Kwong AD, Picchio GR (2010) Viral resistance to specifically targeted antiviral therapies for hepatitis C (STAT Cs) J Antimicrob Chemother 65 202 212
39 Kieffer TL, Sarrazin C, Miller JS, Welker MW, Forestier N, et al (2007) Telaprevir and pegylated interferon alpha 2a inhibit wild type and resistant genotype 1 hepatitis C virus replication in patients Hepatology 46 631 639
40 Sarrazin C, Kieffer TL, Bartels D, Hanzelka B, Muh U, et al (2007) Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir Gastroenterology 132 1767 1777
41 Susser S, Welsch C, Wang Y, Zettler M, Domingues FS, et al (2009) Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus infected patients Hepatology 50 1709 1718
42 Manns MP, Gane E, Rodriguez Torres M, Stoehr A, Yeh CT, et al (2009) MK 7009 significantly improves rapid viral response (RVR) in combination with pegylated interferon alfa 2a and ribavirin in patients with chronic hepatitis c (CHC) genotype 1 infection J Hepatol 50 S384
43 Romano KP, Ali A, Royer WE, Schiffer CA (2010) Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding Proc Natl Acad Sci U S A 107 20986 20991
44 He Y, King MS, Kempf DJ, Lu L, Lim HB, et al (2008) Relative replication capacity and selective advantage profiles of protease inhibitor resistant hepatitis C virus (HCV) NS3 protease mutants in the HCV genotype 1b replicon system Antimicrob Agents Chemother 52 1101 1110
45 Cummings MD, Lindberg J, Lin TI, de Kock H, Lenz O, et al (2010) Induced fit binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target Angew Chem Int Ed Engl 49 1652 1655
46 Bae A, Sun S C, Qi X, Chen X, Ku K, et al (2010) Susceptibility of treatment naive hepatitis C Virus (HCV) clinical isolates to HCV protease inhibitors Antimicrob Agents Chemother 54 5288 5297
47 Lenz O, Verbinnen T, Lin TI, Vijgen L, Cummings MD, et al (2010) In vitro resistance profile of the hepatitis C virus NS3/4A protease inhibitor TMC435 Antimicrob Agents Chemother 54 1878 1887
48 Zhou Y, Muh U, Hanzelka BL, Bartels DJ, Wei Y, et al (2007) Phenotypic and structural analyses of hepatitis C virus NS3 protease Arg155 variants sensitivity to telaprevir (VX 950) and interferon alpha J Biol Chem 282 22619 22628
49 Gottwein JM, Scheel TKH, Jensen TB, Ghanem L, Bukh J (2011) Differential efficacy of protease inhibitors against HCV genotypes 2a, 3a, 5a, and 6a NS3/4A protease recombinant viruses Gastroenterology 141 1067 1079
50 McCauley JA, McIntyre CJ, Rudd MT, Nguyen KT, Romano JJ, et al (2010) Discovery of vaniprevir (MK 7009), a macrocyclic hepatitis C virus NS3/4a protease inhibitor J Med Chem 53 2443 2463
51 Wittekind M, Weinheirner S, Zhang Y, Goldfarb V (2002) Modified forms of hepatitis C NS3 protease for facilitating inhibitor screening and structural studies of protease inhibitor complexes US Patent 6333186
52 Taremi SS, Beyer B, Maher M, Yao N, Prosise W, et al (1998) Construction, expression, and characterization of a novel fully activated recombinant single chain hepatitis C virus protease Protein Sci 7 2143 2149
53 Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, et al (2011) Geneious v5 4 Available from http //www geneious com/
54 Sarkar G, Sommer SS (1990) The ''megaprimer'' method of site directed mutagenesis Biotechniques 8 404 407
55 Gallinari P, Brennan D, Nardi C, Brunetti M, Tomei L, et al (1998) Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus J Virol 72 6758 6769
56 Otwinowski Z, Minor W (1997) Processing of X ray Diffraction Data Collected in Oscillation Mode In C W Carter JRMS, Eds Methods in Enzymology Volume 276 Macromolecular Crystallography, part A p 307 326
57 McCoy AJ, Grosse Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al (2007) Phaser crystallographic software J Appl Crystallogr 40 658 674
58 Collaborative Computational Project N (1994) The CCP4 Suite Programs for Protein Crystallography Acta Crystallographica 50 760 763

59 Davis IW, Leaver Fay A, Chen VB, Block JN, Kapral GJ, et al (2007) MolProbity: all atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35: W375–W383.
60 Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355: 472–475.
61 Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126–2132.
62 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66: 213–221.
63 DeLano WL (2008) The PyMOL Molecular Graphics System. San Carlos, CA: DeLano Scientific LLC.
64 Yao N, Reichert P, Taremi SS, Prosise WW, Weber PC (1999) Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. Structure 7: 1353–1363.
65 Nalam MN, Ali A, Altman MD, Reddy GS, Chellappan S, et al (2010) Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. J Virol 84: 5368–5378.

Appendix D

Secondary Contribution 2: Evaluating the role of macrocyclization in HCV protease inhibitor resistance

4.1 Secondary Contribution

Author contributions: **DS** expressed and purified, various protease used in this study. Additionally, **DS** participated in experimental troubleshooting as well as manuscript preparation.

ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Evaluating the Role of Macrocycles in the Susceptibility of Hepatitis C Virus NS3/4A Protease Inhibitors to Drug Resistance

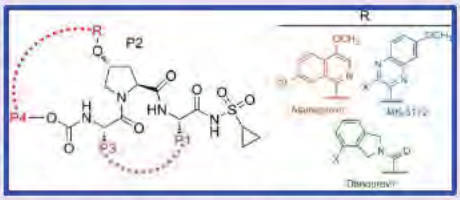
Akbar Ali,[†,§] Cihan Aydin,[†,§] Reinhold Gildemeister,[†] Keith P. Romano, Hong Cao,[†] Ayşegül Özen,[†] Djade Soumana,[†] Alicia Newton,[‡] Christos J. Petropoulos,[‡] Wei Huang,[‡] and Celia A. Schiffer*[,†]

[†]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, United States

[‡]Monogram Biosciences, South San Francisco, California 94080, United States

Supporting Information

ABSTRACT: The hepatitis C virus (HCV) infects an estimated 150 million people worldwide and is the major cause of viral hepatitis, cirrhosis, and liver cancer. The available antiviral therapies, which include PEGylated interferon, ribavirin, and one of the HCV NS3/4A protease inhibitors telaprevir or boceprevir, are ineffective for some patients and cause severe side effects. More potent NS3/4A protease inhibitors are in clinical development, but the long-term effectiveness of these drugs is challenged by the development of drug resistance. Here, we investigated the role of macrocycles in the susceptibility of NS3/4A protease inhibitors to drug resistance in asunaprevir, danoprevir, vaniprevir, and MK-5172, with similar core structures but varied P2 moieties and macrocyclizations. Linear and macrocyclic analogues of these drugs were designed, synthesized, and tested against wild-type and drug-resistant variants R155K, V36M/R155K, A156T, and D168A in enzymatic and antiviral assays. Macrocyclic inhibitors were generally more potent, but the location of the macrocycle was critical for retaining activity against drug-resistant variants: the P1–P3 macrocyclic inhibitors were less susceptible to drug resistance than the linear and P2–P4 macrocyclic analogues. In addition, the heterocyclic moiety at P2 largely determined the inhibitor resistance profile, susceptibility to drug resistance, and the extent of modulation by the helicase domain. Our findings suggest that to design robust inhibitors that retain potency to drug-resistant NS3/4A protease variants, inhibitors should combine P1–P3 macrocycles with flexible P2 moieties that optimally contact with the invariable catalytic triad of this enzyme.



The hepatitis C virus (HCV) infects an estimated 150 million people worldwide and is the major cause of viral hepatitis, cirrhosis, and liver cancer.[1] The current standard of care for HCV-infected patients is a triple combination therapy with PEGylated interferon α (Peg-IFN), ribavirin (RBV), and telaprevir or boceprevir,[2] recently approved direct-acting antiviral agents targeting the viral NS3/4A protease. This new treatment regime is effective for most patients,[3−7] but a combination of viral and host-cell genetic factors, including a human polymorphism at the IL28B gene associated with poor interferon response,[8] causes treatment failure in some patients. In addition, severe side effects associated with Peg-IFN[9] and the rapid emergence of drug resistance against protease inhibitors limit both patient adherence and the effectiveness of triple combination therapy.[10,11] Thus, more robust new direct-acting antivirals are needed against a broader spectrum of resistant HCV variants and genotypes.

The HCV NS3/4A is a bifunctional protein with a chymotrypsin-like serine protease and an ATP-dependent RNA helicase.[12,13] The NS3/4A protease is essential for viral replication playing a pivotal role in polyprotein processing and circumventing host immune response.[14−16] Pharmaceutical companies have invested significant effort in developing NS3/4A protease inhibitors, leading thus far to the discovery of two FDA approved drugs, boceprevir[17] and telaprevir,[18] both of which contain a ketoamide group that form a reversible, covalent bond with the catalytic serine of the protease. In addition, several non-covalent NS3/4A inhibitors are at a variety of stages of clinical development, and include both linear (asunaprevir,[19] BI 201335[20]) and macrocyclic compounds, containing either a P1–P3 (danoprevir,[21] TMC435[22]) or a P2–P4 (vaniprevir,[23] MK-5172[24]) macrocycle. Most of these drug discovery efforts primarily focus on the protease domain alone. However, a recent co-crystal structure of the full-length NS3/4A with an inhibitor interacting with both the helicase and protease domains[25] supports exploring inhibitor-helicase interactions in the design of NS3/4A protease inhibitors.

The NS3/4A protease inhibitors rapidly reduce HCV RNA levels when administered as monotherapy[26−29] and significantly improve treatment outcomes when given in combination with Peg-IFN/RBV.[3,7,30−32] However, HCV evolves very quickly

Received: February 9, 2013
Accepted: April 17, 2013
Published: April 17, 2013

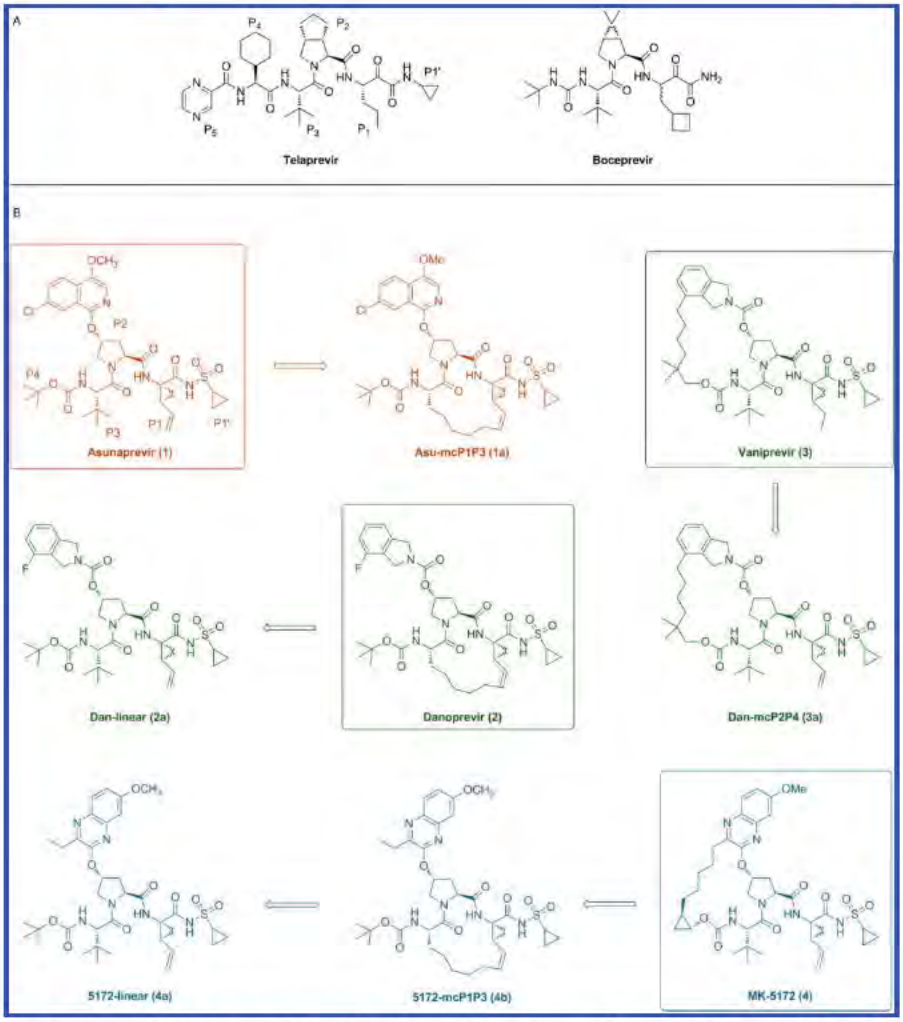


Figure 1. Structures of NS3/4A protease inhibitors: (A) FDA approved drugs telaprevir and boceprevir. (B) Asunaprevir (1) and P1−P3 macrocyclic analogue Asu mcP1P3 (1a); danoprevir (2), linear analogue Dan linear (2a), vaniprevir (3), and P1 olefin analogue Dan mcP2P4 (3a); MK 5172 (4), linear analogue 5172 linear (4a), and P1−P3 macrocyclic analogue 5172 mcP1P3 (4b). The canonical nomenclature for drug moiety positioning is indicated using telaprevir and asunaprevir; arrows indicate corresponding analogues. Asunaprevir, danoprevir/vaniprevir, and MK 5172 inhibitor series, which share the same P2 moiety but have different macrocylization status, are indicated by orange, green, and blue, respectively.

and drug resistance can emerge against direct-acting antiviral agents.[10,33,34] In patients undergoing protease inhibitor therapy, resistance develops due to overlapping but distinct sets of mutations in the NS3/4A protease.[10,34] Notably, single-site mutations at protease residues R155, A156, and D168 confer resistance to nearly all inhibitors in clinical development.[10,34] Mutations at V36, T54, and V36/R155 are also associated with resistance, particularly to ketoamide inhibitors.[35,36] Thus, despite the therapeutic success of NS3/4A protease inhibitors their long-term effectiveness is challenged by drug resistance.

We have previously shown that the NS3/4A protease inhibitors are particularly vulnerable to resistance where the inhibitors contact protease residues beyond the substrate-binding region, defined as the substrate envelope.[37] These

dx.doi.org/10.1021/cb400100g | ACS Chem. Biol. 2013, 8, 1469−1478

residues are not essential for substrate binding and proteolysis, and mutations at these sites can selectively disrupt drug binding with minimal effect on protease function.[37] The non-covalent NS3/4A protease inhibitors protrude from the substrate envelope in the S2 subsite where large P2 inhibitor moieties extensively interact with residues R155, A156, and D168,[37,38] which mutate to confer multi-drug resistance.[10] Recently, we elucidated the molecular basis of drug resistance by determining co-crystal structures of several NS3/4A protease inhibitors in complex with wild-type protease and major drug-resistant variants.[38] Danoprevir and vaniprevir are highly susceptible to mutations at R155 and D168 as these changes disrupt favorable stacking interactions of the isoindoline moiety with R155.[38] However, the P2 quinoxaline moiety in MK-5172 does not directly interact with R155 and D168 but instead the catalytic residues H57 and D81. This unique binding mode explains the retained potency of MK-5172 against R155K and D168A variants.[38] Thus, various P2 moieties appear to differentially impact inhibitor activities against major drug-resistant variants.

The NS3/4A protease inhibitors with macrocycles at either P1 P3 or P2 P4 are reported to have enzyme inhibitory activities better than those of their linear analogues against the wild-type protease, with corresponding enhancements in antiviral potencies.[39,40] However, the role of macrocycles in the susceptibility to drug resistance has not been thoroughly evaluated. In this study, we investigated the effect of varying macrocyclization state on inhibitor activity against drug-resistant variants for four representative protease inhibitors, asunaprevir, danoprevir, vaniprevir, and MK-5172, containing similar core structures but different P2 moieties. We designed and synthesized linear and macrocyclic analogues of these drugs and tested their activities in enzymatic and replicon-based antiviral assays against wild-type and drug-resistant variants. Inhibitor activities against variants of the full-length NS3/4A and the isolated protease domain were compared to assess modulation of inhibitor binding by the helicase domain. The macrocyclic inhibitors were generally more potent than their linear analogues, but the location of the macrocycle significantly affected the potency against drug-resistant variants. The heterocyclic moiety at P2 was the major determinant of inhibitor resistance profiles, susceptibility to drug resistance, and the extent of modulation by the helicase domain. This study elucidates the role of macrocycles and P2 moieties in HCV NS3/4A protease inhibitor resistance and suggests strategies for designing robust drugs against this rapidly evolving virus.

RESULTS AND DISCUSSION

Mutations in NS3/4A protease, particularly at residues R155, A156, and D168, confer resistance to nearly all inhibitors in clinical development. However, the impact of drug resistance mutations on potency varies greatly between the various inhibitors. Elucidating the role of different structural features that affect inhibitor potency against drug-resistant variants is essential for designing more robust NS3/4A protease inhibitors.

Synthesis of Protease Inhibitors with Varied P2 Moieties and Macrocyclizations. The non-covalent NS3/4A protease inhibitors, such as asunaprevir (**1**), danoprevir (**2**), vaniprevir (**3**), and MK-5172 (**4**), contain similar core structures, including an acylsulfonamide-linked P1 cysteine mimic, a substituted P2 proline, a hydrophobic P3 side chain, and a P4 capping group (Figure 1). However, these inhibitors are structurally distinct as heterocyclic moieties at P2 are different and inhibitors are either linear or macrocyclic; macrocyclization is either between the P1 side chain and the P3 side chain (mcP1P3) or the P2 heterocyclic moiety and the P4 capping group (mcP2P4). To evaluate the effect of P2 moieties and macrocylization status on inhibitor activity against drug-resistant variants, we designed and synthesized linear and macrocyclic analogues of asunaprevir, danoprevir, vaniprevir, and MK-5172 (Figure 1). Danoprevir and vaniprevir have many key components in common, and their analogues are clustered together, hence three inhibitor series with the same P2 group but different macrocylization status.

For asunaprevir, a linear acylsulfonamide compound with an ether-linked isoquinoline moiety at P2 proline, a P1 P3 macrocyclic analogue Asu-mcP1P3 (**1a**) was designed. Danoprevir and vaniprevir are both macrocyclic compounds containing a carbamate-linked isoindoline moiety at P2 proline, but the location of the macrocycle is different: danoprevir contains a P1 P3 macrocycle, whereas vaniprevir contains a P2 P4 macrocycle and a saturated P1 side chain. For direct comparison with danoprevir, a danoprevir linear analogue Dan-linear (**2a**) was designed along with a vaniprevir P1 olefin analogue Dan-mcP2P4 (**3a**). For MK-5172, which is a P2 P4 macrocyclic compound containing an ether-linked quinoxaline moiety, a linear analogue 5172-linear (**4a**) and a P1 P3 macrocyclic analogue 5172-mcP1P3 (**4b**) were prepared; in both analogues the hydrocarbon chain of the P2 P4 macrocycle was replaced with an ethyl group at the quinoxaline.

The designed macrocyclization analogues were prepared using synthetic sequences illustrated in Figure 2. Analogous to asunaprevir synthesis, Asu-mcP1P3 (**1a**) was prepared from the intermediate **5**, obtained by the reaction of Boc-*trans*-4-hydroxy-L-proline and 1,7-dichloro-4-methoxyisoquinoline.[41] Coupling of **5** with the P1 P1′ fragment **6**[42] provided the P2 P1′ intermediate **7**. Boc deprotection and coupling of the resulting free amine with the amino acid **9** provided the precursor bis-olefin **10** for ring-closing metathesis. Cyclization of the bis-olefin intermediate **10** was accomplished using a highly efficient ring-closing metathesis catalyst Zhan 1B and provided the target compound Asu-mcP1P3 (**1a**) (Figure 2A).

Dan-linear (**2a**) was synthesized using the reaction sequence analogous to danoprevir.[37] Vaniprevir (**3**) and its P1 olefin analogue Dan-mcP2P4 (**3a**) were synthesized following the synthetic methods reported by McCauley et al.[43]

The MK-5172 analogues 5172-linear (**4a**) and 5172-mcP1P3 (**4b**) were synthesized using the reaction sequences shown in Figure 2B. The intermediate **11** was prepared from 3-chloro-7-methoxyquinoxalin-2-ol and the bosylated Boc-*cis*-4-hydroxy-L-proline using Cs_2CO_3-mediated S_N2 displacement reaction.[24] Palladium-catalyzed vinylation followed by hydrogenation of the double bond provided the common P2 intermediate **12**. The acyclic analogue 5172-linear (**4a**) was assembled from **12** by a four-step reaction sequence, including Boc deprotection, P2 P3 amide coupling, ester hydrolysis, and coupling with the P1 P1′ fragment **6** to provide the target compound 5172-linear (**4a**). The precursor for 5172-mcP1P3, bis-olefin **18**, was synthesized from the P2 fragment **12** in an analogous reaction sequence. Finally, cyclization of the bis-olefin intermediate **18** using ring-closing metathesis catalyst Zhan 1B provided the 5172-mcP1P3 (**4b**).

Macrocyclization Enhances Inhibitor Potency against Wild Type and Drug Resistant Variants. The enzyme

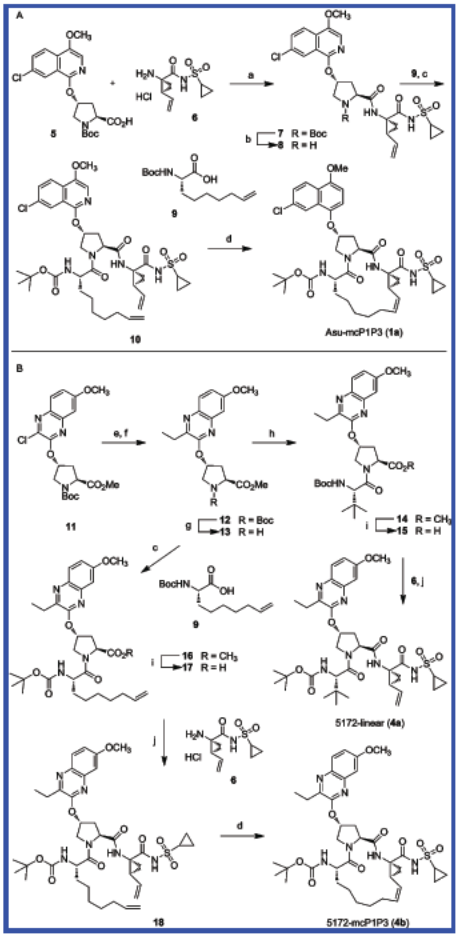


Figure 2. Synthesis of designed protease inhibitors. Reagents and conditions: (a) HATU, DIEA, DMF, CH_2Cl_2; (b) HCl, MeOH; (c) HATU, DIEA, DMF; (d) Zhan 1B catalyst, 1,2 DCE; (e) potassium vinyltrifluoroborate, TEA, $PdCl_2$(dppf) CH_2Cl_2, EtOH; (f) H_2, 10% Pd/C, MeOH, dioxane; (g) 4 N HCl, dioxane; (h) Boc Tle OH, HATU, DIEA, DMF; (i) LiOH, THF, H_2O; (j) HATU, DMAP, DIEA, DMF, CH_2Cl_2.

inhibition constants (K_i) of asunaprevir, danoprevir, vaniprevir, MK-5172, and their respective analogues were determined against genotype 1a full-length NS3/4A, the isolated protease domain, and drug-resistant variants R155K, V36M/R155K, A156T, and D168A of both full-length enzyme and protease (Figure 3a; Supplementary Tables S1 and S2, Figure S1). In addition, half-maximal inhibitory concentrations (IC_{50}) were determined using viral replicon-based inhibition assays with wild-type genotype 1b HCV and drug-resistant variants R155K, V36M/R155K, A156T, and D168A; telaprevir was used as a control (Figure 3b; Supplementary Table S3, Figures S2 S6).

In each inhibitor series, macrocyclic compounds exhibited K_i's lower than those of the corresponding linear analogues against wild-type protease and, in most cases, drug-resistant variants. The P2 P4 macrocyclic compounds were highly active against wild-type protease, with K_i values in the mid to high pM range (K_i 0.46 nM for vaniprevir, K_i 0.22 nM for Dan-mcP2P4, and K_i 0.06 nM for MK-5172), and retained low to high nanomolar activity against resistant variants. The P1 P3 macrocyclic compounds exhibited slightly lower enzymatic activities against wild-type, but overall retained better activities against drug-resistant variants. Notably, 5172-mcP1P3 (A156T K_i 74.3 nM) retained 3-fold better enzymatic activity against A156T variant than the parent compound MK-5172 (A156T K_i 251.3 nM). Interestingly, compared to vaniprevir, Dan-mcP2P4, though only slightly better against wild-type protease, was significantly more active against drug-resistant variants. This signifies the importance of the terminal olefin in the P1 moiety where potential π stacking interactions with F154 are possible,[25,37] likely reducing the off-rate of the P1 olefin analogue compared to vaniprevir.

The antiviral activities of macrocyclic compounds in replicon assays against wild-type HCV and resistant variants correlate with protease binding affinities, and further demonstrate enhanced potency of macrocyclic compounds compared to linear analogues (Figure 3). In general, the P1 P3 macrocyclic compounds exhibited better antiviral potencies against wild-type and drug-resistant variants than the corresponding P2 P4 analogues. Against wild-type virus, P1 P3 macrocyclic inhibitors danoprevir, Asu-mcP1P3, and 5172-mcP1P3 exhibited antiviral potencies in the subnanomolar range with IC_{50} of 0.24, 0.23, and 0.26 nM, respectively. The antiviral potency of Asu-mcP1P3 was significantly better than that of the parent drug, showing a 4-fold enhancement against wild-type and 4 16-fold against resistant variants. Interestingly, 5172-mcP1P3 (WT IC_{50} 0.26 nM) was almost equipotent to MK-5172 (WT IC_{50} 0.29 nM) against wild-type virus and R155K, V36M/R155K, and D168A variants but exhibited 12-fold improved potency against A156T variant (A156T IC_{50} 3.96 nM for 5172-mcP1P3 versus 46.6 nM for MK-5172).

The macrocyclic inhibitors are more potent compared to their linear analogues likely due to the restriction of flexibility of the molecule core around the P2 P3 proline amide bond, which is expected to reduce the conformational entropic penalty for binding thus increasing the overall binding energy.[39,44] The P1 P3 macrocycle restricts the P1 terminus and rotation around the P2 P3 amide bond, which skews the *cis-trans* equilibrium, favoring the *trans* geometry, which is a conserved feature in all protease structures in complex with viral substrate peptides.[37,45] In contrast, the P2 P4 macrocycle restricts the flexibility of the large heterocyclic moiety attached to P2 proline and probably restricts the rotation around the P2 P3 amide bond, in concert with the bulky *tert*-butyl group at P3. In addition, the long hydrophobic chain of the macrocycles reduces the overall polarity of the molecule and makes additional interactions with the protease. Thus, macrocyclization provides several advantages, including preorganization of the molecule in a binding-competent conformation and restriction of the inhibitor geometry to stabilize protease inhibitor interactions.[44]

P2 Moiety Also Determines Inhibitor Susceptibility to Drug Resistance. The susceptibilities of linear and macro-

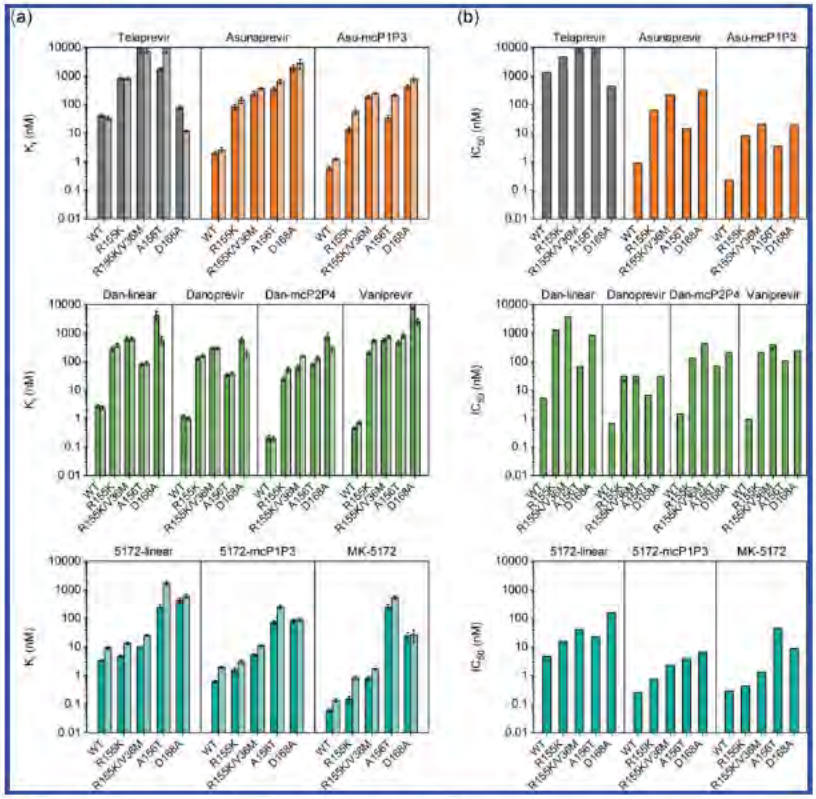


Figure 3. (a) Enzyme inhibition constants for HCV genotype 1a full length NS3/4A and isolated protease domain (dark and light bars) and (b) replicon based half maximal inhibitory concentrations for genotype 1b HCV NS3/4A and drug resistant variants for (first row) telaprevir and asunaprevir series, (second row) danoprevir/vaniprevir series, and (third row) MK 5172 series. Error bars represent standard errors of the mean (n 4); * indicates K_i and IC_{50} values are greater than the highest inhibitor concentration tested.

cyclic protease inhibitors to drug resistance were evaluated by normalizing K_i and IC_{50} values with respect to the corresponding wild-type enzyme (Figure 4) to better ascertain the relative changes. The K_i's against drug-resistant variants generally correlated with IC_{50}'s, but the fold potency losses in replicon assays were in most cases lower than those observed from in vitro enzymatic assays. However, we focused on the fold changes in enzymatic inhibition constants as these reflect the direct effects of drug resistance mutations on inhibitor binding, while antiviral potencies in cellular replicon assays may reflect changes in inhibitor binding when the protease is part of an active replication complex.

Within each inhibitor series of same P2 moiety but different macrocyclization (Figure 4, individual rows) the drug resistance profiles are similar, indicating that the identity of the P2 moiety largely determines the susceptibility to drug resistance. Notably, activity losses for the linear and P1 P3 macrocyclic analogues were similar against all resistant variants. The P2 P4 macrocyclic inhibitors displayed higher fold losses compared to those of linear and P1 P3 macrocyclic compounds against A156T variants and, to a small extent, D168A variants, but the overall resistance profiles are similar in each series. Thus, regardless of the macrocyclization, inhibitors with different P2 moieties exhibit distinct resistance profiles.

The P2 P4 macrocyclic inhibitors are more susceptible to drug resistance possibly because the P2 moiety is constrained to allow stronger interactions with the protein, reducing the inhibitor's ability to adapt to changes in binding environment due to mutations. In all inhibitors, the large heterocyclic P2 moieties bind to the protease in the S2 subsite but vary in how they interact with the primary sites of drug resistance mutations: R155, A156, and D168.[37,38] These mutations lead to disruption of stacking interactions (R155K), pronounced steric clashes (A156T), or disruption of the electrostatic network around the S2 subsite (D168A), but the rigidity of the P2 P4 macrocycle prevents the inhibitor from adjusting to these changes.[38] In contrast, the P1 P3 macrocyclic inhibitors are less susceptible to drug resistance as the location of the

dx.doi.org/10.1021/cb400100g | ACS Chem. Biol. 2013, 8, 1469–1478

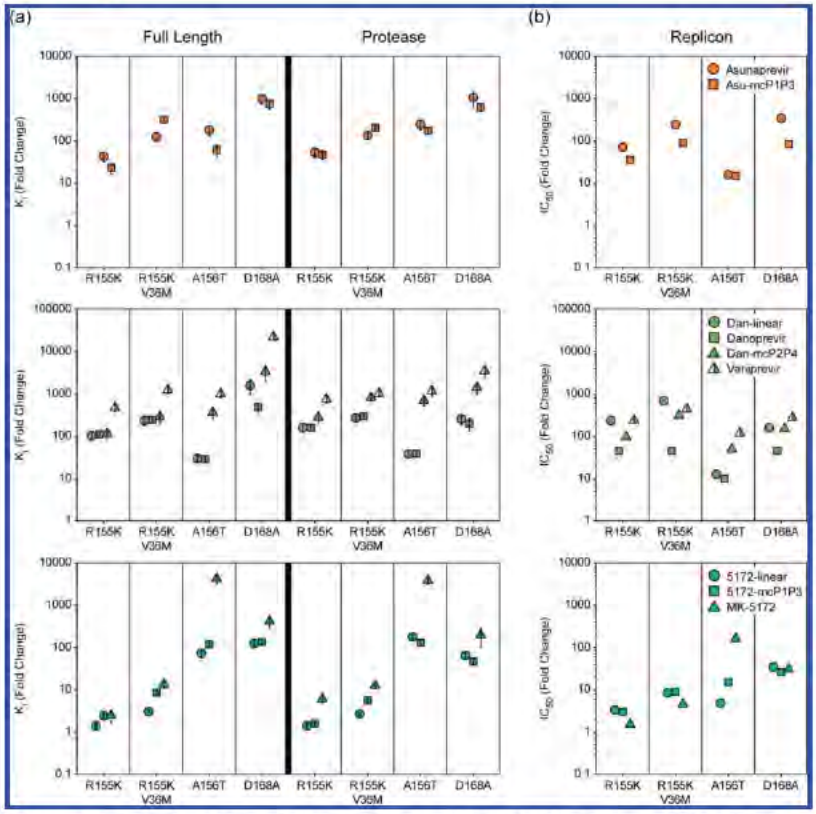


Figure 4. Resistance profiles of protease inhibitors in (a) enzyme inhibition and (b) cell based replicon assays for (first row) asunaprevir series, (second row) danoprevir/vaniprevir series, and (third row) MK 5172 series. Enzyme inhibitory and antiviral activities against mutants were normalized with respect to the wild type full length HCV NS3/4A protease, NS3/4A protease domain, or wild type HCV clone. Error bars represent propagated standard errors.

macrocyclization is distal from the sites of primary drug resistance mutations. Thus, the P1 P3 macrocyclization, although stabilizing the inhibitor's interactions with the protease, allows the P2 moiety to be flexible.

Specific interactions of the P2 moiety with known positions of drug-resistant mutations explain the mechanism behind their distinct resistance profiles. In the danoprevir/vaniprevir series, the aromatic ring of P2 isoindoline stacks on R155 side chain, making favorable cation-π stacking interactions.[38] As R155 conformation is stabilized by hydrogen bond interactions with D168; mutations at R155 and D168 destabilize this packing thereby lowering the affinity of these inhibitors.[38] For asunaprevir and analogues, although there is no reported co-crystal structure, the substituted isoquinoline moiety is expected to interact similarly with R155 and D168. Future co-crystal structures will shed more light on the molecular mechanisms of resistance for these inhibitors.

The in vitro resistance profiles of NS3/4A protease inhibitors generally correlate with resistance observed in patients failing protease inhibitor therapy. Asunaprevir resistance-associated variants mainly include protease variants D168A/E/V/Y,[46] in agreement with significant activity loss against D168A variant in vitro. Danoprevir resistance is primarily associated with R155K variant,[29] while vaniprevir selects for both R155 and D168 variants.[47] In in vitro assays, danoprevir was highly susceptible to R155K mutation but maintained relatively better activity against A156T and D168A variants, while vaniprevir was highly susceptible to all three variants.

We have previously shown that MK-5172 retains relatively better potency against R155K and D168A variants but is highly susceptible to A156T variant, due to steric clashes with the P2 P4 macrocycle.[38] Both the linear and P1 P3 macrocyclic analogues of MK-5172 retain better potency against drug-resistant variants compared to the parent compound. In fact, the P1 P3 macrocyclic analogue 5172-mcP1P3 displays single

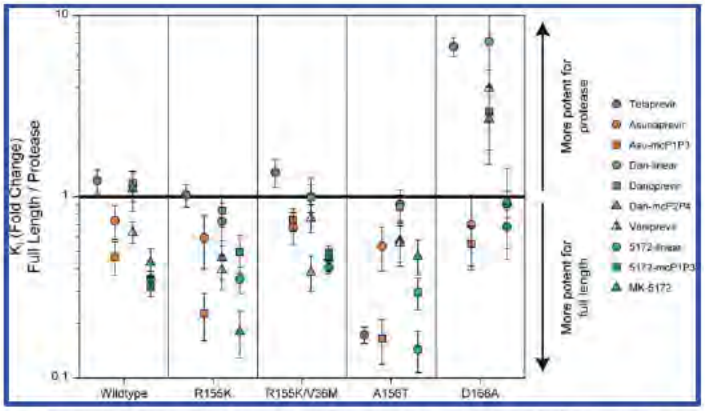


Figure 5. Fold change in inhibitor potencies against full length NS3/4A compared to the isolated HCV NS3/4A protease domain; values lower than 1 indicate higher inhibitor potency against the full length protein compared to protease alone and *vice versa*.

digit nanomolar antiviral potency against all four drug-resistant HCV variants. Thus, compared to MK-5172, 5172-mcP1P3 exhibits a flatter resistance profile. Inhibitors with similar or slightly reduced potency but flatter resistance profiles may prove to be better than those with high potency against wild-type and pivotal resistant variants but reduced activity against other variants such as A156T. These variants may be less competent for replication but could select for secondary mutations that restore viral fitness, rendering the inhibitors highly susceptible to these variants, e.g., R155K versus R155K/V36M.

The activity profile of MK-5172 analogues suggest that even in the absence of P2 P4 macrocycle, the P2 quinoxaline moiety maintains the unique binding conformation observed in MK-5172.[38] In fact, the flexible P2 quinoxaline moiety lacking the P2 P4 macrocycle can be easily modified at the 3-position for structure activity exploration and, together with changes at the P4 capping group, provide opportunities for improving potency against drug-resistant variants and for optimizing pharmacokinetic properties.

Helicase Domain Modulates Protease Inhibitor Potency. The overall drug resistance profiles for the full-length enzyme versus the isolated protease domain did not change significantly (Figure 4). Thus, relative to the effect of drug-resistant mutations, the helicase does not strongly modulate the protease inhibitor activity. Nevertheless, up to an-order-of-magnitude variations in inhibitor potency were observed relative to the isolated protease domain (Figure 5). This observed modulation by the helicase domain was dependent on the identity of the P2 moiety.

For the wild-type enzyme, modulation by the helicase domain was minimal (up to 3-fold preference for the full-length protein). However, with drug-resistant variants R155K and A156T, preference for the full-length protein was more pronounced, with up to 8-fold enhancement in activity for asunaprevir, MK-5172, and analogues, whereas limited modulation was observed for danoprevir and analogues. This trend was not observed for D168A for the asunaprevir and MK-5172 series. Conversely, danoprevir and analogues exhibited significant preference for the isolated protease domain (up to an-order-of-magnitude). Even though none of these protease inhibitors were designed to make any interactions with the helicase domain, the presence of the helicase domain is able to influence their potency. In a full-length NS3/4A crystal structure with an inhibitor similar to danoprevir, the heterocyclic P2 moiety occupies the groove formed between the helicase and the protease active site.[25] Although no specific interactions between the inhibitor P2 moiety and the helicase were observed, solvent exclusion from the otherwise exposed hydrophobic P2 extension may be the reason for the observed potency increase in the asunaprevir and MK-5172 classes. Recently, a new class of allosteric inhibitors were reported, which were shown to interact specifically with the helicase groove to lock the protein in the crystallographic conformation, disallowing substrate access and protein activation by rotation of the helicase domain away from the protease.[48] Further optimization of the heterocyclic P2 moiety to specifically contact the helicase domain may provide inhibitors with better activities and less susceptibility to drug resistance without sacrificing inhibitor potency.

Conclusions. We have found profound influences of macrocyclization status and P2 moieties on the potency of NS3/4A protease inhibitors and their susceptibility to drug resistance. Macrocyclization significantly enhances inhibitor potency, but location of the macrocycle is critical in avoiding drug resistance. Our results provide strong rationale for (1) using P1 P3 macrocyclization to restrict inhibitor geometry and stabilize interactions with the protease thereby improving inhibitor potency and (2) using a flexible P2 moiety to accommodate mutations in the S2 subsite thereby retaining potency against drug-resistant variants. Specifically, protease inhibitors with flexible P2 moieties that avoid direct contact with the S2 subsite, but instead exploit interactions with the essential catalytic residues, are highly promising to retain potency against drug-resistant HCV variants. Additionally, modulation of protease inhibitor potency by the helicase domain suggests that interactions with the helicase domain could be leveraged in future inhibitor development. These strategies may lead to the design of more potent HCV NS3/4A protease inhibitors that are robust against drug resistance.

METHODS

NS3/4A Protease Constructs. The HCV genotype 1a NS3/4A protease domain described previously in a Bristol Myers Squibb patent[49] was synthesized (GenScript) and doned into a pET28a bacterial expression vector. The single chain NS3/4A protease (scNS3/4Apro) contains the protease domain (181 amino acids) and the 12 aa cofactor beta sheet of the NS4A (residues 12−23) tethered to the N terminus of the protease. The codon optimized genotype 1a NS3 helicase (H77, provided by David Frick) was cloned downstream of the protease domain to generate the full length NS3/4A construct. Similar to the protease domain, the single chain full length NS3/4A construct (scNS3/4Afl) contains the NS4A cofactor peptide (residues 12−23) ligated to the N terminus of the protease. Multi drug resistant variants R155K, V36M/R155K A156T, and D168A for both the isolated protease and full length constructs were generated using site directed mutagenesis kit QuikChange II (Stratagene).

Expression and Purification of NS3/4A Protease Constructs. The expression and purification scheme for the isolated protease domain is detailed elsewhere.[37] Briefly, transformed BL21(DE3) *E. coli* expression cells were grown to an A_{600} of 0.6, induced with 1 mM isopropyl 1 thio β D galactopyranoside (IPTG) and incubated with shaking for 5 h at 37 °C. Cells expressing the protein were harvested by centrifugation and stored at −80 °C. Frozen pellets were resuspended in resuspension buffer (50 mM phosphate buffer, 500 mM NaCl, 10% glycerol, 2 mM β mercaptoethanol (β ME), pH 7.5), lysed, and centrifuged to pellet the cell debris, and the resulting supernatant was applied to a nickel column (Qiagen). The column was washed with resuspension buffer, and the protein was eluted with resuspension buffer, containing 200 mM imidazole. The eluate was supplemented with thrombin and dialyzed overnight to deave the His tag and remove the imidazole. The purified protein was flash frozen in liquid nitrogen and stored at −80 °C.

For the expression and the purification of the full length NS3/4A, transformed BL21(DE3) *E. coli* expression cells were grown to an A_{600} of 0.6 at 37 °C, transferred to 20 °C, induced by 0.5 mM IPTG, and incubated with shaking for 4 h. Cells were then harvested via centrifugation; cell pellets were washed with 1X phosphate buffered saline (PBS), repelleted, and stored at −80 °C. Frozen pellets were resuspended in Buffer HT (25 mM HEPES, 500 mM NaCl, 10% glycerol, 0.1% OβG, 2 mM TCEP and 20 mM imidazole, pH 8.0) supplemented with DNase I (Roche) and homogenized using a cell disruptor (Micro Fluidics). Lysed cells were centrifuged to clear the cell debris and applied to a 1 mL HisTrap HP column (GE Life Sciences) using an AKTA Purifier (GE Life Sciences). The protein was washed with Buffer HT supplemented with 40 mM imidazole and eluted with Buffer HT supplemented with 250 mM imidazole. The eluate was dialyzed overnight against Buffer S (25 mM MES, 150 mM NaCl, 10% glycerol, 0.1% OβG, 2 mM TCEP, pH 6.0). Dialyzed protein was applied to a Mono S Column (GE Life Sciences) and eluted with a linear gradient of NaCl up to 1 M. The eluate was judged >90% pure by polacrylamide gel electrophoresis, concentrated, flash frozen, and stored at −80 °C.

Determination of Michaelis−Menten Constants. Protease cleavage assays were performed to determine Michaelis−Menten constants (K_M), which were then used to obtain inhibition constants through Morrison's equation (see below). The assays were done in a final volume of 60 μL containing protease assay buffer (50 mM Tris, 2.5% glycerol, 0.1% OβG, 5 mM TCEP, 1% DMSO, pH 7.5) and up to 20 μM concentration of a modified HCV NS3/4A protease substrate Ac DE Dap(QXL 520) EE Abu ψ [COO]AS C(5 FAMsp) NH2 (Anaspec) in black 96 well flat bottom nonbinding surface half area plates (Corning) at RT. The reaction was initiated by the rapid injection of 10 μL of HCV NS3/4A protease to a final concentration of 20 nM. The fluorescence output from the depsipeptide cleavage product was measured kinetically for at least 1 h using an EnVision plate reader (Perkin Elmer) with excitation wavelength at 485 nm and emission at 530 nm. End point readings were also taken after 6 h. For each construct, three independent experiments were performed.

Inner filter effect in higher concentrations of substrate was corrected using the following scheme: End point readings were subtracted from the background fluorescence values to obtain reaction amplitudes, which were plotted against substrate concentration. The initial linear portion of this curve was fitted with linear regression and extrapolated to maximum substrate concentration to obtain the ideal amplitude line. Corresponding amplitude values in the ideal amplitude line were divided by the actual amplitudes to obtain inner filter effect correction parameters which were then multiplied by initial velocities to obtain corrected velocities. Corrected velocities were plotted against substrate concentrations, and all replicates were fitted to Michaelis−Menten equation globally to obtain Michaelis−Menten constants (K_M), sharing K_M between different data sets. Except full length A156T mutant, K_M values for all the constructs were similar, which signifies that the substrate recognition was not altered significantly with respect to different multi drug resistant mutations and helicase mutations.

Enzyme Inhibition Assays. All enzyme inhibition assays were performed in nonbinding surface 96 well black half area plates (Corning) in a reaction volume of 60 μL. The NS3/4A protease (~1 nM) was preincubated with increasing concentration of drugs in protease reaction buffer for an hour. The reaction was initiated by the rapid injection of 5 μL of HCV NS3/4A protease substrate to a final concentration of 200 nM and kinetically monitored using a Perkin Elmer EnVision plate reader (excitation at 485 nm; emission at 530 nm) for 1 h with 30 s between data points. At least four independent data sets were collected for each inhibitor with each protease construct. Each inhibitor titration included at least 12 inhibitor concentration points, which were globally fit to the Morrison equation to obtain the K_i value.

Initial velocities were obtained from the progress curves and plotted against inhibitor concentrations to get inhibition curves. Resulting curves were fitted to Morrison's equation:

$$\frac{V_i}{V_0} = 1 - [\{([E]_T + [I]_T + K_i^{app}) - \sqrt{([E]_T + [I]_T + K_i^{app})^2 - 4[E]_T[I]_T}\}/2[E]_T]$$

$$K_i^{app} = K_i\left(1 + \frac{[S]}{K_M}\right)$$

where $[E]_T$ is the total enzyme concentration, $[I]_T$ is the total inhibitor concentration, $[S]$ is the substrate concentration, K_M is the Michaelis−Menten constant (obtained from protease deavage assays), V_0 is the initial velocity when there is no drug, V_i is the initial velocity at $[I]_T$, K_i^{app} is the apparent inhibition constant, and K_i is the inhibition constant. Nonlinear regression analyses were performed where independent replicates from different protease constructs with different drugs were fitted globally, sharing K_i.

Cell-Based Drug Susceptibility Assays. Mutations (R155K, V36M/R155K D168A, or A156T) were introduced into the NS3 region of genotype 1b HCV Con1 luciferase reporter replicon using the mega primer method of mutagenesis.[50] Replicon RNA of each protease variant was introduced into Huh7 cells by electroporation. Replication was then assessed in the presence of increasing concentrations of protease inhibitors (telaprevir, asunaprevir, danoprevir, vaniprevir, MK 5172 and respective analogues) by measuring luciferase activity (relative light units) 96 h after electroporation. The drug concentrations required to inhibit replicon replication by 50% (IC_{50}) were calculated directly from the drug inhibition curves.

ASSOCIATED CONTENT

Supporting Information

This material is available free of charge *via* the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: celia.schiffer@umassmed.edu.

Author Contributions

[$]These authors contributed equally to this work.

Notes

The authors declare the following competing financial interest(s): A.N., C.J.P., and W.H. are employees of Monogram Biosciences. C.A.S. has received small "Sponsored Research Grants" from Avila (CelGene) and Merck; however, the work described in this manuscript is supported only by the National Institutes of Health. This does not alter our adherence to all ACS policies on sharing data and materials..

ACKNOWLEDGMENTS

This work was supported by the National Institute of Allergy and Infectious Diseases of the NIH (R01-AI085051). We thank members of the Schiffer laboratory for helpful discussions and N. Kurt-Yilmaz for editorial assistance.

REFERENCES

(1) World Health Organization (WHO). Hepatitis C, Fact Sheet No 164 (July 2012). http://www.who.int/mediacentre/factsheets/fs164/en/index.html; accessed February 2013.

(2) Ghany, M. G., Nelson, D. R., Strader, D. B., Thomas, D. L., and Seeff, L. B. (2011) An update on treatment of genotype 1 chronic hepatitis C virus infection: 2011 practice guideline by the American Association for the Study of Liver Diseases. *Hepatology 54*, 1433−1444.

(3) McHutchison, J. G., Everson, G. T., Gordon, S. C., Jacobson, I. M., Sulkowski, M., Kauffman, R., McNair, L., Alam, J., and Muir, A. J. (2009) Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N. Engl. J. Med. 360*, 1827−1838.

(4) Hézode, C., Forestier, N., Dusheiko, G., Ferenci, P., Pol, S., Goeser, T., Bronowicki, J. P., BourliÃre, M., Gharakhanian, S., Bengtsson, L., McNair, L., George, S., Kieffer, T., Kwong, A., Kauffman, R. S., Alam, J., Pawlotsky, J. M., and Zeuzem, S. (2009) Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N. Engl. J. Med. 360*, 1839−1850.

(5) Kwo, P. Y., Lawitz, E. J., McCone, J., Schiff, E. R., Vierling, J. M., Pound, D., Davis, M. N., Galati, J. S., Gordon, S. C., Ravendhran, N., Rossaro, L., Anderson, F. H., Jacobson, I. M., Rubin, R., Koury, K., Pedicone, L. D., Brass, C. A., Chaudhri, E., and Albrecht, J. K. (2010) Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa 2b and ribavirin in treatment naive patients with genotype 1 hepatitis C infection (SPRINT 1): an open label, randomised, multicentre phase 2 trial. *Lancet 376*, 705−716.

(6) Poordad, F., McCone, J., Bacon, B. R., Bruno, S., Manns, M. P., Sulkowski, M. S., Jacobson, I. M., Reddy, K. R., Goodman, Z. D., Boparai, N., DiNubile, M. J., Sniukiene, V., Brass, C. A., Albrecht, J. K., and Bronowicki, J. P. (2011) Boceprevir for untreated chronic HCV genotype 1 infection. *N. Engl. J. Med. 364*, 1195−1206.

(7) Jacobson, I. M., McHutchison, J. G., Dusheiko, G., Di Bisceglie, A. M., Reddy, K. R., Bzowej, N. H., Marcellin, P., Muir, A. J., Ferenci, P., Flisiak, R., George, J., Rizzetto, M., Shouval, D., Sola, R., Terg, R. A., Yoshida, E. M., Adda, N., Bengtsson, L., Sankoh, A. J., Kieffer, T. L., George, S., Kauffman, R. S., and M.D, S. Z. (2011) Telaprevir for previously untreated chronic hepatitis C virus infection. *N. Engl. J. Med. 364*, 2405−2416.

(8) Ge, D., Fellay, J., Thompson, A. J., Simon, J. S., Shianna, K. V., Urban, T. J., Heinzen, E. L., Qiu, P., Bertelsen, A. H., Muir, A. J., Sulkowski, M., McHutchison, J. G., and Goldstein, D. B. (2009) Genetic variation in IL28B predicts hepatitis C treatment induced viral clearance. *Nature 461*, 399−401.

(9) Fried, M. W. (2002) Side effects of therapy of hepatitis C and their management. *Hepatology 36*, S237−S244.

(10) Sarrazin, C., and Zeuzem, S. (2010) Resistance to direct antiviral agents in patients with hepatitis C virus infection. *Gastroenterology 138*, 447−462.

(11) Hofmann, W. P., and Zeuzem, S. (2011) A new standard of care for the treatment of chronic HCV infection. *Nat. Rev. Gastroenterol. Hepatol. 8*, 257−264.

(12) Kolykhalov, A. A., Mihalik, K., Feinstone, S. M., and Rice, C. M. (2000) Hepatitis C virus encoded enzymatic activities and conserved RNA elements in the 3′ nontranslated region are essential for virus replication in vivo. *J. Virol. 74*, 2046−2051.

(13) Morikawa, K., Lange, C. M., Gouttenoire, J., Meylan, E., Brass, V., Penin, F., and Moradpour, D. (2011) Nonstructural protein 3−4A: the Swiss army knife of hepatitis C virus. *J. Viral Hepat. 18*, 305−315.

(14) Foy, E., Li, K., Wang, C., Sumpter, R., Jr., Ikeda, M., Lemon, S. M., and Gale, M., Jr. (2003) Regulation of interferon regulatory factor 3 by the hepatitis C virus serine protease. *Science 300*, 1145−1148.

(15) Li, K., Foy, E., Ferreon, J. C., Nakamura, M., Ferreon, A. C., Ikeda, M., Ray, S. C., Gale, M., Jr., and Lemon, S. M. (2005) Immune evasion by hepatitis C virus NS3/4A protease mediated cleavage of the Toll like receptor 3 adaptor protein TRIF. *Proc. Natl. Acad. Sci. U.S.A. 102*, 2992−2997.

(16) Chen, Z., Benureau, Y., Rijnbrand, R., Yi, J., Wang, T., Warter, L., Lanford, R. E., Weinman, S. A., Lemon, S. M., Martin, A., and Li, K. (2007) GB virus B disrupts RIG I signaling by NS3/4A mediated cleavage of the adaptor protein MAVS. *J. Virol. 81*, 964−976.

(17) Malcolm, B. A., Liu, R., Lahser, F., Agrawal, S., Belanger, B., Butkiewicz, N., Chase, R., Gheyas, F., Hart, A., Hesk, D., Ingravallo, P., Jiang, C., Kong, R., Lu, J., Pichardo, J., Prongay, A., Skelton, A., Tong, X., Venkatraman, S., Xia, E., Girijavallabhan, V., and Njoroge, F. G. (2006) SCH 503034, a mechanism based inhibitor of hepatitis C virus NS3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells. *Antimicrob. Agents Chemother. 50*, 1013−1020.

(18) Perni, R. B., Almquist, S. J., Byrn, R. A., Chandorkar, G., Chaturvedi, P. R., Courtney, L. F., Decker, C. J., Dinehart, K., Gates, C. A., Harbeson, S. L., Heiser, A., Kalkeri, G., Kolaczkowski, E., Lin, K., Luong, Y. P., Rao, B. G., Taylor, W. P., Thomson, J. A., Tung, R. D., Wei, Y., Kwong, A. D., and Lin, C. (2006) Preclinical profile of VX 950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3 4A serine protease. *Antimicrob. Agents Chemother. 50*, 899−909.

(19) McPhee, F., Sheaffer, A. K., Friborg, J., Hernandez, D., Falk, P., Zhai, G., Levine, S., Chaniewski, S., Yu, F., Barry, D., Chen, C., Lee, M. S., Mosure, K., Sun, L. Q., Sinz, M., Meanwell, N. A., Colonno, R. J., Knipe, J., and Scola, P. (2012) Preclinical profile and characterization of the hepatitis C virus NS3 protease inhibitor asunaprevir (BMS 650032). *Antimicrob. Agents Chemother. 56*, 5387−5396.

(20) White, P. W., Llinàs Brunet, M., Amad, M. a., Bethell, R. C., Bolger, G., Cordingley, M. G., Duan, J., Garneau, M., Lagacé, L., Thibeault, D., and Kukolj, G. (2010) Preclinical characterization of BI 201335, a C terminal carboxylic acid inhibitor of the hepatitis C virus NS3 NS4A protease. *Antimicrob. Agents Chemother. 54*, 4611−4618.

(21) Seiwert, S. D., Andrews, S. W., Jiang, Y., Serebryany, V., Tan, H., Kossen, K., Rajagopalan, P. T., Misialek, S., Stevens, S. K., Stoycheva, A., Hong, J., Lim, S. R., Qin, X., Rieger, R., Condroski, K. R., Zhang, H., Do, M. G., Lemieux, C., Hingorani, G. P., Hartley, D. P., Josey, J. A., Pan, L., Beigelman, L., and Blatt, L. M. (2008) Preclinical characteristics of the hepatitis C virus NS3/4A protease inhibitor ITMN 191 (R7227). *Antimicrob. Agents Chemother. 52*, 4432−4441.

(22) Lin, T. I., Lenz, O., Fanning, G., Verbinnen, T., Delouvroy, F., Scholliers, A., Vermeiren, K., Rosenquist, A., Edlund, M., Samuelsson, B., Vrang, L., de Kock, H., Wigerinck, P., Raboisson, P., and Simmen, K. (2009) In vitro activity and preclinical profile of TMC435350, a potent hepatitis C virus protease inhibitor. *Antimicrob. Agents Chemother. 53*, 1377−1385.

(23) Liverton, N. J., Carroll, S. S., DiMuzio, J., Fandozzi, C., Graham, D. J., Hazuda, D., Holloway, M. K., Ludmerer, S. W., McCauley, J. A., McIntyre, C. J., Olsen, D. B., Rudd, M. T., Stahlhut, M., and Vacca, J.

P. (2010) MK 7009, a potent and selective inhibitor of hepatitis C virus NS3/4A protease. *Antimicrob. Agents Chemother. 54*, 305−311.

(24) Harper, S., McCauley, J. A., Rudd, M. T., Ferrara, M., DiFilippo, M., Crescenzi, B., Koch, U., Petrocchi, A., Holloway, M. K., Butcher, J. W., Romano, J. J., Bush, K. J., Gilbert, K. F., McIntyre, C. J., Nguyen, K. T., Nizi, E., Carroll, S. S., Ludmerer, S. W., Burlein, C., DiMuzio, J. M., Graham, D. J., McHale, C. M., Stahlhut, M. W., Olsen, D. B., Monteagudo, E., Cianetti, S., Giuliano, C., Pucci, V., Trainor, N., Fandozzi, C. M., Rowley, M., Coleman, P. J., Vacca, J. P., Summa, V., and Liverton, N. J. (2012) Discovery of MK 5172, a macrocyclic hepatitis C virus NS3/4a protease inhibitor. *ACS Med. Chem. Lett. 3*, 332−336.

(25) Schiering, N., D'Arcy, A., Villard, F., Simić, O., Kamke, M., Monnet, G., Hassiepen, U., Svergun, D. I., Pulfer, R., Eder, J., Raman, P., and Bodendorf, U. (2011) A macrocyclic HCV NS3/4A protease inhibitor interacts with protease and helicase residues in the complex with its full length target. *Proc. Natl. Acad. Sci. U.S.A. 108*, 21052−21056.

(26) Hinrichsen, H., Benhamou, Y., Wedemeyer, H., Reiser, M., Sentjens, R. E., Calleja, J. L., Forns, X., Erhardt, A., Cronlein, J., Chaves, R. L., Yong, C. L., Nehmiz, G., and Steinmann, G. G. (2004) Short term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients. *Gastroenterology 127*, 1347−1355.

(27) Reesink, H. W., Zeuzem, S., Weegink, C. J., Forestier, N., van Vliet, A., van de Wetering de Rooij, J., McNair, L., Purdy, S., Kauffman, R., Alam, J., and Jansen, P. L. (2006) Rapid decline of viral RNA in hepatitis C patients treated with VX 950: a phase Ib, placebo controlled, randomized study. *Gastroenterology 131*, 997−1002.

(28) Manns, M. P., Bourlière, M., Benhamou, Y., Pol, S., Bonacini, M., Trepo, C., Wright, D., Berg, T., Calleja, J., White, P. W., Stern, J. O., Steinmann, G., Yong, C. L., Kukolj, G., Scherer, J., and Boecher, W. O. (2011) Potency, safety, and pharmacokinetics of the NS3/4A protease inhibitor BI201335 in patients with chronic HCV genotype 1 infection. *J. Hepatol. 54*, 1114−1122.

(29) Lim, S. R., Qin, X., Susser, S., Nicholas, J. B., Lange, C., Herrmann, E., Hong, J., Arfsten, A., Hooi, L., Bradford, W., Nájera, I., Smith, P., Zeuzem, S., Kossen, K., Sarrazin, C., and Seiwert, S. D. (2012) Virologic escape during danoprevir (ITMN 191/RG7227) monotherapy is hepatitis C virus subtype dependent and associated with R155K substitution. *Antimicrob. Agents Chemother. 56*, 271−279.

(30) Manns, M., Reesink, H., Berg, T., Dusheiko, G., Flisiak, R., Marcellin, P., Moreno, C., Lenz, O., Meyvisch, P., Peeters, M., Sekar, V., Simmen, K., and Verloes, R. (2011) Rapid viral response of once daily TMC435 plus pegylated interferon/ribavirin in hepatitis C genotype 1 patients: a randomized trial. *Antiviral Ther. 16*, 1021−1033.

(31) Bacon, B. R., Gordon, S. C., Lawitz, E., Marcellin, P., Vierling, J. M., Zeuzem, S., Poordad, F., Goodman, Z. D., Sings, H. L., Boparai, N., Burroughs, M., Brass, C. A., Albrecht, J. K., and Esteban, R. (2011) Boceprevir for previously treated chronic HCV genotype 1 infection. *N. Engl. J. Med. 364*, 1207−1217.

(32) Zeuzem, S., Andreone, P., Pol, S., Lawitz, E., Diago, M., Roberts, S., Focaccia, R., Younossi, Z., Foster, G. R., Horban, A., Ferenci, P., Nevens, F., MÃ1/4llhaupt, B., Pockros, P., Terg, R., Shouval, D., van Hoek, B., Weiland, O., Van Heeswijk, R., De Meyer, S., Luo, D., Boogaerts, G., Polo, R., Picchio, G., and Beumont, M. (2011) Telaprevir for retreatment of HCV infection. *N. Engl. J. Med. 364*, 2417−2428.

(33) Rong, L., Dahari, H., Ribeiro, R. M., and Perelson, A. S. (2010) Rapid emergence of protease inhibitor resistance in hepatitis C virus. *Sci. Transl. Med. 2*, 30ra32.

(34) Kieffer, T. L., Kwong, A. D., and Picchio, G. R. (2010) Viral resistance to specifically targeted antiviral therapies for hepatitis C (STAT Cs). *J. Antimicrob. Chemother. 65*, 202−212.

(35) Sarrazin, C., Kieffer, T. L., Bartels, D., Hanzelka, B., Muh, U., Welker, M., Wincheringer, D., Zhou, Y., Chu, H. M., Lin, C., Weegink, C., Reesink, H., Zeuzem, S., and Kwong, A. D. (2007) Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology 132*, 1767−1777.

(36) Susser, S., Welsch, C., Wang, Y., Zettler, M., Domingues, F. S., Karey, U., Hughes, E., Ralston, R., Tong, X., Herrmann, E., Zeuzem, S., and Sarrazin, C. (2009) Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus infected patients. *Hepatology 50*, 1709−1718.

(37) Romano, K. P., Ali, A., Royer, W. E., and Schiffer, C. A. (2010) Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding. *Proc. Natl. Acad. Sci. U.S.A. 107*, 20986−20991.

(38) Romano, K. P., Ali, A., Aydin, C., Soumana, D., Özen, A., Deveau, L. M., Silver, C., Cao, H., Newton, A., Petropoulos, C. J., Huang, W., and Schiffer, C. A. (2012) The molecular basis of drug resistance against hepatitis C virus NS3/4A protease inhibitors. *PLOS Pathog. 8*, e1002832.

(39) Tsantrizos, Y. S., Bolger, G., Bonneau, P., Cameron, D. R., Goudreau, N., Kukolj, G., LaPlante, S. R., Llinas Brunet, M., Nar, H., and Lamarre, D. (2003) Macrocyclic inhibitors of the NS3 protease as potential therapeutic agents of hepatitis C virus infection. *Angew. Chem., Int. Ed. Engl. 42*, 1356−1360.

(40) Liverton, N. J., Holloway, M. K., McCauley, J. A., Rudd, M. T., Butcher, J. W., Carroll, S. S., DiMuzio, J., Fandozzi, C., Gilbert, K. F., Mao, S. S., McIntyre, C. J., Nguyen, K. T., Romano, J. J., Stahlhut, M., Wan, B. L., Olsen, D. B., and Vacca, J. P. (2008) Molecular modeling based approach to potent P2 P4 macrocyclic inhibitors of hepatitis C NS3/4A protease. *J. Am. Chem. Soc. 130*, 4607−4609.

(41) Sin, N., Venables, B. L., Scola, P. M., and Wang, A. X. (2009) Hepatitis C virus inhibitors, U.S. Patent 2009/0274652.

(42) Wang, X. A., Sun, L. Q., Sit, S. Y., Sin, N., Scola, P. M., Hewawasam, P., Good, A., C., Chen, Y., and Campbell, A. (2006) Hepatitis C virus inhibitors, U.S. Patent 6995174.

(43) McCauley, J. A., McIntyre, C. J., Rudd, M. T., Nguyen, K. T., Romano, J. J., Butcher, J. W., Gilbert, K. F., Bush, K. J., Holloway, M. K., Swestock, J., Wan, B. L., Carroll, S. S., DiMuzio, J. M., Graham, D. J., Ludmerer, S. W., Mao, S. S., Stahlhut, M. W., Fandozzi, C. M., Trainor, N., Olsen, D. B., Vacca, J. P., and Liverton, N. J. (2010) Discovery of vaniprevir (MK 7009), a macrocyclic hepatitis C virus NS3/4a protease inhibitor. *J. Med. Chem. 53*, 2443−2463.

(44) Marsault, E., and Peterson, M. L. (2011) Macrocycles are great cycles: Applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem. 54*, 1961−2004.

(45) Romano, K. P., Laine, J. M., Deveau, L. M., Cao, H., Massi, F., and Schiffer, C. A. (2011) Molecular mechanisms of viral and host cell substrate recognition by hepatitis C virus NS3/4A protease. *J. Virol. 85*, 6106−6116.

(46) McPhee, F., Hernandez, D., Yu, F., Ueland, J., Monikowski, A., Carifa, A., Falk, P., Wang, C., Fridell, R., Eley, T., Zhou, N., and Gardiner, D. (2013) Resistance analysis of hepatitis C virus genotype 1 prior treatment null responders receiving daclatasvir and asunaprevir. *Hepatology*, DOI: 10.1002/hep.26388.

(47) Manns, M. P., Gane, E., Rodriguez Torres, M., Stoehr, A., Yeh, C. T., Marcellin, P., Wiedmann, R. T., Hwang, P. M., Caro, L., Barnard, R. J. O., and Lee, A. W. (2012) Vaniprevir with pegylated interferon alpha 2a and ribavirin in treatment naïve patients with chronic hepatitis C: A randomized phase II study. *Hepatology 56*, 884−893.

(48) Saalau Bethell, S. M., Woodhead, A. J., Chessari, G., Carr, M. G., Coyle, J., Graham, B., Hiscock, S. D., Murray, C. W., Pathuri, P., Rich, S. J., Richardson, C. J., Williams, P. A., and Jhoti, H. (2012) Discovery of an allosteric mechanism for the regulation of HCV NS3 protein function. *Nat. Chem. Biol. 8*, 920−925.

(49) Wittekind, M., Weinheimer, S., Zhang, Y., and Goldfarb, V. (2002) Modified forms of hepatitis C NS3 protease for facilitating inhibitor screening and structural studies of protease:inhibitor complexes, U.S. Patent 6333186.

(50) Sarkar, G., and Sommer, S. S. (1990) The "megaprimer" method of site directed mutagenesis. *Biotechniques 8*, 404−407.

References

[1] WHO, "Guidelines for the screening, care and treatment of persons with hepatitis c infection," Report ISBN: 978 92 4 154875 5, World Health Organization, 2014.

[2] E. Gower, C. Estes, S. Blach, K. Razavi-Shearer, and H. Razavi, "Global epidemiology and genotype distribution of the hepatitis c virus infection," *J Hepatol*, vol. 61, no. 1 Suppl, pp. S45–57, 2014.

[3] P. Simmonds, J. Bukh, C. Combet, G. Deleage, N. Enomoto, S. Feinstone, P. Halfon, G. Inchauspe, C. Kuiken, G. Maertens, M. Mizokami, D. G. Murphy, H. Okamoto, J. M. Pawlotsky, F. Penin, E. Sablon, I. T. Shin, L. J. Stuyver, H. J. Thiel, S. Viazov, A. J. Weiner, and A. Widell, "Consensus proposals for a unified system of nomenclature of hepatitis c virus genotypes," *Hepatology*, vol. 42, no. 4, pp. 962–73, 2005.

[4] J. P. Messina, I. Humphreys, A. Flaxman, A. Brown, G. S. Cooke, O. G. Pybus, and E. Barnes, "Global distribution and prevalence of hepatitis c virus genotypes," *Hepatology*, p. DOI: 10.1002/hep.27259, 2014.

[5] E. Blanchard, S. Belouzard, L. Goueslain, T. Wakita, J. Dubuisson, C. Wychowski, and Y. Rouille, "Hepatitis c virus entry depends on clathrin-mediated endocytosis," *J Virol*, vol. 80, no. 14, pp. 6964–72, 2006.

[6] D. M. Tscherne, C. T. Jones, M. J. Evans, B. D. Lindenbach, J. A. McKeating, and C. M. Rice, "Time- and temperature-dependent activation of hepatitis c virus for low-ph-triggered entry," *J Virol*, vol. 80, no. 4, pp. 1734–41, 2006.

[7] M. E. Major and S. M. Feinstone, "The molecular virology of hepatitis c," *Hepatology*, vol. 25, no. 6, pp. 1527–38, 1997.

[8] D. Moradpour, F. Penin, and C. M. Rice, "Replication of hepatitis c virus," *Nat Rev Microbiol*, vol. 5, no. 6, pp. 453–63, 2007.

[9] D. Egger, B. Wolk, R. Gosert, L. Bianchi, H. E. Blum, D. Moradpour, and K. Bienz, "Expression of hepatitis c virus proteins induces distinct membrane alterations including a candidate viral replication complex," *J Virol*, vol. 76, no. 12, pp. 5974–84, 2002.

[10] M. Chang, O. Williams, J. Mittler, A. Quintanilla, R. L. J. Carithers, J. Perkins, L. Corey, and D. R. Gretch, "Dynamics of hepatitis c virus replication in human liver.," *Am J Pathol*, vol. 163, pp. 433–444, Aug 2003.

[11] E. Santolini, G. Migliaccio, and N. La Monica, "Biosynthesis and biochemical properties of the hepatitis c virus core protein.," *J Virol*, vol. 68, pp. 3631–3641, Jun 1994.

[12] A. Goffard and J. Dubuisson, "Glycosylation of hepatitis c virus envelope proteins.," *Biochimie*, vol. 85, pp. 295–301, Mar-Apr 2003.

[13] S. Carrere-Kremer, C. Montpellier-Pala, L. Cocquerel, C. Wychowski, F. Penin, and J. Dubuisson, "Subcellular localization and topology of the p7 polypeptide of hepatitis c virus.," *J Virol*, vol. 76, pp. 3720–3730, Apr 2002.

[14] S. D. C. Griffin, L. P. Beales, D. S. Clarke, O. Worsfold, S. D. Evans, J. Jaeger, M. P. G. Harris, and D. J. Rowlands, "The p7 protein of hepatitis c virus forms an ion channel that is blocked by the antiviral drug, amantadine.," *FEBS Lett*, vol. 535, pp. 34–38, Jan 2003.

[15] D. Pavlovic, D. C. A. Neville, O. Argaud, B. Blumberg, R. A. Dwek, W. B. Fischer, and N. Zitzmann, "The hepatitis c virus p7 protein forms an ion channel that is

inhibited by long-alkyl-chain iminosugar derivatives.," *Proc Natl Acad Sci U S A*, vol. 100, pp. 6104–6108, May 2003.

[16] A. Grakoui, D. W. McCourt, C. Wychowski, S. M. Feinstone, and C. M. Rice, "A second hepatitis c virus-encoded proteinase.," *Proc Natl Acad Sci U S A*, vol. 90, pp. 10583–10587, Nov 1993.

[17] I. C. Lorenz, J. Marcotrigiano, T. G. Dentzer, and C. M. Rice, "Structure of the catalytic domain of the hepatitis c virus ns2-3 protease.," *Nature*, vol. 442, pp. 831–835, Aug 2006.

[18] K. E. Reed and C. M. Rice, "Overview of hepatitis c virus genome structure, polyprotein processing, and protein properties," *Curr Top Microbiol Immunol*, vol. 242, pp. 55–84, 2000.

[19] P. Gallinari, D. Brennan, C. Nardi, M. Brunetti, L. Tomei, C. Steinkuhler, and R. De Francesco, "Multiple enzymatic activities associated with recombinant ns3 protein of hepatitis c virus," *J Virol*, vol. 72, no. 8, pp. 6758–6769, 1998.

[20] A. Kanai, K. Tanabe, and M. Kohara, "Poly(u) binding activity of hepatitis c virus ns3 protein, a putative rna helicase.," *FEBS Lett*, vol. 376, pp. 221–224, Dec 1995.

[21] J. L. Kim, K. A. Morgenstern, C. Lin, T. Fox, M. D. Dwyer, J. A. Landro, S. P. Chambers, W. Markland, C. A. Lepre, E. T. O'Malley, S. L. Harbeson, C. M. Rice, M. A. Murcko, P. R. Caron, and J. A. Thomson, "Crystal structure of the hepatitis c virus ns3 protease domain complexed with a synthetic ns4a cofactor peptide.," *Cell*, vol. 87, pp. 343–355, Oct 1996.

[22] D. L. Sali, R. Ingram, M. Wendel, D. Gupta, C. McNemar, A. Tsarbopoulos, J. W. Chen, Z. Hong, R. Chase, C. Risano, R. Zhang, N. Yao, A. D. Kwong, L. Ra-

manathan, H. V. Le, and P. C. Weber, "Serine protease of hepatitis c virus expressed in insect cells as the ns3/4a complex.," *Biochemistry*, vol. 37, pp. 3392–3401, Mar 1998.

[23] J. Gouttenoire, V. Castet, R. Montserret, N. Arora, V. Raussens, J.-M. Ruysschaert, E. Diesis, H. E. Blum, F. Penin, and D. Moradpour, "Identification of a novel determinant for membrane association in hepatitis c virus nonstructural protein 4b.," *J Virol*, vol. 83, pp. 6257–6268, Jun 2009.

[24] J. Gouttenoire, F. Penin, and D. Moradpour, "Hepatitis c virus nonstructural protein 4b: a journey into unexplored territory.," *Rev Med Virol*, vol. 20, pp. 117–129, Mar 2010.

[25] Y. Huang, K. Staschke, R. De Francesco, and S.-L. Tan, "Phosphorylation of hepatitis c virus ns5a nonstructural protein: a new paradigm for phosphorylation-dependent viral rna replication?," *Virology*, vol. 364, pp. 1–9, Jul 2007.

[26] T. Kawai, K. Takahashi, S. Sato, C. Coban, H. Kumar, H. Kato, K. J. Ishii, O. Takeuchi, and S. Akira, "Ips-1, an adaptor triggering rig-i- and mda5-mediated type i interferon induction," *Nature Immunology*, vol. 6, no. 10, pp. 981–8, 2005.

[27] K. Li, E. Foy, J. C. Ferreon, M. Nakamura, A. C. Ferreon, M. Ikeda, S. C. Ray, J. Gale, M., and S. M. Lemon, "Immune evasion by hepatitis c virus ns3/4a protease-mediated cleavage of the toll-like receptor 3 adaptor protein trif," *Proc Natl Acad Sci U S A*, vol. 102, no. 8, pp. 2992–7, 2005.

[28] R. B. Seth, L. Sun, C. K. Ea, and Z. J. Chen, "Identification and characterization of mavs, a mitochondrial antiviral signaling protein that activates nf-kappab and irf 3," *Cell*, vol. 122, no. 5, pp. 669–82, 2005.

[29] M. Reiser, H. Hinrichsen, Y. Benhamou, H. W. Reesink, H. Wedemeyer, C. Avendano, N. Riba, C.-L. Yong, G. Nehmiz, and G. G. Steinmann, "Antiviral efficacy of ns3-serine protease inhibitor biln-2061 in patients with chronic genotype 2 and 3 hepatitis c.," *Hepatology*, vol. 41, pp. 832–835, Apr 2005.

[30] D. Lamarre, P. C. Anderson, M. Bailey, P. Beaulieu, G. Bolger, P. Bonneau, M. Bos, D. R. Cameron, M. Cartier, M. G. Cordingley, A. M. Faucher, N. Goudreau, S. H. Kawai, G. Kukolj, L. Lagace, S. R. LaPlante, H. Narjes, M. A. Poupart, J. Rancourt, R. E. Sentjens, R. St George, B. Simoneau, G. Steinmann, D. Thibeault, Y. S. Tsantrizos, S. M. Weldon, C. L. Yong, and M. Llinas-Brunet, "An ns3 protease inhibitor with antiviral effects in humans infected with hepatitis c virus," *Nature*, vol. 426, no. 6963, pp. 186–189, 2003.

[31] R. B. Perni, S. J. Almquist, R. A. Byrn, G. Chandorkar, P. R. Chaturvedi, L. F. Courtney, C. J. Decker, K. Dinehart, C. A. Gates, S. L. Harbeson, A. Heiser, G. Kalkeri, E. Kolaczkowski, K. Lin, Y.-P. Luong, B. G. Rao, W. P. Taylor, J. A. Thomson, R. D. Tung, Y. Wei, A. D. Kwong, and C. Lin, "Preclinical profile of vx-950, a potent, selective, and orally bioavailable inhibitor of hepatitis c virus ns3-4a serine protease," *Antimicrob. Agents Chemother.*, vol. 50, no. 3, pp. 899–909, 2006.

[32] A. D. Kwong, R. S. Kauffman, P. Hurter, and P. Mueller, "Discovery and development of telaprevir: an ns3-4a protease inhibitor for treating genotype 1 chronic hepatitis c virus," *Nat Biotechnol*, vol. 29, no. 11, pp. 993–1003, 2011.

[33] C. Hézode, N. Forestier, G. Dusheiko, P. Ferenci, S. Pol, T. Goeser, J.-P. Bronowicki, M. BourliÃre, S. Gharakhanian, L. Bengtsson, L. McNair, S. George, T. Kieffer, A. Kwong, R. S. Kauffman, J. Alam, J.-M. Pawlotsky, and S. Zeuzem, "Telaprevir

and peginterferon with or without ribavirin for chronic hcv infection," *N. Engl. J. Med.*, vol. 360, no. 18, pp. 1839–1850, 2009.

[34] J. G. McHutchison, G. T. Everson, S. C. Gordon, I. M. Jacobson, M. Sulkowski, R. Kauffman, L. McNair, J. Alam, and A. J. Muir, "Telaprevir with peginterferon and ribavirin for chronic hcv genotype 1 infection," *N. Engl. J. Med.*, vol. 360, no. 18, pp. 1827–1838, 2009.

[35] B. A. Malcolm, R. Liu, F. Lahser, S. Agrawal, B. Belanger, N. Butkiewicz, R. Chase, F. Gheyas, A. Hart, D. Hesk, P. Ingravallo, C. Jiang, R. Kong, J. Lu, J. Pichardo, A. Prongay, A. Skelton, X. Tong, S. Venkatraman, E. Xia, V. Girijavallabhan, and F. G. Njoroge, "Sch 503034, a mechanism-based inhibitor of hepatitis c virus ns3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells," *Antimicrob. Agents Chemother.*, vol. 50, no. 3, pp. 1013–1020, 2006.

[36] P. Y. Kwo, E. J. Lawitz, J. McCone, E. R. Schiff, J. M. Vierling, D. Pound, M. N. Davis, J. S. Galati, S. C. Gordon, N. Ravendhran, L. Rossaro, F. H. Anderson, I. M. Jacobson, R. Rubin, K. Koury, L. D. Pedicone, C. A. Brass, E. Chaudhri, and J. K. Albrecht, "Efficacy of boceprevir, an ns3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis c infection (sprint-1): an open-label, randomised, multicentre phase 2 trial," *Lancet*, vol. 376, no. 9742, pp. 705–716, 2010.

[37] L. Chatel-Chaix, M. Baril, and D. Lamarre, "Hepatitis c virus ns3/4a protease inhibitors: A light at the end of the tunnel.," *Viruses*, vol. 2, pp. 1752–1765, Aug 2010.

[38] K. Chayama and C. N. Hayes, "Hcv drug resistance challenges in japan: The role of pre-existing variants and emerging resistant strains in direct acting antiviral therapy.," *Viruses*, vol. 7, no. 10, pp. 5328–5342, 2015.

[39] Å. Rosenquist, B. Samuelsson, P.-O. Johansson, M. D. Cummings, O. Lenz, P. Raboisson, K. Simmen, S. Vendeville, H. de Kock, M. Nilsson, A. Horvath, R. Kalmeijer, G. de la Rosa, and M. Beumont-Mauviel, "Discovery and development of simeprevir (tmc435), a hcv ns3/4a protease inhibitor," *J. Med. Chem.*, vol. 57, no. 5, pp. 1673–1693, 2014.

[40] P. Andreone, M. G. Colombo, J. V. Enejosa, I. Koksal, P. Ferenci, A. Maieron, B. Mullhaupt, Y. Horsmans, O. Weiland, H. W. Reesink, J. Rodrigues, L., Y. B. Hu, T. Podsadecki, and B. Bernstein, "Abt-450, ritonavir, ombitasvir, and dasabuvir achieves 97ribavirin in treatment-experienced patients with hcv genotype 1b infection," *Gastroenterology*, vol. 147, no. 2, pp. 359–365.e1, 2014.

[41] J. A. McCauley, C. J. McIntyre, M. T. Rudd, K. T. Nguyen, J. J. Romano, J. W. Butcher, K. F. Gilbert, K. J. Bush, M. K. Holloway, J. Swestock, B. L. Wan, S. S. Carroll, J. M. DiMuzio, D. J. Graham, S. W. Ludmerer, S. S. Mao, M. W. Stahlhut, C. M. Fandozzi, N. Trainor, D. B. Olsen, J. P. Vacca, and N. J. Liverton, "Discovery of vaniprevir (mk-7009), a macrocyclic hepatitis c virus ns3/4a protease inhibitor," *J. Med. Chem.*, vol. 53, no. 6, pp. 2443–2463, 2010.

[42] P. Ingallinella, S. Altamura, E. Bianchi, M. Taliani, R. Ingenito, R. Cortese, R. De Francesco, C. Steinkuhler, and A. Pessi, "Potent peptide inhibitors of human hepatitis c virus ns3 protease are obtained by optimizing the cleavage products," *Biochemistry*, vol. 37, no. 25, pp. 8906–14, 1998.

[43] C. Steinkuhler, G. Biasiol, M. Brunetti, A. Urbani, U. Koch, R. Cortese, A. Pessi, and R. De Francesco, "Product inhibition of the hepatitis c virus ns3 protease," *Biochemistry*, vol. 37, no. 25, pp. 8899–905, 1998.

[44] M. Llinàs-Brunet, M. Bailey, G. Fazal, E. Ghiro, V. Gorys, S. Goulet, T. Halmos, R. Maurice, M. Poirier, M.-A. Poupart, J. Rancourt, D. Thibeault, D. Wernic, and D. Lamarre, "Highly potent and selective peptide-based inhibitors of the hepatitis c virus serine protease: towards smaller inhibitors," *Bioorganic Medicinal Chemistry Letters*, vol. 10, no. 20, pp. 2267–2270, 2000.

[45] M. Hagel, D. Niu, T. St Martin, M. P. Sheets, L. Qiao, H. Bernard, R. M. Karp, Z. Zhu, M. T. Labenski, P. Chaturvedi, M. Nacht, W. F. Westlin, R. C. Petter, and J. Singh, "Selective irreversible inhibition of a protease by targeting a noncatalytic cysteine.," *Nat Chem Biol*, vol. 7, pp. 22–24, Jan 2011.

[46] Y. S. Tsantrizos, "The design of a potent inhibitor of the hepatitis c virus ns3 protease: Biln 2061–from the nmr tube to the clinic," *Biopolymers*, vol. 76, no. 4, pp. 309–23, 2004.

[47] M. Bäck, P.-O. Johansson, F. Wångsell, F. Thorstensson, I. Kvarnström, S. Ayesa, H. Wähling, M. Pelcman, K. Jansson, S. Lindström, H. Wallberg, B. Classon, C. Rydergård, L. Vrang, E. Hamelink, A. Hallberg, Å. Rosenquist, and B. Samuelsson, "Novel potent macrocyclic inhibitors of the hepatitis c virus ns3 protease: Use of cyclopentane and cyclopentene p2-motifs," *Bioorganic Medicinal Chemistry*, vol. 15, no. 22, pp. 7184–7202, 2007.

[48] M. LlinÃ s Brunet, M. D. Bailey, G. Bolger, C. Brochu, A.-M. Faucher, J. M. Ferland, M. Garneau, E. Ghiro, V. Gorys, C. Grand-MaÃ®tre, T. Halmos, N. Lapeyre-Paquette, F. Liard, M. Poirier, M. RhÃ©aume, Y. S. Tsantrizos, and D. Lamarre,

"Structure–activity study on a novel series of macrocyclic inhibitors of the hepatitis c virus ns3 protease leading to the discovery of biln 2061," *J. Med. Chem.*, vol. 47, no. 7, pp. 1605–1608, 2004.

[49] N. J. Liverton, M. K. Holloway, J. A. McCauley, M. T. Rudd, J. W. Butcher, S. S. Carroll, J. DiMuzio, C. Fandozzi, K. F. Gilbert, S. S. Mao, C. J. McIntyre, K. T. Nguyen, J. J. Romano, M. Stahlhut, B. L. Wan, D. B. Olsen, and J. P. Vacca, "Molecular modeling based approach to potent p2-p4 macrocyclic inhibitors of hepatitis c ns3/4a protease," *J Am Chem Soc*, vol. 130, no. 14, pp. 4607–9, 2008.

[50] N. J. Liverton, S. S. Carroll, J. DiMuzio, C. Fandozzi, D. J. Graham, D. Hazuda, M. K. Holloway, S. W. Ludmerer, J. A. McCauley, C. J. McIntyre, D. B. Olsen, M. T. Rudd, M. Stahlhut, and J. P. Vacca, "Mk-7009, a potent and selective inhibitor of hepatitis c virus ns3/4a protease," *Antimicrob. Agents Chemother.*, vol. 54, no. 1, pp. 305–311, 2010.

[51] V. Summa, S. W. Ludmerer, J. A. McCauley, C. Fandozzi, C. Burlein, G. Claudio, P. J. Coleman, J. M. DiMuzio, M. Ferrara, M. Di Filippo, A. T. Gates, D. J. Graham, S. Harper, D. J. Hazuda, C. McHale, E. Monteagudo, V. Pucci, M. Rowley, M. T. Rudd, A. Soriano, M. W. Stahlhut, J. P. Vacca, D. B. Olsen, N. J. Liverton, and S. S. Carroll, "Mk-5172, a selective inhibitor of hepatitis c virus ns3/4a protease with broad activity across genotypes and resistant variants," *Antimicrob. Agents Chemother.*, vol. 56, no. 8, pp. 4161–4167, 2012.

[52] J. M. Cuevas, F. Gonzalez-Candelas, A. Moya, and R. Sanjuan, "Effect of ribavirin on the mutation rate and spectrum of hepatitis c virus in vivo.," *J Virol*, vol. 83, pp. 5760–5764, Jun 2009.

[53] L. Rong, H. Dahari, R. M. Ribeiro, and A. S. Perelson, "Rapid emergence of protease inhibitor resistance in hepatitis c virus," *Sci. Transl. Med.*, vol. 2, no. 30, p. 30ra32, 2010.

[54] M. Martell, J. I. Esteban, J. Quer, J. Genesca, A. Weiner, R. Esteban, J. Guardia, and J. Gomez, "Hepatitis c virus (hcv) circulates as a population of different but closely related genomes: quasispecies nature of hcv genome distribution," *J Virol*, vol. 66, no. 5, pp. 3225–9, 1992.

[55] S. Duffy, L. A. Shackelton, and E. C. Holmes, "Rates of evolutionary change in viruses: patterns and determinants.," *Nat Rev Genet*, vol. 9, pp. 267–276, Apr 2008.

[56] S. Ojosnegros, C. Perales, A. Mas, and E. Domingo, "Quasispecies as a matter of fact: viruses and beyond.," *Virus Res*, vol. 162, pp. 203–215, Dec 2011.

[57] J. Vermehren and C. Sarrazin, "The role of resistance in hcv treatment," *Best Pract. Res. Clin. Gastroenterol.*, vol. 26, no. 4, pp. 487–503, 2012.

[58] P. Halfon and S. Locarnini, "Hepatitis c virus resistance to protease inhibitors," *J Hepatol*, vol. 55, no. 1, pp. 192–206, 2011.

[59] C. Sarrazin and S. Zeuzem, "Resistance to direct antiviral agents in patients with hepatitis c virus infection," *Gastroenterology*, vol. 138, no. 2, pp. 447–462, 2010.

[60] T. L. Kieffer, C. Sarrazin, J. S. Miller, M. W. Welker, N. Forestier, H. W. Reesink, A. D. Kwong, and S. Zeuzem, "Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis c virus replication in patients," *Hepatology*, vol. 46, no. 3, pp. 631–639, 2007.

[61] C. Sarrazin, T. L. Kieffer, D. Bartels, B. Hanzelka, U. Muh, M. Welker, D. Wincheringer, Y. Zhou, H. M. Chu, C. Lin, C. Weegink, H. Reesink, S. Zeuzem,

and A. D. Kwong, "Dynamic hepatitis c virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir," *Gastroenterology*, vol. 132, no. 5, pp. 1767–1777, 2007.

[62] S. Susser, C. Welsch, Y. Wang, M. Zettler, F. S. Domingues, U. Karey, E. Hughes, R. Ralston, X. Tong, E. Herrmann, S. Zeuzem, and C. Sarrazin, "Characterization of resistance to the protease inhibitor boceprevir in hepatitis c virus-infected patients," *Hepatology*, vol. 50, no. 6, pp. 1709–1718, 2009.

[63] M. P. Manns, M. Bourlière, Y. Benhamou, S. Pol, M. Bonacini, C. Trepo, D. Wright, T. Berg, J. Calleja, P. W. White, J. O. Stern, G. Steinmann, C.-L. Yong, G. Kukolj, J. Scherer, and W. O. Boecher, "Potency, safety, and pharmacokinetics of the ns3/4a protease inhibitor bi201335 in patients with chronic hcv genotype-1 infection," *J. Hepatol.*, vol. 54, no. 6, pp. 1114–1122, 2011.

[64] M. P. Manns, E. Gane, M. Rodriguez-Torres, A. Stoehr, C. T. Yeh, and e. al, "Mk-7009 significantly improves rapid viral response (rvr) in combination with pegylated interferon alfa-2a and ribavirin in patients with chronic hepatitis c (chc) genotype 1 infection," *J. Hepatol.*, vol. 50, p. S384, 2009.

[65] M. Manns, H. Reesink, T. Berg, G. Dusheiko, R. Flisiak, P. Marcellin, C. Moreno, O. Lenz, P. Meyvisch, M. Peeters, V. Sekar, K. Simmen, and R. Verloes, "Rapid viral response of once-daily tmc435 plus pegylated interferon/ribavirin in hepatitis c genotype-1 patients: a randomized trial," *Antivir. Ther.*, vol. 16, pp. 1021–1033, 2011.

[66] Y. Jiang, S. W. Andrews, K. R. Condroski, B. Buckman, V. Serebryany, S. Wenglowsky, A. L. Kennedy, M. R. Madduru, B. Wang, M. Lyon, G. A. Doherty, B. T.

Woodard, C. Lemieux, M. G. Do, H. Zhang, J. Ballard, G. Vigers, B. J. Brandhuber, P. Stengel, J. A. Josey, L. Beigelman, L. Blatt, and S. D. Seiwert, "Discovery of danoprevir (itmn-191/r7227), a highly selective and potent inhibitor of hepatitis c virus (hcv) ns3/4a protease," *J. Med. Chem.*, vol. 57, no. 5, pp. 1753–1769, 2013.

[67] M. Prabu-Jeyabalan, E. Nalivaika, and C. A. Schiffer, "Substrate shape determines specificity of recognition for hiv-1 protease: analysis of crystal structures of six substrate complexes," *Structure*, vol. 10, no. 3, pp. 369–381, 2002.

[68] K. P. Romano, A. Ali, C. Aydin, D. Soumana, A. Ozen, L. M. Deveau, C. Silver, H. Cao, A. Newton, C. J. Petropoulos, W. Huang, and C. A. Schiffer, "The molecular basis of drug resistance against hepatitis c virus ns3/4a protease inhibitors," *PLoS Pathog*, vol. 8, no. 7, p. e1002832, 2012.

[69] M. J. Alter, "Epidemiology of hepatitis c virus infection," *World J Gastroenterol*, vol. 13, no. 17, pp. 2436–41, 2007.

[70] A. A. Modi and T. J. Liang, "Hepatitis c: a clinical review," *Oral Dis*, vol. 14, no. 1, pp. 10–4, 2008.

[71] A. D. Kwong, "The hcv revolution did not happen overnight.," *ACS Med Chem Lett*, vol. 5, pp. 214–220, Mar 2014.

[72] X. D. Li, L. Sun, R. B. Seth, G. Pineda, and Z. J. Chen, "Hepatitis c virus protease ns3/4a cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity," *Proc Natl Acad Sci U S A*, vol. 102, no. 49, pp. 17717–22, 2005.

[73] Z. Chen, Y. Benureau, R. Rijnbrand, J. Yi, T. Wang, L. Warter, R. E. Lanford, S. A. Weinman, S. M. Lemon, A. Martin, and K. Li, "Gb virus b disrupts rig-i signaling

by ns3/4a-mediated cleavage of the adaptor protein mavs," *J. Virol.*, vol. 81, no. 2, pp. 964–976, 2007.

[74] D. P. Rotella, "The discovery and development of boceprevir," *Expert Opinion on Drug Discovery*, vol. 8, no. 11, pp. 1439–1447, 2013. PMID: 24079543.

[75] F. McPhee, "Identification and preclinical profile of the novel hcv ns3 protease inhibitor bms-650032," *J Hepatol*, vol. 52, p. S296, 2010.

[76] P. W. White, M. Llinàs-Brunet, M. Amad, R. C. Bethell, G. Bolger, M. G. Cordingley, J. Duan, M. Garneau, L. Lagacé, D. Thibeault, and G. Kukolj, "Preclinical characterization of bi 201335, a c-terminal carboxylic acid inhibitor of the hepatitis c virus ns3-ns4a protease," *Antimicrob. Agents Chemother.*, vol. 54, no. 11, pp. 4611–4618, 2010.

[77] S. D. Seiwert, S. W. Andrews, Y. Jiang, V. Serebryany, H. Tan, K. Kossen, P. T. Rajagopalan, S. Misialek, S. K. Stevens, A. Stoycheva, J. Hong, S. R. Lim, X. Qin, R. Rieger, K. R. Condroski, H. Zhang, M. G. Do, C. Lemieux, G. P. Hingorani, D. P. Hartley, J. A. Josey, L. Pan, L. Beigelman, and L. M. Blatt, "Preclinical characteristics of the hepatitis c virus ns3/4a protease inhibitor itmn-191 (r7227)," *Antimicrob. Agents Chemother.*, vol. 52, no. 12, pp. 4432–4441, 2008.

[78] T.-I. Lin, O. Lenz, G. Fanning, T. Verbinnen, F. Delouvroy, A. Scholliers, K. Vermeiren, A. Rosenquist, M. Edlund, B. Samuelsson, L. Vrang, H. de Kock, P. Wigerinck, P. Raboisson, and K. Simmen, "In vitro activity and preclinical profile of tmc435350, a potent hepatitis c virus protease inhibitor," *Antimicrob. Agents Chemother.*, vol. 53, no. 4, pp. 1377–1385, 2009.

[79] S. Harper, J. A. McCauley, M. T. Rudd, M. Ferrara, M. DiFilippo, B. Crescenzi, U. Koch, A. Petrocchi, M. K. Holloway, J. W. Butcher, J. J. Romano, K. J. Bush,

K. F. Gilbert, C. J. McIntyre, K. T. Nguyen, E. Nizi, S. S. Carroll, S. W. Ludmerer, C. Burlein, J. M. DiMuzio, D. J. Graham, C. M. McHale, M. W. Stahlhut, D. B. Olsen, E. Monteagudo, S. Cianetti, C. Giuliano, V. Pucci, N. Trainor, C. M. Fandozzi, M. Rowley, P. J. Coleman, J. P. Vacca, V. Summa, and N. J. Liverton, "Discovery of mk-5172, a macrocyclic hepatitis c virus ns3/4a protease inhibitor," *ACS Med Chem Lett*, vol. 3, no. 4, pp. 332–6, 2012.

[80] K. P. Romano, A. Ali, W. E. Royer, and C. A. Schiffer, "Drug resistance against hcv ns3/4a inhibitors is defined by the balance of substrate recognition versus inhibitor binding," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, pp. 20986–20991, 2010.

[81] A. Ali, C. Aydin, R. Gildemeister, K. P. Romano, H. Cao, A. Özen, D. Soumana, A. Newton, C. J. Petropoulos, W. Huang, and C. A. Schiffer, "Evaluating the role of macrocycles in the susceptibility of hepatitis c virus ns3/4a protease inhibitors to drug resistance," *ACS Chem. Biol.*, vol. 8, no. 7, pp. 1469–1478, 2013.

[82] F. McPhee, A. K. Sheaffer, J. Friborg, D. Hernandez, P. Falk, G. Zhai, S. Levine, S. Chaniewski, F. Yu, D. Barry, C. Chen, M. S. Lee, K. Mosure, L.-Q. Sun, M. Sinz, N. A. Meanwell, R. J. Colonno, J. Knipe, and P. Scola, "Preclinical profile and characterization of the hepatitis c virus ns3 protease inhibitor asunaprevir (bms-650032)," *Antimicrob. Agents Chemother.*, vol. 56, no. 10, pp. 5387–5396, 2012.

[83] F. McPhee, J. Friborg, S. Levine, C. Chen, P. Falk, F. Yu, D. Hernandez, M. S. Lee, S. Chaniewski, A. K. Sheaffer, and C. Pasquinelli, "Resistance analysis of the hepatitis c virus ns3 protease inhibitor asunaprevir," *Antimicrob. Agents Chemother.*, vol. 56, no. 7, pp. 3670–3681, 2012.

[84] G. T. Everson, K. D. Sims, M. Rodriguez-Torres, C. Hézode, E. Lawitz, M. Bourlière, V. Loustaud-Ratti, V. Rustgi, H. Schwartz, H. Tatum, P. Marcellin,

S. Pol, P. J. Thuluvath, T. Eley, X. Wang, S.-P. Huang, F. McPhee, M. Wind-Rotolo, E. Chung, C. Pasquinelli, D. M. Grasela, and D. F. Gardiner, "Efficacy of an interferon- and ribavirin-free regimen of daclatasvir, asunaprevir, and bms-791325 in treatment-naive patients with hcv genotype 1 infection," *Gastroenterology*, vol. 146, pp. 420–429, 2015/11/28.

[85] K. P. Romano, J. M. Laine, L. M. Deveau, H. Cao, F. Massi, and C. A. Schiffer, "Molecular mechanisms of viral and host cell substrate recognition by hepatitis c virus ns3/4a protease," *J. Virol.*, vol. 85, no. 13, pp. 6106–6116, 2011.

[86] M. P. Manns, E. Gane, M. Rodriguez-Torres, A. Stoehr, C.-T. Yeh, P. Marcellin, R. T. Wiedmann, P. M. Hwang, L. Caro, R. J. O. Barnard, and A. W. Lee, "Vaniprevir with pegylated interferon alpha-2a and ribavirin in treatment-naïve patients with chronic hepatitis c: A randomized phase ii study," *Hepatology*, vol. 56, no. 3, pp. 884–893, 2012.

[87] E. S. Svarovskaia, R. Martin, J. G. McHutchison, M. D. Miller, and H. Mo, "Abundant drug-resistant ns3 mutants detected by deep sequencing in hepatitis c virus-infected patients undergoing ns3 protease inhibitor monotherapy," *Journal of clinical microbiology*, vol. 50, no. 10, pp. 3267–3274, 2012.

[88] V. Cento, C. Mirabelli, R. Salpini, S. Dimonte, A. Artese, G. Costa, F. Mercurio, V. Svicher, L. Parrotta, A. Bertoli, *et al.*, "Hcv genotypes are differently prone to the development of resistance to linear and macrocyclic protease inhibitors," *PLoS One*, vol. 7, no. 7, p. e39652, 2012.

[89] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, "Phaser crystallographic software," *J Appl Crystallogr*, vol. 40, no. Pt 4, pp. 658–674, 2007.

[90] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, "Phenix: a comprehensive python-based system for macromolecular structure solution," *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 2, pp. 213–21, 2010.

[91] 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. 2126–2132, 2004.

[92] P. M. Scola, L.-Q. Sun, A. X. Wang, J. Chen, N. Sin, B. L. Venables, S.-Y. Sit, Y. Chen, A. Cocuzza, D. M. Bilder, S. V. D'Andrea, B. Zheng, P. Hewawasam, Y. Tu, J. Friborg, P. Falk, D. Hernandez, S. Levine, C. Chen, F. Yu, A. K. Sheaffer, G. Zhai, D. Barry, J. O. Knipe, Y.-H. Han, R. Schartman, M. Donoso, K. Mosure, M. W. Sinz, T. Zvyaga, A. C. Good, R. Rajamani, K. Kish, J. Tredup, H. E. Klei, Q. Gao, L. Mueller, R. J. Colonno, D. M. Grasela, S. P. Adams, J. Loy, P. C. Levesque, H. Sun, H. Shi, L. Sun, W. Warner, D. Li, J. Zhu, N. A. Meanwell, and F. McPhee, "The discovery of asunaprevir (bms-650032), an orally efficacious ns3 protease inhibitor for the treatment of hepatitis c virus infection," *J. Med. Chem.*, vol. 57, no. 5, pp. 1730–1752, 2014.

[93] M. Prabu-Jeyabalan, E. A. Nalivaika, K. Romano, and C. A. Schiffer, "Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate," *J. Virol.*, vol. 80, no. 7, pp. 3607–3616, 2006.

[94] A. Ozen, W. Sherman, and C. A. Schiffer, "Improving the resistance profile of hepatitis c ns3/4a inhibitors: Dynamic substrate envelope guided design," *J Chem The-*

ory Comput, vol. 9, no. 12, pp. 5693–5705, 2013.

[95] M. W. Fried, "Side effects of therapy of hepatitis c and their management," *Hepatology*, vol. 36, no. 5 Suppl 1, pp. S237–S244, 2002.

[96] M. W. Fried, M. L. Shiffman, K. R. Reddy, C. Smith, G. Marinos, J. Goncales, F. L., D. Haussinger, M. Diago, G. Carosi, D. Dhumeaux, A. Craxi, A. Lin, J. Hoffman, and J. Yu, "Peginterferon alfa-2a plus ribavirin for chronic hepatitis c virus infection," *N Engl J Med*, vol. 347, no. 13, pp. 975–82, 2002.

[97] R. Bartenschlager, V. Lohmann, and F. Penin, "The molecular and structural basis of advanced antiviral therapy for hepatitis c virus infection," *Nat. Rev. Micro.*, vol. 11, no. 7, pp. 482–496, 2013.

[98] G. Szabo and S. Bala, "Micrornas in liver disease," *Nat. Rev. Gastroenterol. Hepatol.*, vol. 10, no. 9, pp. 542–52, 2013.

[99] T. C. Appleby, R. Anderson, O. Fedorova, A. M. Pyle, R. Wang, X. Liu, K. M. Brendza, and J. R. Somoza, "Visualizing atp-dependent rna translocation by the ns3 helicase from hcv," *J Mol Biol*, vol. 405, no. 5, pp. 1139–53, 2011.

[100] R. A. Love, H. E. Parge, J. A. Wickersham, Z. Hostomsky, N. Habuka, E. W. Moomaw, T. Adachi, and Z. Hostomska, "The crystal structure of hepatitis c virus ns3 proteinase reveals a trypsin-like fold and a structural zinc binding site," *Cell*, vol. 87, no. 2, pp. 331–42, 1996.

[101] Y. Yan, Y. Li, S. Munshi, V. Sardana, J. L. Cole, M. Sardana, C. Steinkuehler, L. Tomei, R. De Francesco, L. C. Kuo, and Z. Chen, "Complex of ns3 protease and ns4a peptide of bk strain hepatitis c virus: a 2.2 a resolution structure in a hexagonal crystal form," *Protein Sci*, vol. 7, no. 4, pp. 837–47, 1998.

[102] N. Yao, P. Reichert, S. S. Taremi, W. W. Prosise, and P. C. Weber, "Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis c virus bifunctional protease-helicase," *Structure*, vol. 7, no. 11, pp. 1353–1363, 1999.

[103] L. G. Xu, Y. Y. Wang, K. J. Han, L. Y. Li, Z. Zhai, and H. B. Shu, "Visa is an adapter protein required for virus-triggered ifn-beta signaling," *Molecular Cell*, vol. 19, no. 6, pp. 727–40, 2005.

[104] E. M. Driggers, S. P. Hale, J. Lee, and N. K. Terrett, "The exploration of macrocycles for drug discovery–an underexploited structural class," *Nat. Rev. Drug Discov.*, vol. 7, no. 7, pp. 608–24, 2008.

[105] E. Marsault and M. L. Peterson, "Macrocycles are great cycles: Applications, opportunities, and challenges of synthetic macrocycles in drug discovery," *J. Med. Chem.*, vol. 54, no. 7, pp. 1961–2004, 2011.

[106] D. I. Soumana, A. Ali, and C. A. Schiffer, "Structural analysis of asunaprevir resistance in hcv ns3/4a protease," *ACS Chem. Biol.*, vol. 9, no. 11, pp. 2485–2490, 2014.

[107] A. Özen, W. Sherman, and C. A. Schiffer, "Improving the resistance profile of hepatitis c ns3/4a inhibitors: dynamic substrate envelope guided design," *J. Chem. Theory Comput.*, vol. 9, no. 12, pp. 5693–5705, 2013.

[108] D. Pan, W. Xue, W. Zhang, H. Liu, and X. Yao, "Understanding the drug resistance mechanism of hepatitis c virus ns3/4a to itmn-191 due to r155k, a156v, d168a/e mutations: a computational study," *Biochim Biophys Acta*, vol. 1820, no. 10, pp. 1526–34, 2012.

[109] B. C. Doak, B. Over, F. Giordanetto, and J. Kihlberg, "Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates," *Chem Biol*, vol. 21, no. 9, pp. 1115–42, 2014.

[110] F. Giordanetto and J. Kihlberg, "Macrocyclic drugs and clinical candidates: what can medicinal chemists learn from their properties?," *J Med Chem*, vol. 57, no. 2, pp. 278–95, 2014.

[111] V. C. Clark, J. A. Peter, and D. R. Nelson, "New therapeutic strategies in hcv: second-generation protease inhibitors," *Liver International*, vol. 33, pp. 80–84, 2013.

[112] E. Lawitz, M. S. Sulkowski, R. Ghalib, M. Rodriguez-Torres, Z. M. Younossi, A. Corregidor, E. DeJesus, B. Pearlman, M. Rabinovitz, N. Gitlin, J. K. Lim, P. J. Pockros, J. D. Scott, B. Fevery, T. Lambrecht, S. Ouwerkerk-Mahadevan, K. Callewaert, W. T. Symonds, G. Picchio, K. L. Lindsay, M. Beumont, and I. M. Jacobson, "Simeprevir plus sofosbuvir, with or without ribavirin, to treat chronic infection with hepatitis c virus genotype 1 in non-responders to pegylated interferon and ribavirin and treatment-naive patients: the cosmos randomised study," *Lancet*, vol. 384, no. 9956, pp. 1756–65, 2014.

[113] Y. He, M. S. King, D. J. Kempf, L. Lu, H. B. Lim, P. Krishnan, W. Kati, T. Middleton, and A. Molla, "Relative replication capacity and selective advantage profiles of protease inhibitor-resistant hepatitis c virus (hcv) ns3 protease mutants in the hcv genotype 1b replicon system," *Antimicrob Agents Chemother*, vol. 52, no. 3, pp. 1101–1110, 2008.

[114] M. Nijhuis, R. Schuurman, D. de Jong, J. Erickson, E. Gustchina, J. Albert, P. Schipper, S. Gulnik, and C. A. Boucher, "Increased fitness of drug resistant hiv-1 protease

as a result of acquisition of compensatory mutations during suboptimal therapy," *Aids*, vol. 13, no. 17, pp. 2349–59, 1999.

[115] S. S. Taremi, B. Beyer, M. Maher, N. Yao, W. Prosise, P. C. Weber, and B. A. Malcolm, "Construction, expression, and characterization of a novel fully activated recombinant single-chain hepatitis c virus protease," *Protein Sci*, vol. 7, no. 10, pp. 2143–2149, 1998.

[116] Z. Otwinowski and W. Minor, "Processing of x-ray diffraction data collected in oscillation mode," 1997.

[117] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, "Phenix: a comprehensive python-based system for macromolecular structure solution," *Acta Crystallogr. D Biol. Crystallogr.*, vol. 66, no. 2, pp. 213–221, 2010.

[118] V. B. Chen, r. Arendall, W. B., J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson, "Molprobity: all-atom structure validation for macromolecular crystallography," *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 1, pp. 12–21, 2010.

[119] A. T. Brunger, "Free r value: a novel statistical quantity for assessing the accuracy of crystal structures," *Nature*, vol. 355, no. 6359, pp. 472–475, 1992.

[120] L. Schrodinger, "The axpymol molecular graphics plugin for microsoft powerpoint, version 1.0." 2010.

[121] M. N. Nalam, A. Ali, M. D. Altman, G. S. Reddy, S. Chellappan, V. Kairys, A. Ozen, H. Cao, M. K. Gilson, B. Tidor, T. M. Rana, and C. A. Schiffer, "Evaluating the substrate-envelope hypothesis: structural analysis of novel hiv-1 protease inhibitors designed to be robust against drug resistance," *J. Virol.*, vol. 84, no. 10, pp. 5368–5378, 2010.

[122] D. E. S. R. N. Y. N. . M.-D. I. T. v. . S. N. Y. N. . Schrödinger Release 2014-3: Desmond Molecular Dynamics System, version 3.9.

[123] H. X. R. O. D. M. P. E. B. A. G. J. L. K.-I. K. M. A. M. F. D. S. J. K. S. Y. S. Kevin J. Bowers, Edmond Chow and D. E. Shaw, "Scalable algorithms for molecular dynamics simulations on commodity clusters," in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, (1188455), p. 746, ACM.

[124] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, 1983.

[125] D. Shivakumar, J. Williams, Y. Wu, W. Damm, J. Shelley, and W. Sherman, "Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field," *Journal of Chemical Theory and Computation*, vol. 6, no. 5, pp. 1509–1519, 2010.

[126] A. Ozen, T. Haliloglu, and C. A. Schiffer, "Dynamics of preferential substrate recognition in hiv-1 protease: redefining the substrate envelope," *J Mol Biol*, vol. 410, no. 4, pp. 726–44, 2011.

[127] S. M. Saalau-Bethell, A. J. Woodhead, G. Chessari, M. G. Carr, J. Coyle, B. Graham, S. D. Hiscock, C. W. Murray, P. Pathuri, S. J. Rich, C. J. Richardson, P. A. Williams,

and H. Jhoti, "Discovery of an allosteric mechanism for the regulation of hcv ns3 protein function," *Nat. Chem. Biol.*, vol. 8, no. 11, pp. 920–925, 2012.

[128] O. Abian, S. Vega, J. Sancho, and A. Velazquez-Campoy, "Allosteric inhibitors of the ns3 protease from the hepatitis c virus," *PLoS One*, vol. 8, no. 7, p. e69773, 2013.

[129] A. Velazquez-Campoy, Y. Kiso, and E. Freire, "The binding energetics of first- and second-generation hiv-1 protease inhibitors: Implications for drug design," *Archives of Biochemistry and Biophysics*, vol. 390, no. 2, pp. 169–175, 2001.

[130] "World health organization (who). hepatitis c, fact sheet no 164 (april 2014): http://www.who.int/mediacentre/factsheets/fs164/en/. accessed november 2014.."

[131] E. Murakami, T. Tolstykh, H. Bao, C. Niu, H. M. M. Steuer, D. Bao, W. Chang, C. Espiritu, S. Bansal, A. M. Lam, M. J. Otto, M. J. Sofia, and P. A. Furman, "Mechanism of activation of psi-7851 and its diastereoisomer psi-7977.," *J Biol Chem*, vol. 285, pp. 34337–34347, Nov 2010.

[132] K. V. Kowdley, E. Lawitz, I. Crespo, T. Hassanein, M. N. Davis, M. DeMicco, D. E. Bernstein, N. Afdhal, J. M. Vierling, S. C. Gordon, J. K. Anderson, R. H. Hyland, H. Dvory-Sobol, D. An, R. G. Hindes, E. Albanis, W. T. Symonds, M. M. Berrey, D. R. Nelson, and I. M. Jacobson, "Sofosbuvir with pegylated interferon alfa-2a and ribavirin for treatment-naive patients with hepatitis c genotype-1 infection (atomic): an open-label, randomised, multicentre phase 2 trial.," *Lancet*, vol. 381, pp. 2100–2107, Jun 2013.

[133] M. Gotte, "Resistance to nucleotide analogue inhibitors of hepatitis c virus ns5b: mechanisms and clinical relevance.," *Curr Opin Virol*, vol. 8, pp. 104–108, Oct 2014.

[134] N. Afdhal, K. R. Reddy, D. R. Nelson, E. Lawitz, S. C. Gordon, E. Schiff, R. Nahass, R. Ghalib, N. Gitlin, R. Herring, J. Lalezari, Z. H. Younes, P. J. Pockros, A. M. Di Bisceglie, S. Arora, G. M. Subramanian, Y. Zhu, H. Dvory-Sobol, J. C. Yang, P. S. Pang, W. T. Symonds, J. G. McHutchison, A. J. Muir, M. Sulkowski, and P. Kwo, "Ledipasvir and sofosbuvir for previously treated hcv genotype 1 infection," *N. Engl. J. Med.*, vol. 370, no. 16, pp. 1483–1493, 2014.

[135] M. S. Sulkowski, D. F. Gardiner, M. Rodriguez-Torres, K. R. Reddy, T. Hassanein, I. Jacobson, E. Lawitz, A. S. Lok, F. Hinestrosa, P. J. Thuluvath, H. Schwartz, D. R. Nelson, G. T. Everson, T. Eley, M. Wind-Rotolo, S.-P. Huang, M. Gao, D. Hernandez, F. McPhee, D. Sherman, R. Hindes, W. Symonds, C. Pasquinelli, and D. M. Grasela, "Daclatasvir plus sofosbuvir for previously treated or untreated chronic hcv infection," *N. Engl. J. Med.*, vol. 370, no. 3, pp. 211–221, 2014.

[136] M. Bourliere, V. Oules, C. Ansaldi, X. Adhoute, and P. Castellani, "Sofosbuvir as backbone of interferon free treatments.," *Dig Liver Dis*, vol. 46 Suppl 5, pp. S212–20, Dec 2014.

[137] M. Llinas-Brunet, M. D. Bailey, G. Bolger, C. Brochu, A.-M. Faucher, J. M. Ferland, M. Garneau, E. Ghiro, V. Gorys, C. Grand-Maitre, T. Halmos, N. Lapeyre-Paquette, F. Liard, M. Poirier, M. Rheaume, Y. S. Tsantrizos, and D. Lamarre, "Structure-activity study on a novel series of macrocyclic inhibitors of the hepatitis c virus ns3 protease leading to the discovery of biln 2061," *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1605–1608, 2004.

[138] A. Arasappan, F. Bennett, S. L. Bogen, S. Venkatraman, M. Blackman, K. X. Chen, S. Hendrata, Y. Huang, R. M. Huelgas, L. Nair, A. I. Padilla, W. Pan, R. Pike, P. Pinto, S. Ruan, M. Sannigrahi, F. Velazquez, B. Vibulbhan, W. Wu, W. Yang,

A. K. Saksena, V. Girijavallabhan, N.-Y. Shih, J. Kong, T. Meng, Y. Jin, J. Wong, P. McNamara, A. Prongay, V. Madison, J. J. Piwinski, K.-C. Cheng, R. Morrison, B. Malcolm, X. Tong, R. Ralston, and F. G. Njoroge, "Discovery of narlaprevir (sch 900518): A potent, second generation hcv ns3 serine protease inhibitor," *ACS Med. Chem. Lett.*, vol. 1, no. 2, pp. 64–69, 2010.

[139] B. Bartolini, E. Giombini, P. Zaccaro, M. Selleri, G. Rozera, I. Abbate, U. V. Comandini, G. Ippolito, M. Solmone, and M. R. Capobianchi, "Extent of hcv ns3 protease variability and resistance-associated mutations assessed by next generation sequencing in hcv monoinfected and hiv/hcv coinfected patients," *Virus Research*, vol. 177, no. 2, pp. 205–208, 2013.

[140] S. L. Bogen, A. Arasappan, F. Bennett, K. Chen, E. Jao, Y.-T. Liu, R. G. Lovey, S. Venkatraman, W. Pan, T. Parekh, R. E. Pike, S. Ruan, R. Liu, B. Baroudy, S. Agrawal, R. Chase, P. Ingravallo, J. Pichardo, A. Prongay, J.-M. Brisson, T. Y. Hsieh, K.-C. Cheng, S. J. Kemp, O. E. Levy, M. Lim-Wilby, S. Y. Tamura, A. K. Saksena, V. Girijavallabhan, and F. G. Njoroge, "Discovery of sch446211 (sch6): A new ketoamide inhibitor of the hcv ns3 serine protease and hcv subgenomic rna replication," *Journal of Medicinal Chemistry*, vol. 49, no. 9, pp. 2750–2757, 2006.

[141] S. Carroll, J. McCauley, S. Ludmerer, S. Harper, V. Summa, M. Rowley, M. Rudd, P. Coleman, N. Liverton, J. Butcher, C. McIntyre, J. Romano, K. Bush, M. Ferrara, B. Crescenzi, A. Petrocchi, M. Difilippo, C. Burlein, J. Dimuzio, D. Graham, C. McHale, M. Stahlhut, A. Gates, C. Fandozzi, N. Trainor, D. Hazuda, J. Vacca, and D. Olsen, "Mk-5172: A novel hcv ns3/4a protease inhibitor with potent activity against known resistance mutants," *J Hepatol*, vol. 52, p. S17, 2010.

[142] S. Harper, V. Summa, N. J. Liverton, and J. A. McCauley, "Macrocyclic quinoxaline compounds as hcv ns3 protease inhibitors," 2010.

[143] D. Niu, M. Hagel, L. Qiao, T. St. Martin, M. Sheets, P. Chaturvedi, M. Labenski, M. Nacht, W. Westlin, R. C. Petter, and J. Singh, "764 avl-181, a potent and selective irreversible hcv protease inhibitor, requires a helper amino acid in the binding microenvironment to form a covalent bond with cys-159," *Journal of Hepatology*, vol. 52, Supplement 1, p. S297, 2010.

[144] A. J. Prongay, Z. Guo, N. Yao, J. Pichardo, T. Fischmann, C. Strickland, J. Myers, J., P. C. Weber, B. M. Beyer, R. Ingram, Z. Hong, W. W. Prosise, L. Ramanathan, S. S. Taremi, T. Yarosh-Tomaine, R. Zhang, M. Senior, R. S. Yang, B. Malcolm, A. Arasappan, F. Bennett, S. L. Bogen, K. Chen, E. Jao, Y. T. Liu, R. G. Lovey, A. K. Saksena, S. Venkatraman, V. Girijavallabhan, F. G. Njoroge, and V. Madison, "Discovery of the hcv ns3/4a protease inhibitor (1r,5s)-n-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3- [2(s)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl] - 6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(s)-carboxamide (sch 503034) ii. key steps in structure-based optimization," *J Med Chem*, vol. 50, no. 10, pp. 2310–2318, 2007.

[145] L. Delang, I. Vliegen, M. Froeyen, and J. Neyts, "Comparative study of the genetic barriers and pathways towards resistance of selective inhibitors of hepatitis c virus replication.," *Antimicrob Agents Chemother*, vol. 55, pp. 4103–4113, Sep 2011.

[146] E. J. Gane, C. A. Stedman, R. H. Hyland, X. Ding, E. Svarovskaia, W. T. Symonds, R. G. Hindes, and M. M. Berrey, "Nucleotide polymerase inhibitor sofosbuvir plus ribavirin for hepatitis c.," *N Engl J Med*, vol. 368, pp. 34–44, Jan 2013.

[147] M. Buti, J. Llaneras, M. Riveiro-Barciela, and R. Esteban, "Therapy for hepatitis c genotype 3: moving forward.," *J Viral Hepat*, vol. 22, pp. 683–690, Sep 2015.

[148] D. R. Nelson, J. N. Cooper, J. P. Lalezari, E. Lawitz, P. J. Pockros, N. Gitlin, B. F. Freilich, Z. H. Younes, W. Harlan, R. Ghalib, G. Oguchi, P. J. Thuluvath, G. Ortiz-Lasanta, M. Rabinovitz, D. Bernstein, M. Bennett, T. Hawkins, N. Ravendhran, A. M. Sheikh, P. Varunok, K. V. Kowdley, D. Hennicken, F. McPhee, K. Rana, and E. A. Hughes, "All-oral 12-week treatment with daclatasvir plus sofosbuvir in patients with hepatitis c virus genotype 3 infection: Ally-3 phase iii study.," *Hepatology*, vol. 61, pp. 1127–1135, Apr 2015.

[149] S. B. Jensen, S. B. N. Serre, D. G. Humes, S. Ramirez, Y.-P. Li, J. Bukh, and J. M. Gottwein, "Substitutions at ns3 residue 155, 156, or 168 of hepatitis c virus genotypes 2 to 6 induce complex patterns of protease inhibitor resistance.," *Antimicrob Agents Chemother*, vol. 59, pp. 7426–7436, Dec 2015.

[150] F. McPhee, D. Hernandez, F. Yu, J. Ueland, A. Monikowski, A. Carifa, P. Falk, C. Wang, R. Fridell, T. Eley, N. Zhou, and D. Gardiner, "Resistance analysis of hepatitis c virus genotype 1 prior treatment null responders receiving daclatasvir and asunaprevir," *Hepatology*, vol. 58, no. 3, pp. 902–911, 2013.

[151] N. M. King, M. Prabu-Jeyabalan, E. A. Nalivaika, and C. A. Schiffer, "Combating susceptibility to drug resistance: lessons from hiv-1 protease," *Chem. Biol.*, vol. 11, no. 10, pp. 1333–1338, 2004.

[152] M. N. Nalam, A. Ali, G. S. Reddy, H. Cao, S. G. Anjum, M. D. Altman, N. K. Yilmaz, B. Tidor, T. M. Rana, and C. A. Schiffer, "Substrate envelope-designed potent hiv-1 protease inhibitors to avoid drug resistance," *Chem Biol*, vol. 20, no. 9, pp. 1116–24, 2013.

[153] J. A. O'Meara, C. T. Lemke, C. Godbout, G. Kukolj, L. Lagacé, B. Moreau, D. Thibeault, P. W. White, and M. Llinàs-Brunet, "Molecular mechanism by which a potent hepatitis c virus ns3-ns4a protease inhibitor overcomes emergence of resistance," *Journal of Biological Chemistry*, vol. 288, no. 8, pp. 5673–5681, 2013.

[154] N. . Collaborative Computational Project, "The ccp4 suite: Programs for protein crystallography.," *Acta Crystallographica*, vol. 50, no. 5, pp. 760–763, 1994.

[155] W. Humphrey, A. Dalke, and K. Schulten, "Vmd: visual molecular dynamics," *J Mol Graph*, vol. 14, no. 1, pp. 33–8, 27–8, 1996.

[156] D. Xu, C. J. Tsai, and R. Nussinov, "Hydrogen bonds and salt bridges across protein-protein interfaces.," *Protein Eng*, vol. 10, pp. 999–1012, Sep 1997.

[157] C. Lin, A. D. Kwong, and R. B. Perni, "Discovery and development of vx-950, a novel, covalent, and reversible inhibitor of hepatitis c virus ns3.4a serine protease," *Infect Disord Drug Targets*, vol. 6, no. 1, pp. 3–16, 2006.

[158] C. Lin, K. Lin, Y. P. Luong, B. G. Rao, Y. Y. Wei, D. L. Brennan, J. R. Fulghum, H. M. Hsiao, S. Ma, J. P. Maxwell, K. M. Cottrell, R. B. Perni, C. A. Gates, and A. D. Kwong, "In vitro resistance studies of hepatitis c virus serine protease inhibitors, vx-950 and biln 2061: structural analysis indicates different resistance mechanisms," *J Biol Chem*, vol. 279, no. 17, pp. 17508–14, 2004.

[159] E. Lawitz, E. Gane, B. Pearlman, E. Tam, W. Ghesquiere, D. Guyader, L. Alric, J.-P. Bronowicki, L. Lester, W. Sievert, R. Ghalib, L. Balart, F. Sund, M. Lagging, F. Dutko, M. Shaughnessy, P. Hwang, A. Y. M. Howe, J. Wahl, M. Robertson, E. Barr, and B. Haber, "Efficacy and safety of 12 weeks versus 18 weeks of treatment with grazoprevir (mk-5172) and elbasvir (mk-8742) with or without

ribavirin for hepatitis c virus genotype 1 infection in previously untreated patients with cirrhosis and patients with previous null response with or without cirrhosis (c-worthy): a randomised, open-label phase 2 trial," *Lancet*, pp. DOI: 10.1016/S0140–6736(14)61795–5, 2014.

[160] S. Chellappan, V. Kairys, M. X. Fernandes, C. Schiffer, and M. K. Gilson, "Evaluation of the substrate envelope hypothesis for inhibitors of hiv-1 protease," *Proteins*, vol. 68, no. 2, pp. 561–567, 2007.

[161] J. M. Chen, S. L. Xu, Z. Wawrzak, G. S. Basarab, and D. B. Jordan, "Structure-based design of potent inhibitors of scytalone dehydratase: displacement of a water molecule from the active site," *Biochemistry*, vol. 37, no. 51, pp. 17735–44, 1998.

[162] C. Liu, S. T. Wrobleski, J. Lin, G. Ahmed, A. Metzger, J. Wityak, K. M. Gillooly, D. J. Shuster, K. W. McIntyre, S. Pitt, D. R. Shen, R. F. Zhang, H. Zhang, A. M. Doweyko, D. Diller, I. Henderson, J. C. Barrish, J. H. Dodd, G. L. Schieven, and K. Leftheris, "5-cyanopyrimidine derivatives as a novel class of potent, selective, and orally active inhibitors of p38 map kinase," *Journal of Medicinal Chemistry*, vol. 48, no. 20, pp. 6261–6270, 2005.

[163] M. T. Rudd, J. W. Butcher, K. T. Nguyen, C. J. McIntyre, J. J. Romano, K. F. Gilbert, K. J. Bush, N. J. Liverton, M. K. Holloway, S. Harper, M. Ferrara, M. DiFilippo, V. Summa, J. Swestock, J. Fritzen, S. S. Carroll, C. Burlein, J. M. DiMuzio, A. Gates, D. J. Graham, Q. Huang, S. McClain, C. McHale, M. W. Stahlhut, S. Black, R. Chase, A. Soriano, C. M. Fandozzi, A. Taylor, N. Trainor, D. B. Olsen, P. J. Coleman, S. W. Ludmerer, and J. A. McCauley, "P2-quinazolinones and bis-macrocycles as new templates for next-generation hepatitis c virus ns3/4a protease

inhibitors: discovery of mk-2748 and mk-6325.," *ChemMedChem*, vol. 10, pp. 727–735, Apr 2015.

[164] D. I. Soumana, N. K. Yilmaz, K. L. Prachanronarong, C. Aydin, A. Ali, and C. A. Schiffer, "Structural and thermodynamic effects of macrocyclization in hcv ns3/4a inhibitor mk-5172," *ACS Chemical Biology*, vol. 0, no. ja, p. null, 0. PMID: 26682473.

[165] E. Freire, "A thermodynamic approach to the affinity optimization of drug candidates.," *Chem Biol Drug Des*, vol. 74, pp. 468–472, Nov 2009.

[166] E. Freire, "Do enthalpy and entropy distinguish first in class from best in class?," *Drug Discov Today*, vol. 13, pp. 869–874, Oct 2008.

[167] U. Shah, C. Jayne, S. Chackalamannil, F. Velázquez, Z. Guo, A. Buevich, J. A. Howe, R. Chase, A. Soriano, S. Agrawal, M. T. Rudd, J. A. McCauley, N. J. Liverton, J. Romano, K. Bush, P. J. Coleman, C. Grisé-Bard, M.-C. Brochu, S. Charron, V. Aulakh, B. Bachand, P. Beaulieu, H. Zaghdane, S. Bhat, Y. Han, J. P. Vacca, I. W. Davies, A. E. Weber, and S. Venkatraman, "Novel quinoline-based p2–p4 macrocyclic derivatives as pan-genotypic hcv ns3/4a protease inhibitors," *ACS Med. Chem. Lett.*, vol. 5, no. 3, pp. 264–269, 2014.

[168] A. Lopes, M. Schmidt Am Busch, and T. Simonson, "Computational design of protein-ligand binding: modifying the specificity of asparaginyl-trna synthetase.," *J Comput Chem*, vol. 31, pp. 1273–1286, Apr 2010.

[169] H. Brandstetter, A. Kuhne, W. Bode, R. Huber, W. von der Saal, K. Wirthensohn, and R. A. Engh, "X-ray structure of active site-inhibited clotting factor xa. implications for drug design and substrate recognition.," *J Biol Chem*, vol. 271, pp. 29988–29992, Nov 1996.

[170] A. Y. M. Howe, S. Black, S. Curry, S. W. Ludmerer, R. Liu, R. J. O. Barnard, W. Newhard, P. M. T. Hwang, D. Nickle, C. Gilbert, L. Caro, M. J. DiNubile, and N. Mobashery, "Virologic resistance analysis from a phase 2 study of mk-5172 combined with pegylated interferon/ribavirin in treatment-naive patients with hepatitis c virus genotype 1 infection," *Clin. Infect. Dis.*, p. DOI: 10.1093/cid/ciu696, 2014.

[171] "World health organization (who). west nile virus, fact sheet no 354 (july 2011): http://www.who.int/mediacentre/factsheets/fs354/en/. accessed november 2014.."

[172] H. Wu, S. Bock, M. Snitko, T. Berger, T. Weidner, S. Holloway, M. Kanitz, W. E. Diederich, H. Steuber, C. Walter, D. Hofmann, B. Weissbrich, R. Spannaus, E. G. Acosta, R. Bartenschlager, B. Engels, T. Schirmeister, and J. Bodem, "Novel dengue virus ns2b/ns3 protease inhibitors.," *Antimicrob Agents Chemother*, vol. 59, pp. 1100–1109, Feb 2015.

[173] B. Falgout, M. Pethel, Y. M. Zhang, and C. J. Lai, "Both nonstructural proteins ns2b and ns3 are required for the proteolytic processing of dengue virus nonstructural proteins," *J. Virol.*, vol. 65, no. 5, pp. 2467–2475, 1991.

[174] L. Zhang, P. M. Mohan, and R. Padmanabhan, "Processing and localization of dengue virus type 2 polyprotein precursor ns3-ns4a-ns4b-ns5," *J. Virol.*, vol. 66, no. 12, pp. 7549–7554, 1992.

[175] S. Chandramouli, J. S. Joseph, S. Daudenarde, J. Gatchalian, C. Cornillez-Ty, and P. Kuhn, "Serotype-specific structural differences in the protease-cofactor complexes of the dengue virus family," *J. Virol.*, vol. 84, no. 6, pp. 3059–3067, 2010.

[176] P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, and U. Hommel, "Structural basis for the activation

of flaviviral ns3 proteases from dengue and west nile virus," *Nat. Struct. Mol. Biol.*, vol. 13, no. 4, pp. 372–373, 2006.

[177] M. A. M. Behnam, D. Graf, R. Bartenschlager, D. P. Zlotos, and C. D. Klein, "Discovery of nanomolar dengue and west nile virus protease inhibitors containing a 4-benzyloxyphenylglycine residue.," *J Med Chem*, vol. 58, pp. 9354–9370, Dec 2015.