Extended interaction networks with HCV protease NS3-4A substrates explain the lack of adaptive capability against protease inhibitors

Georg Dultz, Tetsuro Shimakami, Markus Schneider, Kazuhisa Murai, Daisuke Yamane, Antoine Marion, Tobias M. Zeitler, Claudia Stross, Christian Grimm, Rebecca M. Richter, Katrin Bäumer, MinKyung Yi, Ricardo M. Biondi, Stefan Zeuzem, Robert Tampé, Iris Antes, Christian M. Lange, Christoph Welsch

1Department of Internal Medicine 1, Goethe University Hospital Frankfurt, 60590 Frankfurt a.M., Germany
2Department of Gastroenterology, Kanazawa University Hospital, Kanazawa, Japan
3Technical University Munich, Center for Integrated Protein Science Munich (CIPS) and Department of Life Science, 85354 Freising-Weihenstephan, Germany
4Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya-Ku, Tokyo, Japan
5Department of Microbiology and Immunology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA
6Instituto de Investigación en Biomedicina de Buenos Aires (IBioBA) - CONICET - Partner Institute of the Max Planck Society, Buenos Aires, Argentina
7University Center for Infectious Diseases, University Hospital Frankfurt, 60590 Frankfurt a.M., Germany
8Institute of Biochemistry, Biocenter and Cluster of Excellence-Macromolecular Complexes (CEF-MC), Goethe University Frankfurt, 60438 Frankfurt a.M., Germany

*Present address: Department of Gastroenterology and Hepatology, University Hospital Essen and University of Duisburg-Essen, Essen, Germany.

§G.D. and T.S. contributed equally to this work.

#Corresponding author: Christoph Welsch
E-mail: christoph.welsch@kgu.de

Running Title: Structural constraints and adaptive capability in HCV NS3-4A

Keywords: hepatitis C virus (HCV), serine protease (NS3-4A), structure constraints, evolution, adaptation, mitochondrial antiviral signaling protein (MAVS), protease inhibitor, resistance mutation, replicative fitness
ABSTRACT

Inhibitors against the NS3-4A protease of hepatitis C virus have proven to be useful drugs in the treatment of HCV infection. Although variants have been identified with mutations that confer resistance to these inhibitors, the mutations do not restore replicative fitness and no secondary mutations that rescue fitness have been found. To gain insight into the molecular mechanisms underlying the lack of fitness compensation, we screened known resistance mutations in infectious HCV cell culture with different genomic backgrounds. We observed that the Q41R mutation of NS3-4A efficiently rescues the replicative fitness in cell culture for virus variants containing mutations at NS3-D168. To understand how the Q41R mutation rescues activity, we performed protease activity assays complemented by molecular dynamics simulations, which showed that protease-peptide interactions far outside the targeted peptide cleavage sites mediate substrate recognition by NS3-4A and support protease cleavage kinetics. These interactions shed new light on the mechanisms by which NS3-4A cleaves its substrates, viral polyproteins and a prime cellular antiviral adaptor protein, the mitochondrial antiviral signaling protein MAVS. Peptide binding is mediated by an extended hydrogen bond network in NS3-4A that was effectively optimized for protease-MAVS binding in D168-variants with rescued replicative fitness from NS3-Q41R. In the protease harboring NS3-Q41R, the N-terminal cleavage products of MAVS retained high affinity to the active site, rendering the protease susceptible for potential product inhibition. Our findings reveal delicately balanced protease-peptide interactions in viral replication and immune escape that likely restrict the protease adaptive capability and narrow the virus evolutionary space.

INTRODUCTION

Viruses under treatment with specific antivirals undergo rapid adaptive selection to maximize their fit for escape from drug pressure. Thereby, increased replicative fitness can lead to decreased drug sensitivity of viral variants (1) and determine the probability with which resistant variants will persist (2). Based on a deterministic model (3-6) hepatitis C virus (HCV) variants under selective pressure with direct-acting antiviral agents (DAA) are selected by their degree of resistance and capability to replicate and assemble infectious virus particles (7). Protease inhibitors (PI) against the serine protease NS3-4A of HCV are burdened with a particularly high risk for drug resistance development that inflicts considerable fitness costs (7, 8). Although compensatory second-site mutations that rescue fitness deficits from drug resistance mutations are pivotal drivers of viral quasispecies evolution (3, 9, 10), they are virtually not observed upon PI treatment failure in the NS3-4A protease of HCV. While the impact of mutations on PI resistance has been well characterized, their role in replicative fitness of viral variants is less understood. The NS3-4A protease plays an essential role in the HCV replication cycle (11-13). In particular, NS3-4A is involved in RNA synthesis by proteolytically processing nonstructural proteins from the viral polyprotein. In addition, the capacity to cleave and inactivate mitochondrial antiviral-signaling protein (MAVS) (13) (also called Cardif, VISA, and IPS-1 (15-17)) in the RIG-I-signaling pathway is a cardinal feature of NS3-4A, by which HCV blocks induction of interferon-(IFN)-β. MAVS cleavage is a key mechanism in the selection of drug resistance mutations against macrocyclic PIs and is important for promoting viral persistence. Novel PI-based treatment regimens still appear to fail in some patients, specifically those with advanced liver disease or liver cirrhosis (18). Importantly, most of these difficult to treat patients show dominant resistant variants with D168A or D168V amino acid substitutions in NS3-4A at the time of virologic failure (18). We previously showed that the D168 mutant proteases are characterized by increased capacity to cleave MAVS and suppress IFN-β induction (19). The reconstitution of a tight network of electrostatic interactions between protease and the peptide substrate in those variants allows much stronger binding of MAVS than in other protease variants or wild type HCV (19).

In the present study, we unraveled a molecular mechanism in PI-resistant D168 variants that affect the cutting efficiency of peptide substrates by a second-site mutation in the viral protease, NS3-Q41R. By interference with critically balanced protease-peptide interactions, the second-site mutation rescues...
replicative fitness deficits of D168 variants in infectious cell culture. The finding that NS3-Q41R, that itself is reported to confer resistance to macrocyclic PIs (20), has a fitness-compensatory effect in D168 variants is surprising, given the very low prevalence of variants harboring Q41R in PI-experienced patients (<0.1%) (21). Our findings highlight a previously unknown molecular mechanism for MAVS that potentially restricts virus evolution under PI pressure, providing a new paradigm for variant selection from HCV-infected cells.

RESULTS

Rational approach to characterize replicative fitness of PI-resistant variants. We aimed to identify PI resistance-associated amino acid substitutions in NS3-4A that lead to high replicative fitness and potentially confer a general drug resistance phenotype to the virus (1). Resistance mutations were identified by a review of published data from both cell culture-based and clinical studies of PIs (7).

Replicative fitness of D168 variants is enhanced by the second-site mutation NS3-Q41R. To characterize the impact of PI-resistance mutations on RNA replication, we constructed two comprehensive panels, each with 23 resistance mutations in derivatives of pH77S (22) (Fig. 1, A and B). The two panels only differed by the cell culture adaptive mutation Q41R in their NS3-4A protease backbone (7) (Fig. 1B). We measured GLuc activity from NS3-Q41/GLuc2A RNA and NS3-Q41R/GLuc2A RNA in media collected at 24-h intervals after transfection. NS3-Q41R showed no loss of replication capacity compared to NS3-Q41, although the replication capacity of mutants in both backbones, in general, was more compromised in NS3-Q41R (Fig. 1C). Interestingly, NS3-Q41R served to promote the replication of the NS3-D168 variants A/E/H/V despite acting synergistically to impair the replication of other key PI-resistant variants, such as R155K. Compared to wild type (WT), the GLuc activity of D168A was compensated to approximately 91% by NS3-Q41R (161% increase compared to NS3-Q41; fold change (FC) 2.6) and that in D168E was boosted to approximately 141% by NS3-Q41R (44% increase compared to NS3-Q41; FC 1.4). For the clinically relevant D168A/E mutants with increased replicative fitness by NS3-Q41R, we further measured the ability to produce infectious viruses in NS3-Q41 and NS3-Q41R protease backbones using a sensitive GLuc assay (Fig. 2). Both mutants produced more infectious viruses in the NS3-Q41R backbone than in NS3-Q41, which corresponded with the respective results on their RNA replication capacity (Fig. 2). D168 mutants are signature mutations selected from pressure with macrocyclic PIs, such as simprevir, asunaprevir, paritaprevir and glecaprevir (18, 23). To assess the ability of a macrocyclic PI to inhibit the RNA replication of PI-resistant variants in NS3-Q41 and NS3-Q41R protease backbones, antiviral EC50 values were determined from the concentration of simprevir required to cause a 50% reduction in the secretion of GLuc by RNA-transfected cells. The NS3-Q41R backbone caused an increase in EC50 for most PI-resistance mutations tested with on average higher-level PI resistance than the respective mutants in an NS3-Q41 genomic background. All D168 mutants tested in the NS3-Q41R backbone showed high-level PI resistance against simprevir (FC 32-156 compared to the WT protease NS3-Q41); see Table 1 and Fig. S2 for details on the replicative fitness and drug resistance.

For further in vitro characterization, we selected the clinically most relevant PI-resistance associated mutations D168A and D168E that showed rescued replicative fitness due to NS3-Q41R, but omitted D168H (clinically less important) and D168V (only weak fitness compensation). D168T was selected from the D168 mutants as a negative control without impact on replicative fitness in protease backbones ± Q41R.

NS3-Q41R had no effect on the protease fold. To characterize the fitness-compensatory mechanism of NS3-Q41R in D168 variants, we expressed and purified the NS3-4A wild-type (WT) fusion and different mutant forms (see Experimental Procedures and Fig. S1). We first tested if D168A/E/T ± NS3-Q41R had an effect on the protease stability (Fig. 3 and Table 2). The impact of mutants on protein folding was determined by a thermal shift assay, where protein denaturation at increasing temperatures was revealed by increased fluorescence (see Experimental Procedures).
The NS3-Q41 protease WT was stable to temperature with a half maximal denaturation, $T_{\text{m}}$, at 54.9°C (Fig. 3). NS3-Q41R had no effect on the protein stability. A significant difference in protein stability was observed between D168E and D168A/T mutant proteases irrespective of the NS3-4A protease background ± Q41R ($p < 0.0001$) (Fig. 3). As we showed previously (19), the impact on protein fold in D168 mutants is likely related to weakening or stabilization of a local salt bridge and H-bond network in the outer ligand-binding site of the protease. In contrast to D168A, the D168E mutant can compensate the missing interactions of the D168 residue with R155 and R123 and stabilize the protein fold (19). Our present study confirms the modeling results published earlier (19). The melting temperatures reflect a stable protease fold in D168E versus destabilization in D168A (Fig. 3). Importantly, NS3-Q41R showed no difference to the NS3-Q41 WT background neither as a single mutant nor as a second-site mutation with D168A/E/T (Fig. 3 and Table 2). Thus, we considered it unlikely that the increased replicative fitness in the NS3-Q41R background in D168A/E could be due to a modulation of protein stability in the NS3-4A protease.

**NS3-Q41R can modulate the kinetic properties of the protease for viral polyprotein substrates and MAVS.** Next, we analyzed the protease enzymatic activity using an *in vitro* fluorescence resonance energy transfer (FRET)-based assay that measures the cleavage of a viral polyprotein substrate. We tested the kinetic properties of NS3-Q41 WT and D168A/E/T ± NS3-Q41R mutants using a peptide substrate derived from the viral NS4A/4B polyprotein cleavage site (Fig. 4A and Table 2) comprising an excited donor molecule and a quencher (see Experimental Procedures). In this assay, fluorescence is released when the protease recognizes and cleaves the FRET substrate. We determined the amount of substrate needed to obtain half of the protease maximum rate of reaction ($K_M$) and substrate turnover rates ($k_{\text{cat}}$). The protease with the NS3-Q41R backbone exhibited a lower $K_M$ than the NS3-Q41 WT ($p = 0.08$; FC 0.5). No difference was found in the substrate turnover $k_{\text{cat}}$ between the two genomic backgrounds. The D168A mutant in NS3-Q41 showed a significantly lower $k_{\text{cat}}$ ($p = 0.0008$; FC 1.1) and a slight decrease in $K_M$ ($p = 0.24$; FC 0.3) compared to D168-WT. A further decrease in $K_M$ ($p = 0.13$; FC 0.8) paralleled by a significant increase in $k_{\text{cat}}$ ($p = 0.003$; FC 0.7) for D168A was found in the NS3-Q41R backbone. While differences in $K_M$ were not significant for D168E ± NS3-Q41R, $k_{\text{cat}}$ was significantly higher for D168E in the NS3-Q41R versus NS3-Q41 backbone ($p = 0.01$; FC 0.2) (Fig. 4A). No such effect was observed for D168T ± NS3-Q41R (Fig. 4A), that was tested as a D168 variant not compensated for replicative fitness deficits by NS3-Q41R (Fig. 1C). Although $K_M$ and $k_{\text{cat}}$ varied in rather narrow ranges between 1.7 and 3.1 µM and 0.09 and 0.19, respectively, the kinetic constants reflect the fitness-compensatory effect of NS3-Q41R for D168A/E variants that we observed in infectious cell culture (Fig. 1C).

Notably, NS3-Q41R is not observed as a fitness-compensatory mutation in PI-failure patients with D168 variants (24). As we showed previously, efficient MAVS cleavage is a characteristic feature of D168 variants and a key escape mechanism under selective pressure with macrocyclic PIs (19). Hence, we reasoned that NS3-Q41R could have varied the kinetic properties of D168A variants towards MAVS. To characterize the effect of D168A and NS3-Q41R on the protease MAVS cleavage kinetics, we designed a novel FRET-peptide incorporating the MAVS cleavage site. As done for NS4A/4B, $K_M$ and $k_{\text{cat}}$ of D168A ± NS3-Q41R were determined for the MAVS peptide (Fig. 4B). In contrast to NS4A/4B, we observed vast variations in the kinetic constants for MAVS. Compared to the NS3-Q41 protease WT, NS3-Q41R showed an approximately 5-fold increase in $k_{\text{cat}}$ ($p = 0.008$; FC 4.7), although the $K_M$ was significantly higher than in the WT ($p = 0.02$; FC 4.8). A highly significant and fundamental change with inverse kinetic constants was observed when NS3-Q41R was present as a second-site mutation in the D168A-protease (Fig. 4B). Here, D168A showed a more than 16-fold increase in $k_{\text{cat}}$ compared to D168A without NS3-Q41R ($p = 0.0003$; FC 16.5). In parallel, the Michaelis-Menten constant, $K_M$, was significantly decreased ($p < 0.05$; FC 7.1) (Fig. 4B). Thus, there was large reversal in $K_M$ and $k_{\text{cat}}$ for D168A upon the addition of NS3-Q41R. Next, we calculated the ratio $k_{\text{cat}}/K_M$, the ‘specificity constant’, to compare the relative rates of NS3-4A protease cleavage on
the competing substrates NS4A/4B and MAVS (Table 2). We found the specificity constant of the D168A-protease for MAVS more than 100-fold increased when harboring NS3-Q41R and 3.3-fold higher than for NS4A/4B.

**NS3-Q41R increases the ability of the protease and D168 variants to cleave MAVS.** The above results using peptide substrates in vitro suggested a fitness advantage due to NS3-Q41R in D168 variants via increased MAVS cleavage capacity. To test this hypothesis in cells, we prepared expressed constructs for D168-WT and D168A/E/T mutants in protease genomic backgrounds ± NS3-Q41R. The expressed constructs were transfected into U-2 OS cells that express high levels of endogenous MAVS (19). Western blot was employed after transfection with NS3-4A expression constructs to detect full-length and cleaved MAVS (Fig. 5A). Our data revealed that NS3-Q41 constructs harboring D168E/T had a comparable capacity to cleave endogenous MAVS compared to the NS3-Q41 protease WT. A notable exception is D168A in the NS3-Q41 protease that displayed a significantly higher potency to cleave MAVS than the NS3-Q41 protease WT, which is in line with our previous findings (19). Interestingly, the addition of NS3-Q41R increased the protease capacity to cleave MAVS for all D168 variants tested (Fig. 5A). Most notably, full-length MAVS was not visible already after 12 h of transfection in D168A + NS3-Q41R. We observed complete cleavage of MAVS after 16 h of transfection for D168A/E/T in the NS3-Q41R background. Remarkably, we observed no degradation of the cleavage products for D168A/E/T in NS3-Q41R after 12 h of transfection (Fig. 5A).

Next, we tested the capacity of MAVS in U-2 OS cells to induce IFN-β for D168-WT and D168A/E/T mutants in protease genomic backgrounds ± NS3-Q41R after transfection with an IFN-β-reporter plasmid (Fig. 5B). As expected, over-expression of MAVS resulted in strong activation of the IFN-β reporter. As previously shown (19), we find IFN-β activity significantly inhibited by D168A in the NS3-Q41 protease backbone. Moreover, we find that proteases harboring NS3-Q41R or NS3-Q41R + D168E significantly attenuated the IFN-β activity. The competence of the Q41R-D168A protease to interfere with MAVS-induced IFN-β-reporter activation did not differ significantly from the Q41 background. A potential explanation is that peptides from MAVS cleavage could play an inhibitory role at the protease active site. Such a model could explain that the Q41R-D168A protease with high $k_{cat}$ but low $K_M$ for MAVS was not superior in inhibiting IFN-β activity compared to the D168A mutant protease (Fig. 4B and 5B).

**An extended hydrogen bond network around the protease active site plays a critical role in substrate recognition.** We previously showed that the mutant D168A reconstitutes electrostatic interactions between the protease and MAVS facilitating stronger binding of the MAVS peptide to the D168A mutant than to the D168-WT protease (19). To characterize the molecular mechanisms of fitness compensation by NS3-Q41R as a second-site mutation in the D168A protease, we performed molecular dynamics (MD) simulations of the NS3-Q41 protease WT, as well as the Q41R-D168A mutant protease, bound to (i) the viral polyprotein junction NS4A/4B or (ii) MAVS (Fig. 6, left panel). For each simulation, we analyzed the H-bond network surrounding the protease active center (Fig. 6, right panel) and calculated peptide-protease interaction energies (Table 3).

A comparison of the simulated systems showed that the D168-WT protease with NS3-Q41 bound to NS4A/4B formed a well-balanced and stable extended H-bond network (Fig. 6, A and B). The core of this network consisted of H-bonds within the protease catalytic triad D81-H57-S139. This core network was extended by H-bond contacts to Q41 and R155-D168-R123. The two mutants D168A and NS3-Q41R substantially disturbed this extended network, weakening its coupling to the protease active site (Fig. 6, C and D). Instead, R41 and R123 participated in more H-bonds (+ ~ 1.3) and π-π contacts (+42% of simulation time) to the NS4A/4B peptide substrate than in the WT protease. For the Q41R-D168A mutant protease, the interaction energy between protease and the NS4A/4B non-cleaved substrate increased by approximately 29% compared to the WT protease (Table 3).

A similar analysis of the H-bond network for the WT protease bound to MAVS identified a substantially weaker stabilization of the
catalytic site (Fig. 6, E and F) than in the NS4A/4B-protease complex (Fig. 6, A and B). The destabilization was a direct consequence of polar MAVS residues competing with intraprotease H-bonds within the extended network. This changed considerably due to the mutants D168A and NS3-Q41R (Fig. 6, G and H). Q41R participated in strong H-bonds (+0.98) with the protease active site while establishing π-π interactions (+ 81%) to MAVS at the same time, rigidifying the complex around the active site (Fig. 6H). In addition, D168A gained strong H-bonds between MAVS and R155 (+ 1.31) / R123 (+ 0.21). These changes in the bonding network leveraged the affinity of the N-terminal MAVS and NS4A/4B cleavage products to the Q41R-D168A protease (77% versus 9% gain in protease-peptide interaction energy respectively) (Table 3). Summing up effects from protease mutations and substrate binding, NS3-Q41R and D168A showed substantial impact on the H-bond network at the protease active site. Both mutations split the H-bond network and freed protease residues for additional interactions with the substrates, explaining the observed decreases in $K_M$ values. Binding of the natural substrate NS4A/4B yielded a stable extended H-bond network in the WT protease, which was weakened by NS3-Q41R and D168A. In contrast, MAVS binding to the protease naturally disturbed this network in the WT protease. Here, the PI-resistance mutation D168A and its second-site mutation NS3-Q41R showed unique compensatory interactions in the protease bound to MAVS and stabilization of the protease active site surroundings. This aligns well with the $k_{cat}$ trends from our enzyme activity assays and Michaelis-Menten kinetics (Fig. 4).

**DISCUSSION**

Fitness compensation by a second-site mutation in the HCV protease NS3-4A, as we show here, has not been noted previously (21, 24). We find NS3-Q41R to permit a fitness recovery for viral variants with resistance-associated amino acid substitutions at D168 in infectious cell culture at high-level resistance against a macrocyclic PI (Fig. 1C and Table 1). Our findings emphasize the capability of NS3-Q41R to yield ‘superfit’ PI-resistant D168 variants. However, NS3-Q41R is not observed in patients though the selection of D168 variants in advanced-stage liver disease is responsible for the vast majority of PI-treatment failures (18). We considered the underlying mechanisms important to understand structural constraints and adaptive capability in HCV NS3-4A.

To characterize the molecular mechanism of fitness compensation by NS3-Q41R, we used enzyme kinetics from purified protein and complementary molecular dynamics simulations. As previously shown, a characteristic feature of D168-protease mutants is the molecular adaptation of the protease active site for cleavage and inactivation of MAVS (19), a key cellular antiviral adaptor protein whose expression level directly determines antiviral immunity (14-17). Strikingly, we find the D168A mutant protease boosted to 16-fold higher MAVS cleavage efficiency by NS3-Q41R, while the efficiency for the viral polyprotein substrate was only 2-fold increased (Fig. 4). This is important, given that blocking RIG-I signaling by targeting MAVS is a key strategy in several highly diverse classes of viruses to evade type I-interferon immune responses (25-27). Conversely, MAVS in primates evolved under strong positive selection, driven by antagonism from ancient ‘paleovirus’ infections (28). As a consequence of protease-peptide coevolution, viral proteases can be highly adapted for MAVS cleavage already in the wild-type protein (28). In the HCV protease NS3-4A, we find that NS3-Q41R had a strong impact on the catalytic activity and turnover in D168 mutant proteases, thereby differentially tuning the protease for viral polyprotein substrates and MAVS. Noteworthy, the changes in the protease kinetic constants for the viral polyprotein substrate NS4A/4B reflect the fitness-compensatory effect of NS3-Q41R that we observed in infectious cell culture. Of note, the Huh-7.5 cells that we were using here are defective for signaling via RIG-I and MAVS (29). Hence, the MAVS cleavage capacity of the protease mutants should not have made an impact on the replicative fitness in our cell culture model of HCV infection.

Mechanistically, we find a delicately balanced extended hydrogen bond network around the protease active site that was altered by D168 mutants as well as the second-site mutation NS3-Q41R. Governed by changes on the
opposite ends of this network (Fig. 6), the protease active site environment is effectively optimized for MAVS binding by D168A and NS3-Q41R (Fig. 6H and Table 3). In line with these findings, we show that NS3-Q41R in cell culture shifted the balance in protease-substrate recognition in the D168 variants towards MAVS (Fig. 5). The presence of hydrophilic and charged residues in the peptide substrate between the P4 and P3' sites is likely a key factor determining how D168 and Q41 mutations influence binding and catalysis. Therefore, based on our observations for NS4A/4B and MAVS, the binding of other cellular substrates of the NS3-4A protease might also be affected by mutation of these sites: Primarily hydrophobic substrates like the membrane-associated glutathione peroxidase GPx8, the E3 ubiquitin ligase component UV-damaged DNA-binding protein 1 (DDB1) and the cleavage site b within T-cell protein tyrosine phosphatase (TC-PTP), a phosphatase involved in growth factor signaling, could behave similarly to NS4A/4B. While hydrophilic residues of the TIR domain-containing adaptor inducing interferon-β (TRIF) or the TC-PTP cleavage site a might interact with the extended network along the lines of MAVS (Fig. S3).

Notably, we find that the MAVS cleaved products did not get degraded fast and retained high affinity to D168 protease variants when harboring NS3-Q41R (Fig. 5 and Table 3). Hence, we speculated that the cleaved products might compete with protease substrates at the active site, a mechanism that potentially could downregulate the protease catalytic efficiency. Adding to this notion, we calculated protease-peptide interaction energy differences from the molecular dynamics simulations that in the Q41R-D168A mutant context showed a remarkable gain of 77% for the N-terminal MAVS cleavage product compared to only 9% for the viral polyprotein peptide (Table 3). In support of a direct impact of MAVS on the protease catalytic activity, preincubation of the purified protein with MAVS peptides showed more potent blockade of enzyme activity in a protease harboring NS3-Q41R than in the wild-type protease (Fig. 7).

It is difficult to compare directly the in vitro biochemistry with the cellular degradation of MAVS. There are several strong limitations for this comparison such as ambiguous concentrations of viral protease, polyprotein substrates and MAVS expression levels in cells. However, from the biochemistry data we clearly conclude that mutations in the NS3-4A protease can have vast differences in the cutting efficiency of peptide substrates. Our data suggest that protease variants already optimized for MAVS cleavage can be easily overfitted for this peptide substrate by second-site mutations in NS3-4A. Such a mechanism could interfere with RNA synthesis and immune escape and hence prevent selection of fitness-compensatory mutations such as NS3-Q41R in patients. The mechanism could potentially explain previous observations from a chimpanzee infected with a cell-culture derived virus containing NS3-Q41R that initially was poorly fit but evolved to establish long-term persistent infection (30). Yi and coworkers observed an unusual 4- to 6-week delay before detectable viremia in this animal during which the level of intrahepatic innate immune responses increased substantially. They found that the cell-culture adapted H77S.2 strain had reverted R41 back to Q41 wild-type before viremia (30), whereas other chimpanzees infected by NS3-Q41 wild-type strains were viremic within a week (31, 32).

Collectively, our data support a model in which balancing between MAVS cleavage capability and efficiency of polyprotein processing likely restricts the selection of drug-resistant HCV protease variants in vivo. Our data suggest that the negative selection of fitness-compensatory mutations in NS3-4A to some extent is a consequence of protease-MAVS overfitting, which is important because it can prevent fitness compensation and hence ultimately contribute to extinction of PI-resistant HCV variants upon treatment failure (2-5).

EXPERIMENTAL PROCEDURES

Cell culture studies. Central to our studies has been the use of an infectious molecular clone of genotype 1a H77 virus (22). The clone pH77S.2 was derived from pH77S and differs from pH77S.3 by the cell-culture adaptive mutation NS3-Q41R in the NS3-4A protease domain (Fig. 1B). When transfected into permissive Huh-7.5 cells (22), genome-length RNA transcribed from pH77S.2 and pH77S.3 replicates efficiently and produces a virus that
is infectious in naïve cells (22). To monitor replication, the Gaussia luciferase (GLuc) sequence, fused at its C terminus to the foot-and-mouth disease virus (FMDV) 2A autoprotease, was inserted between p7 and NS2 (7).

**Site-directed mutagenesis and transfection.** Primers for NS3-4A amplification were used with restriction sites for NheI and EcoRI and designed as follows: GT1a-NS3-NheI-For (ATG ATG GCT AGC GCC ACC ATG GCC CCC ATC ACG GCG TAC GCC CAG CAG) and GT1a-NS3-EcoRI-Rev (ATG ATG GAA TTC TCA CTA GCA CTC). Amplified PCR products were subcloned into the expression vector pcDNA3.1(+) from Invitrogen (Darmstadt, Germany). NS3-4A mutant constructs in NS3-Q41 and NS3-Q41R backbones were generated by site-directed mutagenesis. Primers for the mutagenesis of D168 mutants were designed as follows: GT1a-NS3-D168A-For (GCT AAA GCG GTG GCC TTT ATC CCT GTG) and GT1a-NS3-D168A-Rev (CAC AGG GAT AAA GGC CAC CGC TTT AGC), GT1a-NS3-D168E-For (GCT AAA GCG GTG GAG TTT ATC CCT GTG) and GT1a-NS3-D168E-Rev (CAC AGG GAT AAA CTC CAC CGC TTT AGC), GT1a-NS3-D168T-For (GCT AAA GCG GTG ACC TTT ATC CCT GTG) and GT1a-NS3-D168T-Rev (CAC AGG GAT AAA GGT CAC CGC TTT AGC). All generated constructs were controlled by sequencing with Eurofins. Transfection was performed using the X-tremeGene HP DNA transfection reagent (Roche, Basel, Switzerland) according to the manufacturer’s instructions.

**Viral fitness and antiviral resistance.** Genome-length RNA was transcribed from the mutated pH77S.2, pH77S.3, pH77S.2/Gluc2A and pH77S.3/GLuc2A plasmids in vitro, and studies to assess antiviral resistance and viral fitness carried out as described previously (7).

**MAVS cleavage and IFN-β promoter activation.** To assess MAVS cleavage in cells, we used U-2 OS human osteosarcoma cells (obtained from the American Type Culture Collection, Rockville, MD) and Western blotting. U-2 OS cells express high levels of endogenous MAVS and are highly permissive for transfection with NS3-4A expression constructs as previously described (19). IFN-β promoter activation assays were based on luciferase reporter assays as described previously (33).

**Antibodies.** Western blot analyses were performed using the following antibodies: mouse monoclonal antibody to Cardif/MAVS (Adri-1) (#A3002702, AdipoGen, Liestal, Switzerland); anti-β-actin mouse monoclonal antibody (AC-15) (Sigma-Aldrich, Hamburg, Germany); goat antibody to HCV NS3 region (#8C07014, Meridian Life Science, Memphis, TN, USA).

**Protein biochemistry.** For transformation and recombinant NS3-4A protein expression, the pET-15b vector incorporating the single chain proteolytic domain (NS4A21-32-GSGS-NS33-181) (34) of WT and mutant NS3-4A was transformed into E. coli Rosetta pLysS (DE3) competent cells (Novagen) according to the manufacturer’s protocol. Recombinant protein production and purification was carried out as described previously (Fig. S1) (34).

**Thermal shift assay.** Stability of the protein fold of expressed WT and mutant NS3-4A was measured by fluorescence-based thermal shift (35) using Sypro Orange, a temperature-stable fluorophore that exhibit enhanced fluorescence upon interacting with unfolded proteins. The impact of mutations on protein unfolding patterns is characterized by fluorescence emission curves. Melting temperatures (T_m) from purified NS3-4A protease mutants are determined by fitting the sigmoidal melt curve to the Boltzmann equation.

**Protease activity assay.** The enzymatic activity and Michaelis-Menten kinetics of WT and mutant NS3-4A proteases were measured by a fluorescence-based protease assay (36, 37). To measure peptide cleavage kinetics, a viral polyprotein substrate derived from the NS4A/4B cleavage site was purchased from Anaspec (SensoLyte 520 HCV, Anaspec). To determine enzyme kinetics for MAVS cleavage, we designed a novel FRET peptide incorporating the MAVS cleavage site, produced by Anaspec (Ac – EREVPCHRPS – NH2).

**MAVS product inhibition assay.** We used an untagged MAVS peptide incorporating the MAVS cleavage site (Ac – EREVPCHRPS – NH2) to determine the influence of MAVS
cleavage products on the cleavage kinetics of viral polyprotein substrates. Small amounts of purified NS3-4A protease (NS3-Q41 WT or NS3-Q41R, 50 nM) were preincubated with MAVS peptide (50 µM) for 0, 10 or 20 min at room temperature to allow different time spans for cleavage product formation. After preincubation, 2 µM of the NS4A/4B FRET peptide (SensoLyte 520 HCV, Anaspec) was added and reaction velocity measured as described above. Noteworthy, MAVS and NS3-4A are shown to colocalize in infected cells (38). Hence, the “effective concentration” of MAVS can theoretically reach levels higher than 50 micromolar.

**Statistical analyses.** Comparisons of measurements were analyzed by two-sided T-tests with a level of significance set at a minimum of \( p = 0.05 \).

**Computational studies.** Molecular systems were prepared from a crystal structure of the NS3-4A protease WT bound to the N-terminal cleavage product of MAVS (PDB-3RC5) (39). We modeled both full-length substrates, MAVS and NS4A/4B, prior to the peptide bond cleavage. The P’1-4 residues of MAVS, missing in the experimental structure, were rebuilt based on the crystal structure of bovine Chymotrypsin bound to Eglin C (PDB-1ACB) (40), as previously performed (41). Topologies and starting coordinates for molecular dynamics simulations of protease WT and mutants bound to MAVS and NS4A/4B substrates were prepared using the leap program of the AmberTools15 (42), minimized in gas phase while restraining the protein backbone. The structures were then solvated and neutralized in a box of SPC/E water molecules (43) using a minimum distance of 14.0 Å between the boundary and the protein atoms. All calculations were performed with the Amber14/AmberTools15 suite of programs (41) using the ff03.r1 Amber force field (44, 45). After minimization and adjustment of the box size to reach a target density of 1,000 kg/m³, the systems were gradually heated up to 300K in the NVT ensemble (45). Finally, production runs were performed for 250 ns in the NPT ensemble at 300K and 1 bar. Both heat-up and production runs used a time step of 1 fs, and SHAKE bond length constraints were applied to all bonds involving hydrogen atoms. Temperature and pressure were controlled with Langevin dynamics (collision frequency of 4.0 ps⁻¹) and isotropic position scaling, respectively.

Analysis of the trajectories was performed on frames extracted from the last 100 ns of the NPT production simulations. Analysis and visualization of residue-interaction networks was carried out using our in-house interaction network plugin CONAN (Manuscript in preparation) for Cytoscape 3 (46). Hydrogen bonds present in each time frame were extracted from MD trajectories using cpptraj (46) with a maximum donor-acceptor distance of 3.5 Å and a minimum donor-hydrogen-acceptor angle of 135°. Edge widths in the networks shown represent the average number of H-bonds during simulation time. Only edges above a minimum simulation presence of 10% are shown. Cation-π stacking interactions between arginine (Arg) and histidine (His) residues were analyzed by considering sidechain distances and relative orientations. The sidechain plane was defined for the Arg and His residues based on the heavy atoms of the side chain guanidinium group and imidazole ring, respectively. In this scheme, an interaction was counted as present if the distance between the center of mass of the two planes did not exceed 5.0 Å, and if the angle between the plane normal vectors was found in the ranges of 0-30° or 150-180°. Protein-substrate interaction energies were obtained from the “Delta G gas” term as calculated by MMPBSA.py (AmberTools17) (47-49). The error of uncorrelated samples \( \sigma_e \) was estimated by an autocorrelation time analysis (48), using a weighted single exponential fit (weight factor \( 1 / (x+1) \)), where \( x \) is the integer distance between frames. To reduce noise in the fit, only autocorrelation values larger than 0.1 were used.

**DATA AVAILABILITY**

All data that support the findings of this study are available from the corresponding author upon request.

**ACKNOWLEDGMENTS**

We thank Hirai Feng and Stanley M. Lemon (The University of North Carolina at Chapel Hill) for helpful discussions and critically reviewing the manuscript, Yolanda Martinez and Marie Fiedler for expert technical
assistance and Charles M. Rice (The Rockefeller University) for supplying the Huh-7.5 cells.

**FUNDING AND ADDITIONAL INFORMATION**

The German Research Foundation supported this work; Cluster of Excellence – Macromolecular Complexes CEF-MC, CRC 807, SPP 1623 and RTG 1986 to R.T., Center for Integrated Protein Science Munich CIPS\textsuperscript{M} to I.A., SFB 749/C8 to A.M., SFB 1035 to M.S. and WE 4388/6-1 to C.W. M.S. was supported by the IGSSE Graduate School. G.D. was supported by a Goethe University-Frankfurt Tandem Research Scholarship. T.S. and C.W. acknowledge support by the DAAD-JSPS Joint Research Program, C.M.L., C.W. and S.Z. by the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz (LOEWE) of the State of Hessen, Research Center for Translational Medicine and Pharmacology TMP, and Novel Drug Targets against Poverty-Related and Neglected Tropical Infectious Diseases DRUID, and the Else Kröner-Fresenius Foundation (EKFS), Research Training Group Translational Research Innovation - Pharma TRIP.

**CONFLICT OF INTEREST**

The authors declare that they have no conflicts of interest with the contents of this article.
REFERENCES

1. Sheldon, J., Beach, N.M., Moreno, E., Gallego, I., Piñeiro, D., Martínez-Salas, E., Gregori, J., Quer, J., Esteban, J.I., Rice, C.M., Domingo, E., and Perales, C. (2014) Increased replicative fitness can lead to decreased drug sensitivity of hepatitis C virus. *J. Virol.* 88, 12098-12111.

2. Doncheva, N.T., Domingues, F.S., McGivern, D.R., Shimakami, T., Zeuzem, S., Lengauer, T., Lange, C.M., Albrecht, M., and Welsch, C. (2019) Near-Neighbor Interactions in the NS3-4A Protease of HCV Impact Replicative Fitness of Drug-Resistant Viral Variants. *J Mol Biol.* 431, 2354-2368.

3. Domingo, E., Sheldon, J., and Perales, C. (2012) Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* 76, 159-216.

4. 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.

5. Pawlotsky, J.M. (2016) Hepatitis C Virus Resistance to Direct-Acting Antiviral Drugs in Interferon-Free Regimens. *Gastroenterology* 151, 70-86.

6. Guedj, J., and Neumann, A.U. (2010) Understanding hepatitis C viral dynamics with direct-acting antiviral agents due to the interplay between intracellular replication and cellular infection dynamics. *J. Theor. Biol.* 267, 330-340.

7. Shimakami, T., Welsch, C., Yamane, D., McGivern, D.R., Yi, M., Zeuzem, S., and Lemon, S.M. (2011) Protease inhibitor-resistant hepatitis C virus mutants with reduced fitness from impaired production of infectious virus. *Gastroenterology* 140, 667-675.

8. Götte, M., Berghuis, A.A., Matlashewski, G., Sheppard, D., and Wainberg, M.A. (2017) Handbook of Antimicrobial Resistance, 1st Ed., Springer-Verlag, New York.

9. Cervera, H., Lalić, J., and Elena, S.F. (2016) Efficient escape from local optima in a highly rugged fitness landscape by evolving RNA virus populations. *Proc. Biol. Sci.* 283, 20160984.

10. Acevedo, A., Brodsky, L., and Andino, R. (2014) Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Science* 345, 686-690.

11. Bartenschlager, R., Lohmann, V., and Penin, F. (2013) The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nat. Rev. Microbiol.* 11, 482-446.

12. Ma, Y., Yates, J., Liang, Y., Lemon, S.M., and Yi, M. (2008) NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *J. Virol.* 82, 7624-7639.

13. McGivern, D.R., Masaki, T., Lovell, W., Hamlett, C., Saalau-Bethell, S., and Graham, B. (2015) Protease Inhibitors Block Multiple Functions of the NS3/4A Protease-Helicase during the Hepatitis C Virus Life Cycle. *J. Virol.* 89, 5362-5370.

14. Seth, R.B., Sun, L., Ea, C.K., and Chen, Z.J. (2005) Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* 122, 669-682.

15. Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R., and Tschopp, J. (2005) Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* 437, 1167-1172.

16. Xu, L.G., Wang, Y.Y., Han, K.J., Li, L.Y., Zhai, Z., and Shu, H.B. (2005) VISA is an adapter required for virus-triggered IFN-beta signaling. *Mol. Cell.* 19, 727-740.

17. Kawai, T., Takahashi, K., Sato, S., Coban, C., Kumar, H., Kato, H., Ishii, K.J., Takeuchi, O., and Akira, S. (2005) IPS-1, an adapter triggering RIG-I and Mda5-mediated type I interferon induction. *Nat. Immunol.* 6, 981-988.

18. Poordad, F., Hezode, C., Trinh, R., Kowdley, K.V., Zeuzem, S., Agarwal, K., Shiffman, M.L., Wedemeyer, H., Berg, T., Yoshida, E.M., Forns, X., Lovell, S.S., Da Silva-Tillmann, B., Collins, C.A., Campbell, A.L., Podscadeck, T., and Bernstein, B. (2014) ABT-450/r-ombitasvir and dasabuvir with ribavirin for hepatitis C with cirrhosis. *N. Engl. J. Med.* 370, 1973-1982.

19. Welsch, C., Haselow, K., Gouttenoire, J., Schneider, M., Morikawa, K., Martinez, Y., Susser, S., Sarrazin, C., Zeuzem, S., Antes, I., Moradpour, D., and Lange, C.M. (2015) Hepatitis C virus variants resistant to macrolyclic NS3-4A inhibitors subvert IFN-β induction by efficient MAVS cleavage. *J. Hepatol.* 62, 779-784.

20. Jiang, M., Mani, N., Lin, C., Ardzinski, A., Nelson, M., Reagan, D., Bartels, D., Zhou, Y., Nicolas, O., Rao, B.G., Müh, U., Hanzelka, B., Tigges, A., Rijnbrand, R., and Kieffer, T.L.
(2013) In vitro phenotypic characterization of hepatitis C virus NS3 protease variants observed in clinical studies of telaprevir. *Antimicrob. Agents Chemother.* **57**, 6236-6245.

21. Wu, R., Geng, D., Chi, X., Wang, X., Gao, X., Xu, H., Shi, Y., Guan, Y., Wang, Y., Jin, J., Ding, Y., and Niu, J. (2019) Computational analysis of naturally occurring resistance-associated substitutions in genes NS3, NS5A, and NS5B among 86 subtypes of hepatitis C virus worldwide. *Infect. Drug Resist.* **12**, 2987-3015.

22. Yi, M., Villanueva, R.A., Thomas, D.L., Wakita, T., and Lemon, S.M. (2006) Production of infectious genotype 1a hepatitis C virus (Hutchinson strain) in cultured human hepatoma cells. *Proc. Natl. Acad. Sci. U S A* **103**, 2310-2315.

23. Krishnan, P., Schnell, G., Tripathi, R., Ng, T., Reisch, T., Beyer, J., Dekhtyar, T., Irvin, M., Xie, W., Larsen, L., Mensa, F., Pilot-Matias, T., and Collins, C. (2017) Pooled resistance analysis in HCV genotype 1–6-infected patients treated with glecaprevir/pibrentasvir in phase 2 and 3 clinical trials. *J. Hepatol.* **66**, S500.

24. Welzel, T.M., Bhardwaj, N., Hedskog, C., Chodavarapu, K., Camus, G., McNally, J., Brainard, D., Miller, M.D., Mo, H., Svarovskaia, E., Jacobson, I., Zeuzem, S., and Agarwal, K. (2017) Global epidemiology of HCV subtypes and resistance-associated substitutions evaluated by sequencing-based subtype analyses. *J. Hepatol.* **8278**, 30148-30154.

25. Choi, H.J., Park, A., Kang, S., Lee, E., Lee, T.A., Ra, E.A., Lee, J., Lee, S., and Park, B. (2018) Human cytomegalovirus-encoded US9 targets MAVS and STING signaling to evade type I interferon immune responses. *Nat. Commun.* **9**, 125.

26. Wang, B., Xi, X., Lei, X., Zhang, X., Cui, S., Wang, J., Jin, Q., and Zhao, Z. (2013) Enterovirus 71 protease 2Apro targets MAVS to inhibit anti-viral type I interferon responses. *PLoS Pathog.* **9**, e1003231.

27. Feng, H., Sander, A.L., Moreira-Soto, A., Yamane, D., Drexler, J.F., and Lemon, S.M. (2019) Hepatovirus 3ABC proteases and evolution of mitochondrial antiviral signaling protein (MAVS). *J. Hepatol.* **71**, 25-34.

28. Patel, M.R., Loo, Y.M., Horner, S.M., Gale, M. Jr, and Malik, H.S. (2012) Convergent evolution of escape from hepaciviral antagonism in primates. *PLoS Biol.* **10**, e1001282.

29. Sumpter, R. Jr, Loo, Y.M., Foy, E., Li, K., Yoneyama, M., Fujita, T., Lemon, S.M., and Gale, M. Jr. (2005) Regulating intracellular antiviral defense and permissiveness to hepatitis C virus RNA replication through a cellular RNA helicase, RIG-I. *J. Virol.* **79**, 2689-2699.

30. Yi, M., Hu, F., Joyce, M., Saxena, V., Welsch, C., Chavez, D., Guerra, B., Yamane, D., Veselenaik, R., Pyles, R., Walker, C.M., Tyrell, L., Bourne, N., Lanford, R.E., and Lemon, S.M. (2014) Evolution of a cell culture-derived genotype 1a hepatitis C virus (H77S.2) during persistent infection with chronic hepatitis in a chimpanzee. *J Virol.* **88**, 3678-3694.

31. Kolykhalov, A.A., Agapov, E.V., Blight, K.J., Mihalik, K., Feinestone, S.M., and Rice, C.M. (1997) Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* **277**, 570-574.

32. Yanagi, M., St Claire, M., Shapiro, M., Emerson, S.U., Purcell, R.H., and Bukh, J. (1998) Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo. *Virology* **244**, 161-172.

33. Meylan, E., Martinon, F., Thome, M., Gschwendt, M., and Tschopp, J. (2002) RIP4 (DIK/PKK), a novel member of the RIP kinase family, activates NF-kappa B and is processed during apoptosis. *EMBO Rep.* **3**, 1201-1208.

34. Taremi, S.S., Beyer, B., Maher, M., Yao, N., Prosise, W., Weber, P.C., and Malcolm, B.A. (1998) Construction, expression, and characterization of a novel fully activated recombinant single-chain hepatitis C virus protease. *Protein Sci.* **7**, 2143-2149.

35. Niesen, F.H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212-2221.

36. Copeland, R.A. (2013) Evaluation of enzyme inhibitors in drug discovery: A guide for medicinal chemists and pharmacologists, 2nd Ed., Wiley-Interscience, New Jersey.

37. Beran, R.K., and Pyle, A.M. (2008) Hepatitis C viral NS3-4A protease activity is enhanced by the NS3 helicase. *J. Biol. Chem.* **283**, 29929-29937.

38. Ferreira, A.R., Magalhães, A.C., Camões, F., Gouveia, A., Vieira, M., Kagan, J.C., and Ribeiro, D. (2016) Hepatitis C virus NS3-4A inhibits the peroxisomal MAVS-dependent antiviral signalling response. *J Cell Mol Med.* **20**, 750-757.
39. 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.

40. Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H.P., Ascenzi, P., and Bolognesi, M. (1992) Crystal and molecular structure of the bovine alpha-chymotrypsin-eglin c complex at 2.0 A resolution. *J. Mol. Biol.* **225**, 107-123.

41. Rodriguez, A., Oliva, C., González, M., van der Kamp, M., and Mulholland, A.J. (2007) Comparison of different quantum mechanical/molecular mechanics boundary treatments in the reaction of the hepatitis C virus NS3 protease with the NS5A/5B substrate. *J. Phys. Chem. B* **111**, 12909-12915.

42. Case, D.A., Berryman, J.T., Betz, R.M., Cerutti, D.S., Cheatham, D.E. III, Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T.S., LeGrand, S., Li, P., Luchko, T., Luo, R., Madej, B., Merz, K.M., Monard, G., Needham, P., Nguyen, H., Nguyen, H.T., Omelyan, I., Onufriev, A., Roe, D.R., Roitberg, A., Salomon-Ferrer, R., Simmerling, C.L., Smith, W., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu, X., and Kollman, P.A. (2015) AMBER 2015, University of California, San Francisco.

43. Berendsen, H.J.C., Grigera, J.R., and Straatsma, T.P. (1987) The Missing Term in Effective Pair Potentials. *J Phys Chem* **91**, 6269-6271.

44. Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999-2012.

45. Duell, E.R., Glaser, M., Le Chapelain, C., Antes, I., Groll, M., and Huber, E.M. (2016) Sequential Inactivation of Gliotoxin by the S-Methyltransferase TmtA. *ACS Chem. Biol.* **11**, 1082-1089.

46. Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431-432.

47. Roe, D.R., and Cheatham, T.E. 3rd. (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory. Comput.* **9**, 3084-3095.

48. Miller, B.R. 3rd, McGee, T.D. Jr, Swails, J.M., Homeyer, N., Gohlke, H., and Roitberg, A.E. (2012) MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory. Comput.* **8**, 3314-3321.

49. Grossfield, A., and Zuckerm. (2009) Quantifying uncertainty and sampling quality in biomolecular simulations. *Annu. Rep. Comput. Chem.* **5**, 23-48.

50. Prongay, A.J., Guo, Z., Yao, N., Pichardo, J., Fischmann, T., Strickland, C., Myers, J. Jr, Weber, P.C., Beyer, B.M., Ingram, R., Hong, Z., Prossie, W.W., Ramanathan, L., Taremi, S.S., Yarosh-Tomaine, T., Zhang, R., Senior, M., Yang, R.S., Malcolm, B., Arasappan, A., Bennett, F., Bogen, S.L., Chen, K., Jao, E., Liu, Y.T., Lovey, R.G., Saksena, A.K., Venkatraman, S., Girjavallabhan, V., Njorge, F.G., and Madison, V. (2007) Discovery of the HCV NS3/4A protease inhibitor (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(S)-[[(1,1-dimethylthethyl)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.* **50**, 2310-2318.

51. Schechter, I., and Berger, A. (2012) On the size of the active site in proteases. I. Papain. 1967. *Biochim. Biophys. Res. Commun.* **425**, 497-502.
proteomics identifies the membrane-associated peroxidase GPx8 as a cellular substrate of the hepatitis C virus NS3-4A protease. *Hepatology*. 59, 423-433.

55. Kang, X., Chen, X., He, Y., Guo, D., Guo, L., Zhong, J., and Shu, H.B. (2013) DDB1 is a cellular substrate of NS3/4A protease and required for hepatitis C virus replication. *Virology*. 435, 385-394.
ABBREVIATIONS AND NOMENCLATURE

c.l.       cleaved
DAA       direct acting antiviral agent
DNA       deoxyribonucleic acid
EC50      half maximal effective concentration
FC        fold change
FI        fold induction
f.l.      full length
FMDV      foot-and-mouth disease virus
FRET      Förster/fluorescence resonance energy transfer
GLuc      Gaussia luciferase
GT        genotype
HCV       hepatitis C virus
Huh-7.5   hepatocyte derived cellular carcinoma cell line
IFN       interferon
kcat      substrate turnover rate
K_M       maximum rate of reaction (Michaelis-Menten constant)
MAVS      mitochondrial antiviral signaling protein
MD        molecular dynamics
n/a       not applicable
n/d       not done
ns        nanosecond
NS        non-structural
N-terminal amino-terminal
PCR       polymerase chain reaction
PDB       Protein Databank
PI        protease inhibitor
RIG-I      retinoic acid inducible gene I
RNA       ribonucleic acid
S.D.      standard deviation
Tm        temperature with a half maximal denaturation
U-2 OS    human bone osteosarcoma epithelial cell line
WT        wildtype
### Table 1. HCV RNA replication and EC_{50} of simeprevir for PI-resistant variants.

| Residue | Variant | Replication capacity (%) * | EC_{50} (nM) §  |
|---------|---------|-----------------------------|------------------|
|         |         | NS3-Q41R | NS3-Q41 | NS3-Q41R | NS3-Q41 |
| WT      |         | 100      | 100     | 6.4, 0.3 | 6.4, 0.3 |
| V36     | A       | 50.4     | 89.7    | 56, 3.5  | 14, 6.2  |
|         | G       | 2.6      | 20.8    | n/a      | 42, 5.2  |
|         | L       | 76.7     | 100.8   | 38, 2.6  | 12, 2.0  |
|         | M       | 27.8     | 94.7    | 37, 2.6  | 11, 1.3  |
|         |         |          |         |          |          |
| Q41     | R #     | n/a      | 99.9    | 29, 0.4  | 29, 0.4  |
| F43     | S       | 3.2      | 70.2    | 197, 24  | 67, 3.2  |
| T54     | A       | 14.5     | 68.0    | 7.3, 0.2 | 3.2, 0.6 |
| Q80     | R       | 24.1     | 51.4    | 239, 16  | 48, 5.3  |
| S138    | T       | 1.4      | 0.2     | 185, 35  | n/a      |
|         | K       | 23.2     | 47.1    | >1000    | 259, 8.8 |
|         | T       | 4.6      | 71.1    | 162, 21  | 85, 3.7  |
|         | G       | 2.7      | 45.8    | n/a      | 74, 5.7  |
|         | W       | 2.5      | 28.1    | n/a      | 544, 20  |
| R155    | G       | 41.7     | 45.8    | 329, 26  | 84, 7.7  |
|         | S       | 62.5     | 50.9    | 5.8, 1.0 | 2.1, 1.0 |
|         | T       | 0.88     | 2.5     | n/a      | 270, 21  |
|         | V       | 1.0      | 0.3     | n/a      | n/a      |
| A156    | A       | 90.5     | 34.7    | >1000    | 2047, 270 |
|         | E       | 141.0    | 97.9    | 504, 65  | 120, 4.8 |
|         | H       | 80.0     | 55.4    | >1000    | 2228, 103 |
|         | I       | 20.1     | 43.7    | >1000    | 6935, 469 |
|         | T       | 84.7     | 97.9    | >1000    | 2025, 77 |
|         | V       | 62.5     | 47.0    | >1000    | 8070, 624 |
|         | Y       | 55.4     | 54.2    | n/d      | 2640, 584 |

**Table legend:** HCV RNA replication and EC_{50} of simeprevir for PI-resistant variants in an NS3-Q41 (H77S.3) or NS3-Q41R (H77S.2) genomic background. * Replication capacity relative to WT (%); § EC_{50} (nM; mean, S.D.); # H77S.2 is R41 and H77S.3 is Q41; n/a: Replication too low to measure EC_{50}; n/d: Mutant not prepared.
Table 2. Enzyme kinetic constants and protein melting temperatures for NS3-4A protease variants.

| Table legend: | Kinetic constants (upper panel) and melting temperatures (lower panel) were obtained by a FRET-based protease and thermal shift assay, respectively. Data shown represent the mean ± S.D. from at least three independent experiments. |
| --- | --- |
| **Substrate** | **Variant** | **Enzyme kinetics** |
| | | $K_M$ (µM) | $k_{cat}$ (/s) | $k_{cat}/K_M$ 1/(s·µM) |
| NS4A/4B | WT (Q41) | 3.11 ± 0.29 | 0.19 ± 0.011 | 0.062 ± 0.008 |
| | Q41R | 2.11 ± 0.14 | 0.18 ± 0.006 | 0.085 ± 0.008 |
| | D168A | 2.45 ± 0.23 | 0.09 ± 0.004 | 0.035 ± 0.004 |
| | Q41R-D168A | 1.71 ± 0.23 | 0.15 ± 0.009 | 0.087 ± 0.015 |
| | D168E | 1.79 ± 0.11 | 0.12 ± 0.003 | 0.067 ± 0.005 |
| | Q41R-D168E | 1.83 ± 0.10 | 0.14 ± 0.001 | 0.074 ± 0.005 |
| | D168T | 2.03 ± 0.15 | 0.13 ± 0.005 | 0.064 ± 0.006 |
| | Q41R-D168T | 1.69 ± 0.14 | 0.12 ± 0.004 | 0.071 ± 0.007 |
| MAVS | WT (Q41) | 7.82 ± 0.70 | 0.058 ± 0.002 | 0.008 ± 0.001 |
| | Q41R | 37.32 ± 4.42 | 0.275 ± 0.021 | 0.007 ± 0.001 |
| | D168A | 14.41 ± 2.91 | 0.029 ± 0.003 | 0.002 ± 0.000 |
| | Q41R-D168A | 1.77 ± 0.14 | 0.507 ± 0.011 | 0.287 ± 0.024 |

| **Variant** | **Melting temperature** |
| --- | --- |
| | $T_m$ (°C) |
| WT (Q41) | 54.9 ± 0.3 |
| Q41R | 55.1 ± 0.3 |
| D168A | 48.7 ± 0.1 |
| Q41R-D168A | 48.3 ± 0.2 |
| D168E | 55.9 ± 0.2 |
| Q41R-D168E | 55.8 ± 0.3 |
| D168T | 47.1 ± 0.1 |
| Q41R-D168T | 48.5 ± 0.2 |
### Table 3. Average interaction energies of the NS3-4A protease-substrate complexes.

| Variant     | Substrate   | ΔE<sup>b</sup> | ΔΔE<sup>c</sup> | σ(ΔE)<sup>d</sup> |
|-------------|-------------|----------------|-----------------|------------------|
| WT (Q41)    | NS4A/4B     | -558.85        | -3%             | 8.34             |
|             | N-term NS4A/4B<sup>a</sup> | -541.94        |                 | 3.79             |
| Q41R-D168A  | NS4A/4B     | -719.46        | +9%             | 11.33            |
|             | N-term NS4A/4B<sup>a</sup> | -783.06        |                 | 7.80             |
| WT (Q41)    | MAVS        | -184.57        | +49%            | 6.16             |
|             | N-term MAVS<sup>a</sup> | -274.70        |                 | 6.22             |
| Q41R-D168A  | MAVS        | -237.56        | +77%            | 4.22             |
|             | N-term MAVS<sup>a</sup> | -420.34        |                 | 3.29             |

**Table legend:** Average interaction energies in kcal/mol from the last 100 ns of MD simulations of the NS3-4A protease-substrate complexes.  
<sup>a</sup> N-terminal cleavage product;  
<sup>b</sup> Average interaction energy without solvent;  
<sup>c</sup> Relative interaction energy difference of the Q41R-D168A mutant compared to WT;  
<sup>d</sup> Error of uncorrelated samples.
FIGURES

Figure 1. Replicative fitness of PI-resistant variants in NS3±Q41R protease genomic backgrounds. (A) Structural context of PI-resistance mutations depicted on the NS3-4A protease structure from PDB-2OC8 (50) with PI bound to the ligand-binding site (purple stick model); ligand-binding pockets S4 to S1’ (51). The protein surface is partially transparent; protease domain: green, NS4A: light blue. Q41 and D168 are shown as blue stick models located at opposite ends of the protease substrate-binding pocket. (B) The H77S.2 molecular clone contains six cell-culture adaptive mutations (not shown) throughout the viral genome; the only cell-culture adaptive mutation located in the NS3-4A protease domain, NS3-Q41R (depicted by arrow), is reverted back to NS3-Q41 WT in H77S.3 (6). (C) GLuc activity secreted by RNA-transfected cells normalized to pH77S.3 and pH77S.2 WT. Blue shading indicates PI-resistant variants with enhanced replicative fitness by NS3-Q41R. Data shown represent
the mean ± S.D. from at least three independent experiments; *, \( p < 0.001 \); **, \( p < 0.0001 \); by two-sided T-test. WT: wild type; AAG: negative control.
Figure 2.
Impact of NS3-Q41R on the infectious virus production of D168 variants. Comparison of infectious virus yield from D168A/E-protease mutants within NS3-Q41 and NS3-Q41R backbones. Infectious virus production determined from supernatant fluids at 72 and 96 h infected to naïve Huh-7.5 cells as determined by GLuc activity assay. Data shown represent the mean ± S.D. from at least three independent experiments.
Figure 3.
Impact of D168 mutants on the protein fold of the NS3-4A protease ± NS3-Q41R. Data from real-time thermal stability assay using Sypro Orange, a temperature-stable fluorophore that exhibit enhanced fluorescence upon interacting with unfolded proteins. The thermal stability of WT protease and mutants was assessed under increasing incubation temperatures. The impact of mutations on protein unfolding patterns is characterized by fluorescence emission curves (left). Melting temperatures ($T_m$) from purified NS3-4A protease mutants are determined by fitting the sigmoidal melt curve to the Boltzmann equation (right). Error bars represent the mean ± S.D. from at least three independent experiments; *, $p \leq 0.05$; **, $p \leq 0.01$; by two-sided T-test.
Figure 4.
Impact of D168 mutants and NS3±Q41R on protease enzymatic function. (A) Reaction velocity and Michealis-Menten kinetics as assessed from purified protein of protease WT and D168 mutants ± NS3-Q41R and the natural polyprotein substrate NS4A/4B and (B) assessed for the MAVS peptide (each at different peptide concentrations). Reaction constants $K_M$ and $k_{cat}$ were calculated after nonlinear regression curve fitting. Error bars represent the mean ± S.D. from at least three independent experiments; *, $p < 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$; by two-sided T-test.
Figure 5.
Impact of D168 mutants and NS3±Q41R on MAVS cleavage capacity and IFN-β activation. (A) U-2 OS cells were left untransfected (mock), transfected with a GFP expression plasmid (control of transfection efficiency) and transfected with expression constructs harboring NS3-Q41 or NS3-Q41R ± D168A/E/T. MAVS cleavage as assessed by Western blot; one out of five representative experiments is shown. (B) U-2 OS cells co-transfected with IFN-β reporter plasmids, a full-length MAVS expression construct, and D168A/E/T mutant protease constructs ± NS3-Q41R, subsequently analyzed for IFN-β-dependent luciferase activity. Renilla luciferase (RLuc) activity determined to normalized IFN-β activity. Error bars represent the mean ± S.D. from at least three independent experiments; *, p ≤ 0.05; **, p ≤ 0.01; by two-sided T-test; f.l., full-length; cl., cleaved form; FI, fold induction.
Figure 6.
Molecular characterization of protease-substrate interactions. Illustrations of representative protease binding-site conformations (left panel: A, C, E and G) and H-bond networks of protease-substrate complex simulations (right panel: B, D, F and H). (A) 3-D structure and (B) residue-interaction network of the NS4A/4B substrate bound to the NS3-Q41 protease WT or (C) 3-D structure and (D) residue-interaction network for NS4A/4B bound to the Q41R-D168A protease; (E) 3-D structure and (F) residue-interaction network for MAVS bound to the NS3-Q41 protease WT or (G) 3-D structure and (H) residue-interaction network for MAVS bound to the Q41R-D168A protease. Hydrogen bonds present in each time frame were extracted from MD trajectories. Edge widths (single lines) in the residue networks shown represent the average number of H-bonds during simulation time.
and correspond to the individual bond strength; a quantitative measure for the bond strength was
provided by corresponding numbers. Only edges above a minimum occurrence limit of 0.1 are shown.
Cation-π stacking interactions between arginine and histidine residues (time fraction within π stacking
distance cutoff) are shown as double lines with caps. Atom colors: Substrate (green); protease catalytic
triad (red); extended H-bond network of the catalytic triad (blue). H-bonds are shown in magenta (A,
C, E and G) and black (B, D, F and H).
Figure 7. Inhibition of the NS3-4A protease ± Q41R by MAVS cleavage products. Relative reaction velocity for NS4A/4B cleavage as assessed for the purified NS3-Q41 protease WT and NS3-Q41R mutant protease (measured by FRET protease assay) and different preincubation time with MAVS peptide (0, 10 and 20 min). Reference velocity (100%) is the cleavage reaction for the NS4A/4B FRET peptide without MAVS. Error bars represent the mean ± S.D. from at least three independent experiments; *, p ≤ 0.05; **, p ≤ 0.01; by two-sided T-test.
Extended interaction networks with HCV protease NS3-4A substrates explain the lack of adaptive capability against protease inhibitors

Georg Dultz, Tetsuro Shimakami, Markus Schneider, Kazuhisa Murai, Daisuke Yamane, Antoine Marion, Tobias M Zeitler, Claudia Stross, Christian Grimm, Rebecca M Richter, Katrin Bäumer, MinKyung Yi, Ricardo M Biondi, Stefan Zeuzem, Robert Tampé, Iris Antes, Christian M Lange and Christoph Welsch

J. Biol. Chem. published online August 3, 2020

Access the most updated version of this article at doi: 10.1074/jbc.RA120.013898

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts