Epistatic interactions promote persistence of NS3-Q80K in HCV infection by compensating for protein folding instability

Georg Dultz[1,‡], Sanjay K. Srikakulam[2,3,4,‡], Michael Konetschnik[1,‡], Tetsuro Shimakami[5], Nadezhda T. Doncheva[6], Julia Dietz[1], Christoph Sarrazin[1], Ricardo M. Biondi[7], Stefan Zeuzem[1,8], Robert Tampé[9], Olga V. Kalinina[2,10,11], and Christoph Welsch[1,8,*]

From the [1]Department of Internal Medicine 1, Goethe University Hospital Frankfurt, Frankfurt am Main, Germany; [2]Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, Saarland University Campus, Saarbrücken, Germany; [3]Graduate School of Computer Science, [4]Interdisciplinary Graduate School of Natural Product Research, Saarland University, Saarbrücken, Germany; [5]Department of Gastroenterology, Kanazawa University Hospital, Kanazawa, Japan; [6]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark; [7]Molecular Targeting, Instituto de Investigación en Biomedicina de Buenos Aires (IBioBA) - CONICET - Partner Institute of the Max Planck Society, Buenos Aires, Argentina; [8]University Center for Infectious Diseases, University Hospital Frankfurt, Frankfurt am Main, Germany; [9]Institute of Biochemistry, Biocenter, Goethe University Frankfurt, Frankfurt am Main, Germany; [10]Medical Faculty, Saarland University, Homburg, Germany; [11]Center for Bioinformatics, Saarland Informatics Campus, Saarbrücken, Germany

Edited by Craig Cameron

The Q80K polymorphism in the NS3-4A protease of the hepatitis C virus is associated with treatment failure of direct-acting antiviral agents. This polymorphism is highly prevalent in genotype 1a infections and stably transmitted between hosts. Here, we investigated the underlying molecular mechanisms of evolutionarily conserved coevolving amino acids in NS3-Q80K and revealed potential implications of epistatic interactions in immune escape and variants persistence. Using purified protein, we characterized the impact of epistatic amino acid substitutions on the physicochemical properties and peptide cleavage kinetics of the NS3-Q80K protease. We found that Q80K destabilized the protease protein fold ($p < 0.0001$). Although NS3-Q80K showed reduced peptide substrate turnover ($p < 0.0002$), replicative fitness in an H77S.3 cell culture model of infection was not significantly inferior to the WT virus. Epistatic substitutions at residues 91 and 174 in NS3-Q80K stabilized the protein fold ($p < 0.0001$) and leveraged the WT protease stability. However, changes in protease stability inversely correlated with enzymatic activity. In infectious cell culture, these secondary substitutions were not associated with a gain of replicative fitness in NS3-Q80K variants. Using molecular dynamics, we observed that the total number of residue contacts in NS3-Q80K mutants correlated with protein folding stability. Changes in the number of contacts reflected the compensatory effect on protein folding instability by epistatic substitutions. In summary, epistatic substitutions in NS3-Q80K contribute to viral fitness by mechanisms not directly related to RNA replication. By compensating for protein-folding instability, epistatic interactions likely protect NS3-Q80K variants from immune cell recognition.

The polymorphism Q80K in the NS3-4A protease of hepatitis C virus (HCV) dates back to a single substitution event around the 1940s in the United States and is associated with reduced treatment response on direct-acting antivirals (DAAs) (1–4). Depending on the geographic region, NS3-Q80K is found in up to 48% of untreated genotype 1a–infected patients, whereas it is rarely seen in genotype 1b infection (approximately 0.5%) (5). Although the resistance level for NS3-Q80K against protease inhibitors (PIs) was determined to be rather low, only approximately 10-fold change (FC) (6), its clinical consequences are striking. In large interferon-based clinical trials of the PI simeprevir, Q80K was highly associated with genotype 1a–treatment failure leading to response rates not superior to pegylated interferon-α plus ribavirin (7). Dietz *et al.* (8) reported an increase from 35% to 56% of Q80K as a resistance-associated amino acid substitution in genotype 1a for simeprevir in combination with the nucleotide NS5B inhibitor sofosbuvir. A similar although weaker increase in Q80K is reported for the 3D regimen (PrOD), a triple combination comprising the PI paritaprevir/ritonavir, the NS5A inhibitor ombitasvir, and the non-nucleoside NS5B inhibitor dasabuvir. Notably, 41% of individuals with failure on 3D and NS5A-RASs harbored NS3-Q80K, which is a higher prevalence compared with 35% in DAA-naïve patients from this study (8). The long phylogenetic history of NS3-Q80K, however, is in sharp contrast with those of other drug resistance mutations and indicates that it does not arise as a response to drug selection (1).

By reconstruction of the evolutionary history of the Q80K polymorphism, epistatic residue interactions are identified

descendent from a single NS3-Q80K lineage that may compensate stabilizing or functional properties of the protease structure, S174N and A91S/T[1]. Epistasis is widely recognized as a key phenomenon that drives the dynamics of evolution. It occurs when the combined effect of two or more mutations differs from the sum of their individual effects and arguably reflects molecular interactions that affect the function of a protein (9, 10). Characterizing high-order epistasis between three or more mutations can unveil hidden intramolecular interaction networks that underlie essential molecular mechanisms related to protein functions and their evolution (11). Multidomain proteins such as the NS3 protein of HCV are complex molecular machines with their biophysical and functional properties not simply additive. The N-terminal protease of NS3 is characterized by a bilobal structure that requires the structural complementation by the NS4A peptide derived from the viral polyprotein (12). By proteolytically processing nonstructural proteins from the viral polyprotein, the protease is involved in RNA synthesis (13). In addition, the capacity to cleave and inactivate the mitochondrial antiviral-signaling protein MAVS (also called Cardif, VISA, and IPS-1) in the retinoic acid–inducible gene I signaling pathway is a cardinal feature of NS3-4A, by which HCV blocks induction of interferon-β (14–18). The NS4A peptide intercalates into the protease structure to support proper protein folding and full enzymatic activity (13). A linker connects the protease domain to a C-terminal RNA helicase with both domains being highly interdependent (19, 20). By interference with the NS3 helicase domain, the protease modulates infectious virus particle assembly (20, 21). The NS3-4A protease is a prime drug target against HCV (20). Its structure and function is conserved along the family of flaviviridae and is a target for new drugs against emerging human pathogens, such as dengue, West Nile, or yellow fever virus (22). RASs in NS3-4A of HCV frequently lead to considerable fitness costs that for most mutations result from reductions in RNA replication capacity and for some in impairment in their ability to produce infectious virus (23).

In the present study, we show that epistatic interactions in the NS3-4A protease stabilize the protein fold of NS3-Q80K variants and leverage protein stability to the WT level. In agreement with the overall protein stability of NS3-Q80K, we find the total number of noncovalent residue interactions increasing upon amino acid substitution at epistatic interaction sites, although some of these residues do not immediately interact with residues at position 80. Changes in protein stability because of epistatic interactions showed a trade-off against protease enzymatic activity and replicative fitness deficits in infectious cell culture. The finding that epistatic residue interactions are not associated with a gain-of-function phenotype in replicative fitness of NS3-Q80K variants appears surprising, given the high prevalence of these substitutions in genotype 1a. Our data suggest that immune-escape mechanisms could contribute to the overall fitness of NS3-Q80K variants *in vivo*.

Results

Rational approach

We explored epistasis and mutual dependence of distinct sites in the NS3-4A protease structure as a possible mechanism to explain the prevalence of NS3-Q80K in genotype 1a HCV infection. The mutant patterns are selected based on an article by McCloskey *et al.* (1) We analyzed sequence data of genotype 1a–infected patients with NS3-Q80K for epistatic secondary substitutions at residue 91 and/or 174 based on sequence information from DAA-experienced patients from the Frankfurt Resistance Database and sequence information from DAA-naïve patients from the European HCV database (24) (see supporting information). In DAA-experienced patients, we found triple mutant patterns, that is Q80K-A91S-S174N and Q80K-A91T-S174N, selected over double mutant patterns, Q80K-A91S/T and Q80K-S174N (Fig. 1). Subsequently, we characterized the impact of epistatic residue interactions on the physicochemical properties of the protease by synthetic biology using purified protein and complementary bioinformatics approaches. We then used a cell culture model of HCV infection to assess implications of epistatic residue interactions in NS3-4A for the replicative fitness of NS3-Q80K variants.

Epistatic secondary substitutions in NS3-Q80K stabilize the protease protein fold

To characterize the impact of NS3-Q80K and its epistatic amino acid substitutions S174N and A91S/T on the protease protein fold, we expressed and purified the NS3-4A WT fusion and different mutant forms (see Experimental procedures section). We tested if NS3-Q80K ± S174N and/or A91S/T had an effect on the protease stability (Figs. 2 and 3 and
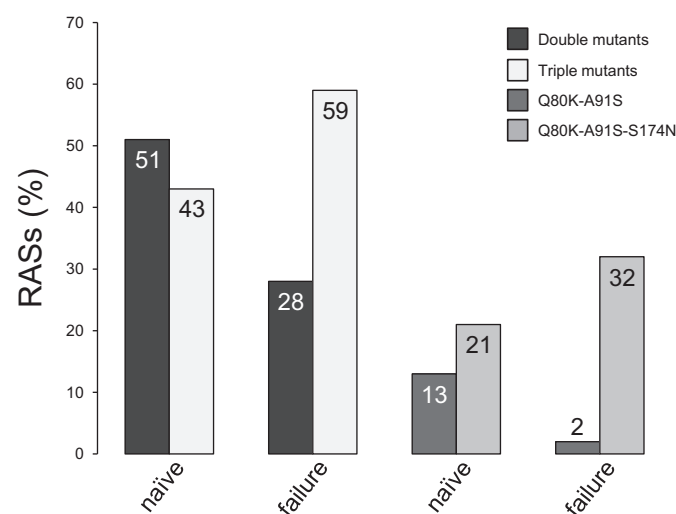

Figure 1. Prevalence of epistatic secondary substitutions in DAA-naïve and DAA-experienced patients harboring NS3-Q80K. Comparison of mutant patterns in patients with genotype 1a HCV infection harboring NS3-Q80K between the European HCV database (24), euHCVdb (DAA naïve, n = 100), and DAA-experienced patients harboring NS3-Q80K from the Frankfurt Resistance Database (upon DAA failure, n = 41). Data shown represent percentages (%) of patients from which the respective variants were isolated as major variants before DAA treatment or upon treatment failure. DAA, direct-acting antiviral; HCV, hepatitis C virus.

Table 1). The impact of mutants on protein folding was determined by a thermal shift assay, where protein denaturation at increasing temperatures is revealed by increased fluorescence (25). The NS3-4A protease WT was stable to temperature with a half maximal denaturation, T_m, at 50.6 °C (Figs. 2 and 3), whereas the T_m of NS3-Q80K revealed a significant destabilizing effect on protein stability, T_m at 44.1 °C ($p < 0.0001$) (Figs. 2 and 3). We find that all epistatic amino acid substitutions shifted the denaturation curve to higher temperatures and stabilized the protease fold of the respective expressed NS3-Q80K mutant proteins (Fig. 2). All mutants, but Q80K-A91S, showed compensation in protein stability that equals the level of the WT protein (Fig. 3). Notably, the double mutants Q80K-A91T (T_m at 49.6 °C) ($p < 0.0001$) and Q80K-S174N (T_m at 48.4 °C) ($p < 0.0001$) and both triple mutants, Q80K-A91S-S174N (T_m at 48.6 °C) ($p = 0.018$) and Q80K-A91T-S174N (T_m at 50.1 °C) ($p < 0.0001$), showed no difference in protein stability with the WT protease (Fig. 3 and Table 1). Only weak stabilization of the protease fold ($p < 0.04$) was observed for Q80K-A91S (Fig. 2). Among all mutant patterns, the melting curve of the triple mutant Q80K-A91T-S174N closest resembles the curve of the WT protease protein (Fig. 2). Interestingly, the denaturation curves do not show the expected denaturation curves parallel to the denaturation curve of the WT. Instead, the Q80K mutation shows a clear deviation from the expected curve and reproducibly possesses a shoulder during denaturation (Fig. 2*A*). The result suggests that NS3-Q80K differentially affect the two subdomains of the protease that build the bilobal structure of the protein. In solution, upon heating, the protease subdomain comprising the Q80K mutation may denature separately. Notably, epistatic secondary substitutions that compensate Q80K also show a clear shoulder (Fig. 2, *B–F*) and a denaturation curve that does not parallel the denaturation of the WT protein. It appears then that NS3-Q80K not only destabilizes the protein but also may affect the communication between the two lobes of the protease. Since the active site of the protease is located in the region corresponding to the connection between the two lobes of the protease, the NS3-Q80K mutation and the epistatic secondary substitutions may affect the protease catalytic activity.

Effect of mutations on the catalytic activity of the protease

We considered it likely that epistatic amino acid substitutions in NS3-4A might rescue replicative fitness deficits
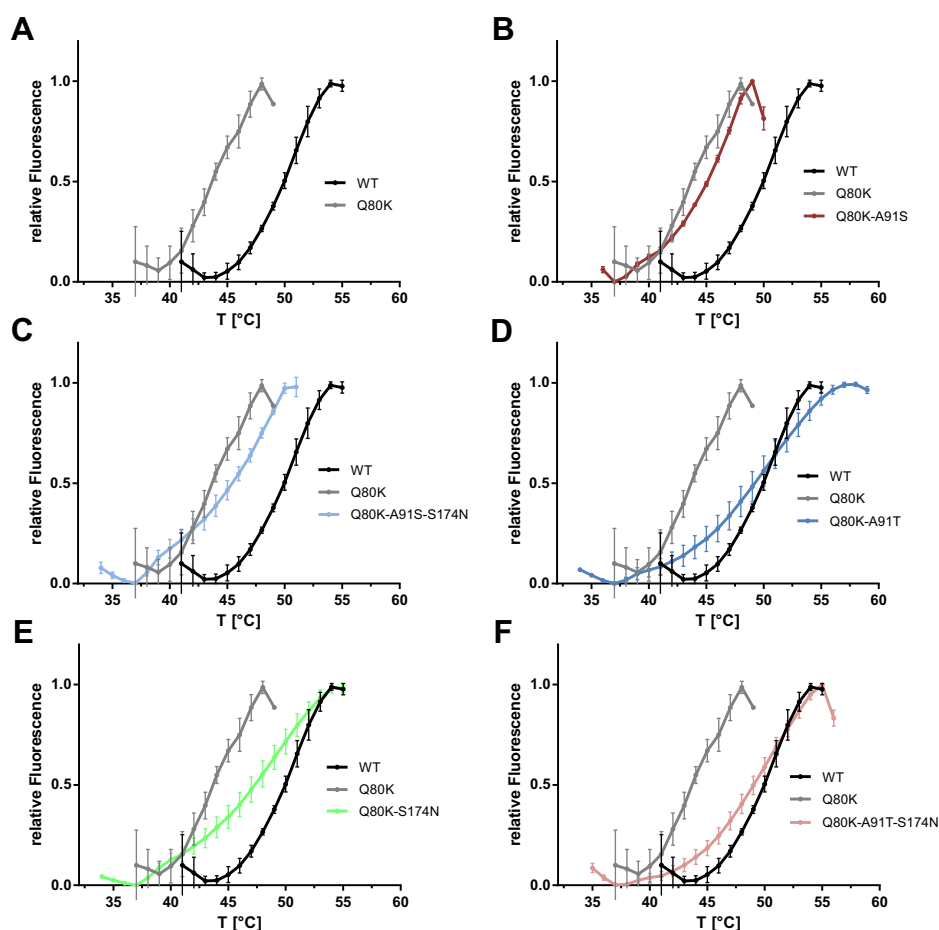
Figure 2. Thermal shift of the NS3-4A protease harboring NS3-Q80K and epistatic amino acid substitutions. Data from real-time thermal stability assay using Sypro Orange, a temperature-stable fluorophore that exhibit enhanced fluorescence upon interacting with unfolded proteins. *A*, thermal stability of the WT protease and the NS3-Q80K mutant, as well as (*B–F*), mutants harboring epistatic secondary substitutions assessed under increasing incubation temperatures. Graphs showing impact of mutations on protease protein unfolding patterns characterized by fluorescence emission curves.
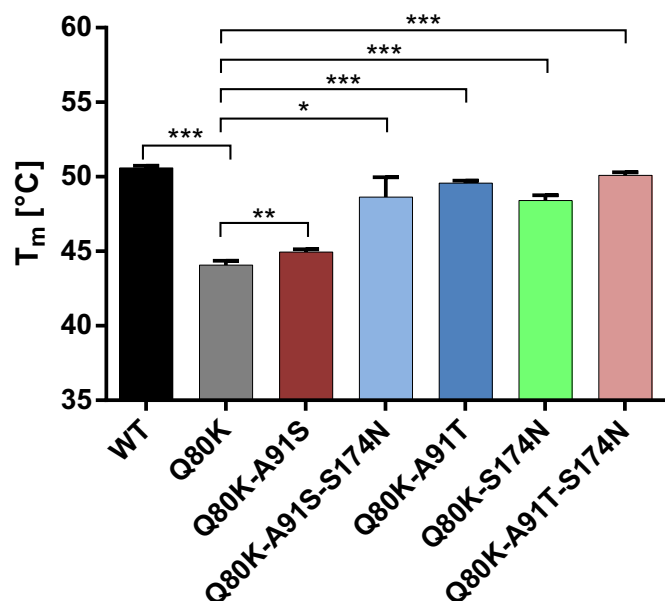
Figure 3. Impact of NS3-Q80K and epistatic amino acid substitutions on the protease protein fold. T_m from purified NS3-4A protease WT and mutants are determined by fitting the sigmoidal melt curve as shown in Figure 2 to the Boltzmann equation. Error bars represent the mean ± SD from at least three independent experiments; *$p \leq 0.05$; **$p \leq 0.01$; and ***$p \leq 0.001$; by two-sided *t* test.

from NS3-Q80K and potentially explain the high prevalence of this polymorphism. To assess a potential direct impact of the stabilized protein fold on the protease catalytic activity, we analyzed the enzymatic activity using an *in vitro* FRET-based assay that measures the cleavage of a viral polyprotein substrate derived from the viral NS4A/4B polyprotein cleavage site (26). In this assay, fluorescence is released when the protease recognizes and cleaves the FRET substrate (see Experimental procedures section). We tested the kinetic properties of the NS3-4A protease harboring NS3-Q80K ± S174N and/or A91S/T. We determined the amount of substrate needed to obtain half of the protease maximum rate of reaction (K_M) and substrate turnover rates (k_{cat}). The protease harboring NS3-Q80K exhibited a lower substrate turnover rate ($p = 0.0002$; FC = 0.33) and lower Michaelis–Menten constant, K_M, than the NS3-Q80 protease WT ($p = 0.03$; FC = 0.53). Except for the Q80K-A91S-S174N triple mutant (with 8.16 µM), K_M, in all mutant patterns tested, was significantly lower than in NS3-Q80K and only varied in rather narrow

ranges between 0.14 and 0.67 µM (Fig. 4 and Table 1). Roughly, the low K_M indicates higher substrate affinity in the NS3-Q80K protease comprising epistatic amino acid substitutions than in the NS3-Q80 protease WT or NS3-Q80K variant without epistatic secondary substitutions. The protease comprising epistatic secondary substitutions likely requires a lower concentration of peptide substrate to achieve the maximum rate of reaction. However, this was not paralleled by an increase in the substrate turnover rates, k_{cat}, which were significantly inferior to WT for all mutant patterns tested. Only the Q80K-A91S double mutant (equal level; $p = 0.1$) and the Q80K-A91S-S174N triple mutant (increased substrate turnover; $p = 0.015$) showed substrate turnover that leveled up with NS3-Q80K. Noteworthy, in full-length NS3, the helicase domain can stimulate protease activity (19, 20). In our enzymatic activity assay, however, the expressed protease constructs lack the NS3 helicase domain (27) and potentially do not capture indirect mechanisms related to the interaction of the protease and helicase domain of NS3.

Collectively, our biochemistry data show that Q80K clearly decreases the catalytic turnover of the NS3-4A protease, whereas epistatic secondary substitutions do not compensate on this. The *in vitro* catalytic turnover of a cellular protease, however, may not reflect the efficiency by which it accomplishes its physiological role, since the concentration of protease substrates *in vivo* may be limiting. Hence, we conclude that the catalytic turnover may not be limiting viral replication. Alternatively, the destabilization on the Q80K lobe of the protease, as suggested by the temperature stability experiments, could render a compensating effect. We speculated that destabilization of protein regions potentially enable optimized interaction with protease physiological substrates. In line with this idea, the decreased K_M toward the polyprotein peptide substrate (Fig. 4B) already suggests that peptide substrates may bind with higher affinity to NS3-4A in the context of the destabilized protein.

Protein dynamics and inter-residue contacts explain protease fold stabilization in NS3-Q80K harboring epistatic secondary substitutions

The Q80 residue resides in the protease–helicase domain interface of NS3 close to the protease active site (Fig. 5A). The epistatic site S174 resides in close proximity to Q80 (approximately 3.9 Å distance), exhibiting hydrogen bonds between

Table 1
Enzyme kinetic constants and protease protein T_ms for the WT protease and NS3-Q80K variants with epistatic amino acid substitutions

Variant	Enzyme kinetics		Variant	T_m (°C)
	K_M (µM)	k_{cat} (/min)		
WT (Q80)	2.71 ± 0.31	5.46 ± 0.26	WT (Q80)	50.6 ± 0.15
Q80K	1.44 ± 0.24	1.8 ± 0.1	Q80K	44.1 ± 0.29
Q80K-A91S	0.56 ± 0.06	1.56 ± 0.06	Q80K-A91S	45.0 ± 0.19
Q80K-A91T	0.48 ± 0.13	0.4 ± 0.03	Q80K-A91T	49.6 ± 0.16
Q80K-S174N	0.14 ± 0.03	0.48 ± 0.02	Q80K-S174N	48.4 ± 0.37
Q80K-A91S-S174N	8.16 ± 4.62	0.13 ± 0.04	Q80K-A91S-S174N	48.6 ± 1.33
Q80K-A91T-S174N	0.67 ± 0.09	2.41 ± 0.11	Q80K-A91T-S174N	50.1 ± 0.21

Kinetic constants and melting temperatures obtained by FRET-based protease and thermal shift assay, respectively. Data shown represent the mean ± SD from at least three independent experiments.
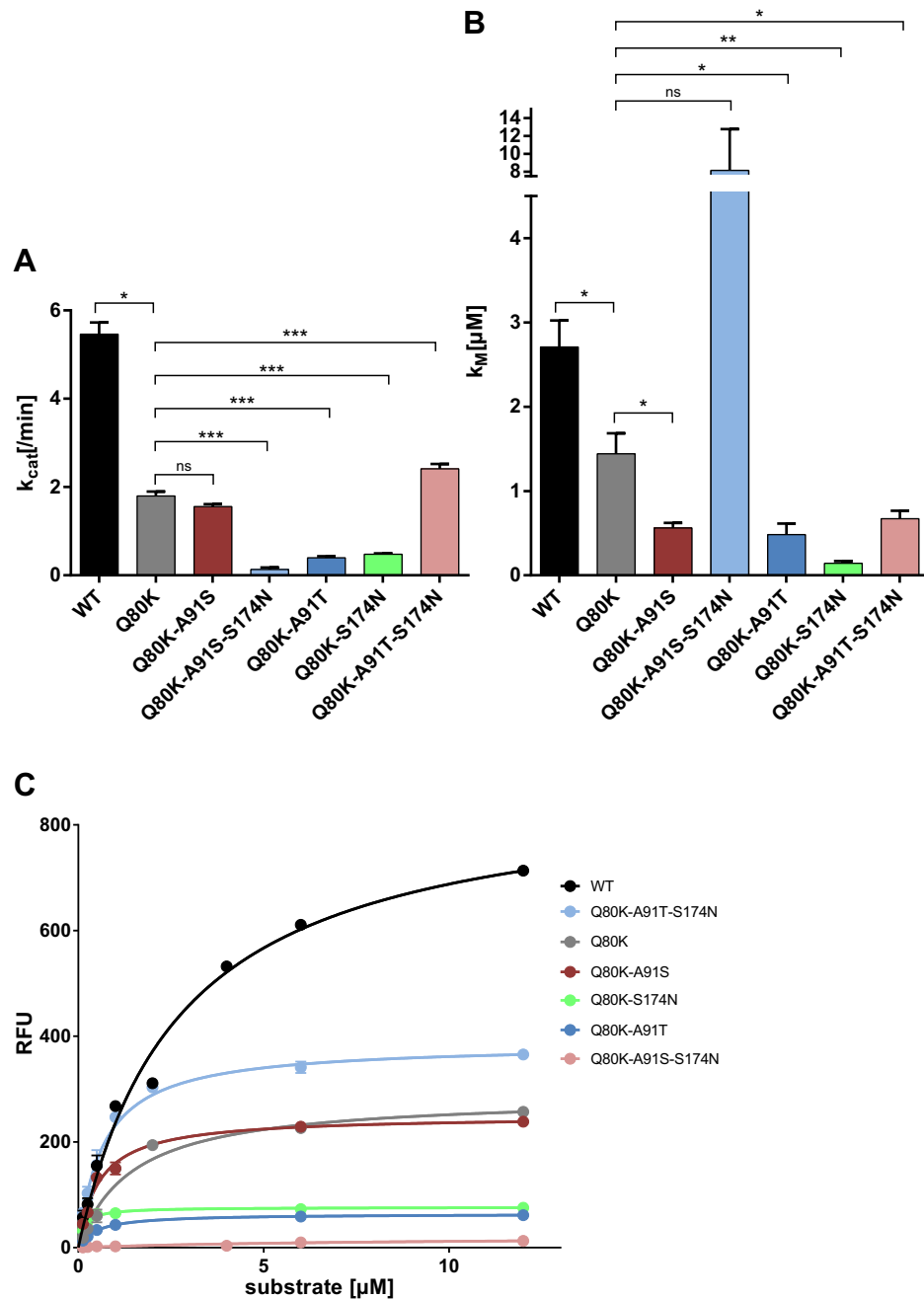
Figure 4. Impact of NS3-Q80K and epistatic amino acid substitutions on the protease enzymatic function and reaction velocity. Reaction velocity and Michaelis–Menten kinetics as assessed from purified protein of WT protease and NS3-Q80K mutants and the natural polyprotein substrate NS4A/4B. *A*, reaction constants k_{cat} and (*B*) K_M calculated after nonlinear regression curve fitting. Error bars represent the mean ± SD from at least three independent experiments; *$p \leq 0.05$; **$p \leq 0.01$; and ***$p \leq 0.001$; by two-sided *t* test. *C*, kinetic progress curves with reaction velocities (RFU, relative fluorescence unit) for the NS3-4A protease and the natural polyprotein substrate NS4A/4B as assessed from purified protein of WT protease and NS3-Q80K mutants.

both residues. In contrast, the second epistatic site A91 is more than 20 Å distant from residue Q80 (Fig. 5*A*). For simulations, we created protease mutant structures corresponding to NS3-Q80K with or without epistatic amino acid substitutions (equivalent to mutants from Table 1).

During molecular dynamics (MD) simulation, we observe destabilization of the protease structure upon NS3-Q80K mutation in the N-terminal lobe of the protease (approximately residues 7–78). Destabilization from NS3-Q80K in part is alleviated by the epistatic secondary substitutions S174N

and/or A91S/T. The modeling results are in line with the particular protein denaturation curves that we observed for NS3-Q80K and epistatic secondary substitutions (Fig. 5, *B–F*). They indirectly validate our hypothesis on the differential impact of mutations on the bilobal structure of the protease. Particularly, mutant patterns comprising S174N or A91T showed fold-stabilizing potential, as witnessed by root mean square fluctuation values for each trajectory (Fig. 5*B*). We have quantified this observation by extracting 181 frames from the concatenated MD trajectory and computing the number of

SASBMB

*J. Biol. Chem.* (2021) 297(3) 101031 **5**
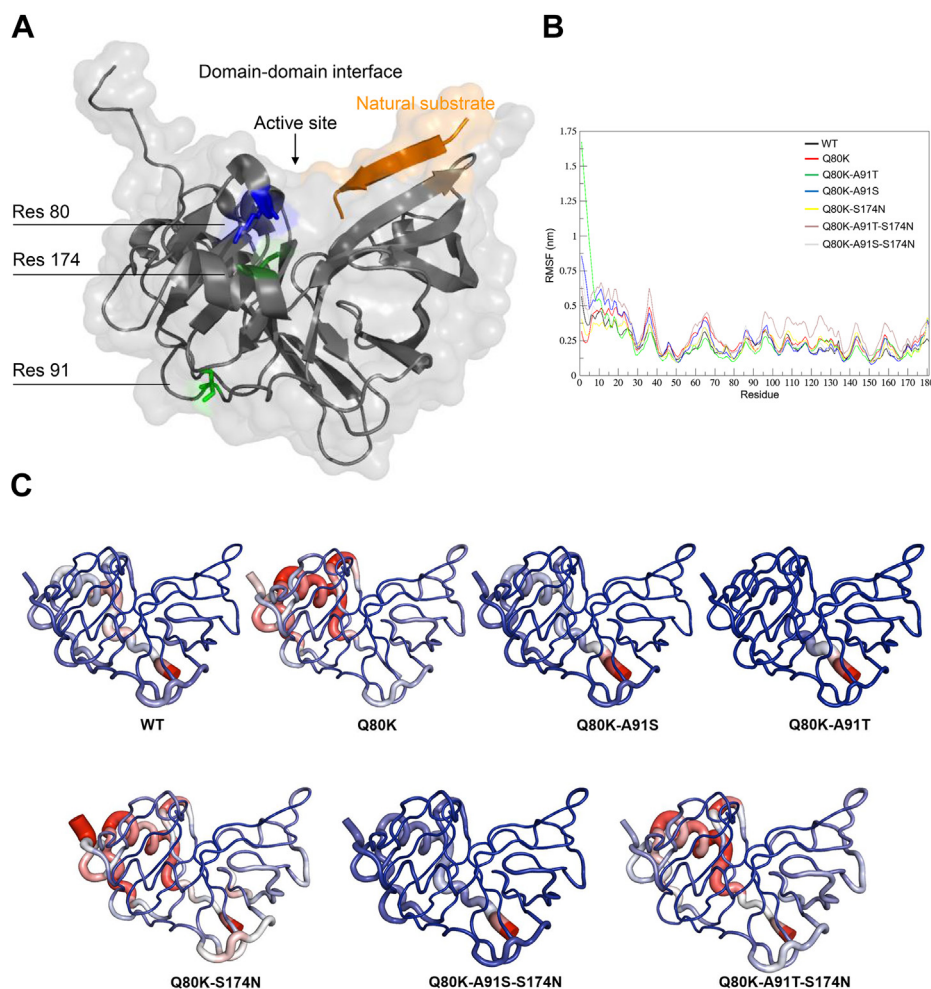
Figure 5. Protease fold and protein flexibility estimated by root mean square fluctuations (RMSFs). *A*, protease structure with cocrystallized natural peptide substrate (in *orange*) (from Protein Data Bank [PDB] ID: 3M5O (62)); the protein backbone given as *ribbon* model with transparent surface depiction. Residues of interest highlighted as *stick* models and colored as follows: residue 80, *blue*; residue 91, *light green*; and residue 174, *dark green*. *B*, impact of NS3-Q80K and epistatic secondary substitutions on the protease protein flexibility estimated by RMSFs over time for different NS3-Q80K mutant patterns. *C*, protease protein flexibility averaged over time for protease WT (from PDB ID: 1CU1 (48)) and NS3-Q80K mutants represented as tubes. Epistatic secondary substitutions differentially affect the bilobal fold of the protease. We used GROMACS to generate RMSF B-factors added to the PDB file. We then used PyMOL defaults to visualize flexible, stable, and intermediate residues. Highly flexible residues are shown in *red*, highly stable residues are shown in *blue*, and the intermediate flexible residues are shown in *white*. The size of the tubes is proportional to the observed RMSF values in those regions.

contacts for each of the amino acid substitutions at positions 80, 91, and 174 (Fig. 5 and Fig. S1). We observe that while the number of contacts of individual positions fluctuates, and in particular, the number of contacts of position 80 decreases upon Q80K mutation, the total number of contacts agrees well with the stability measurements (Fig. 3) and the root mean square fluctuation values from the MD trajectories. Interestingly, new contacts and hydrogen bonds of the mutated amino acid at position 91 contribute most to this trend. This finding agrees with previous evidence that long-range contacts generally make greater thermodynamic impact than contacts that are near in sequence and structure (28).

Epistatic secondary substitutions impair the replicative fitness of NS3-Q80K variants

Based on the biochemistry data and MD simulations (see previously), we speculated that epistatic interactions in NS3-Q80K by stabilizing the protease protein fold could support the replicative fitness of viral variants by mechanisms related to the protease–helicase domain interplay. To assess the replication capacity of NS3-Q80K ± epistatic secondary substitutions, the amino acid substitutions were created within the background of the genotype 1a H77S.3 genome, and their impact on replication of the viral RNA was determined in RNA-transfected cells. H77S.3/Gaussia luciferase (GLuc)2A RNA was assessed from NS3-Q80 protease WT and the NS3-Q80K protease ± epistatic secondary substitutions by measuring GLuc activity in supernatant media collected at 24-h intervals after transfection of synthetic RNA, as described previously. Results were normalized to the GLuc activity present at 8 h after transfection, as this represents GLuc expressed directly by the transfected input RNA. An H77S.3 virus comprising a protease with Q80-A91-S174 sequence pattern is denoted WT virus. This virus was able to replicate efficiently in the cell culture model of infection (Fig. 6).
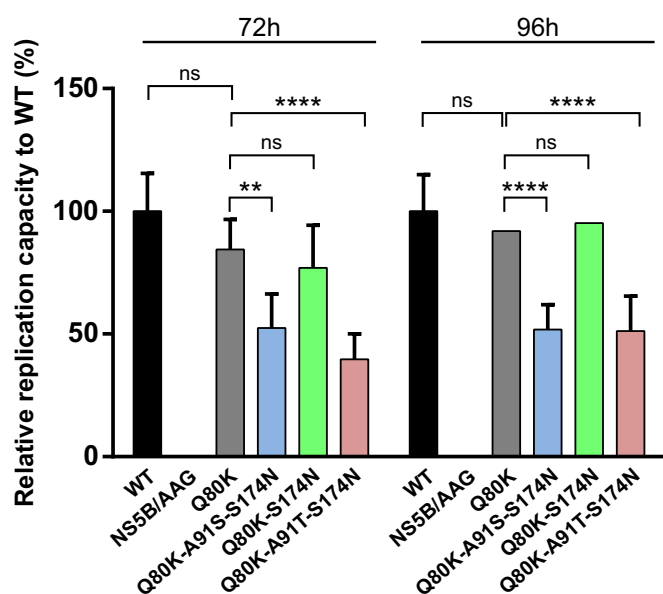
Figure 6. Impact of epistatic amino acid substitutions on RNA replication of NS3-Q80K variants. Medium was collected and replaced at 8, 24, 48, 72, and 96 h after transfection of H77S.3/GLuc2A RNAs carrying the indicated mutations, and GLuc activity determined at 72 h (*left panel*) and 96 h in time (*right panel*). Results normalized to the 8 h GLuc activity represent the mean of triplicate samples and are representative of multiple experiments. Error bars represent the mean ± SD from at least three independent experiments; *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$; and ****$p \leq 0.0001$; by two-way ANOVA. GLuc, Gaussia luciferase.

We find that the Q80K polymorphism in the H77S.3 backbone showed only minimal loss of replication capacity at 72 h after transfection compared with NS3-Q80 protease WT while it reached WT level at 96 h. Unexpectedly, the mutant patterns from our study did not increase the replicative fitness of NS3-Q80K variants. We find the GLuc activity of the triple mutant Q80K-A91T-S174N at 72 and 96 h close to 50% or lower compared with the WT virus (48.8% decrease compared with WT at 96 h; FC = 0.51). In contrast, the GLuc activity of the double mutant Q80K-S714N was only moderately impaired at 72 h (23.1% decrease compared with WT; FC = 0.77) and was not significantly different at 96 h (4.8% decrease compared with WT; FC = 0.95). However, none of the epistatic amino acid substitutions demonstrated enhanced RNA replication capacity compared with the parental H77S.3/GLuc2A RNA or NS3-Q80K.

Discussion

Fitness-compensatory mutations are pivotal drivers of viral quasispecies evolution, but structural constraints restrict their evolutionary space (29–31). We previously reported the delicately balanced protease–peptide interactions in NS3-4A that limit the protease adaptive capability to preclude peptide substrate overfitting and competitive enzyme inhibition (32). As a consequence, secondary substitutions to rescue fitness deficits are observed only rarely in NS3-4A. A study by McCloskey *et al.* (1), however, identified amino acid substitutions in NS3-4A that epistatically "interact" with NS3-Q80K and obviously contribute to the phylogenetic stability

of the polymorphism in viral variants. The molecular mechanisms underlying epistatic interactions in NS3-Q80K likely differ from previously observed fitness-compensatory mechanisms in PI-resistance mutations (18, 32–35). Albeit clinically important in treatment failure with DAA agents, they so far remained unexplained.

One of the constraints on fast-evolving viruses, such as hepatitis C, and the evolution of proteins is the stability of the folded "native" protein state (36). Essentially, the epistatic amino acid substitutions from our study stabilized the NS3-Q80K protein fold and leveraged the WT protease level. By MD simulations, we find changes in the total number of contacts between protease residues in NS3-Q80K and epistatic amino acid substitutions that reflect destabilizing *versus* stabilizing effects on the protease protein structure. This is in agreement with previous findings from Peres-da-Silva *et al.* (37) reporting variations in hydrogen bond occupancies from epistatic interactions in NS3-Q80K. The total number of contacts in our structural models agreed well with our protein stability measurements. This reflects the compensatory effect on the protein stability from epistatic amino acid substitutions in NS3-Q80K. Interestingly, the bioinformatics analysis shows that Q80K mutation destabilizes the protein by affecting the N-terminal subdomain of the protease fold. The shape of the experimental temperature stability curves can reflect the differential stability of different domains of a protein, with increased formation of a shoulder when the T_ms between two domains are more distant between each other (38). We suggest that, by differentially affecting the stability of the two subdomains of the protease fold, the mutations change the shape of the denaturation curves: the start of the denaturation shifts to lower temperatures for the N-terminal domain of the protease because of destabilization by Q80K with a corresponding small shoulder around 45 °C, unmasking the bilobal protease fold. A more pronounced change in the shape of the denaturation curve is then observed for Q80K with additional mutations at A91 and S174N that reflect the Q80K early denaturation with the extended stability to higher temperatures because of epistatic substitutions. However, thermodynamic stability is subject to various conflicting forces that trade-off against each other, in particular protein function and specificity, which illustrates the complexity of the evolutionary forces (36, 39). Accordingly, we observed a trade-off in protease catalytic activity from increased protein stability in NS3-Q80K harboring epistatic secondary substitutions that also limited the replication fitness of the respective variants. Based on our data, however, it seems unlikely that replicative fitness is a limiting factor in the selection of NS3-Q80K variants in virus evolution. We find no significant deficit in the replication fitness of NS3-Q80K variants that require rescue by secondary substitutions. Murai *et al.* (40) observed a replication boost of Q80K-R155K variants in a genotype 1a infectious cell culture model of infection. However, by using purified protein in our protease activity assay, we find that none of the epistatic secondary substitutions could even leverage the NS3-Q80K protease functional level, but some nearly abolished peptide substrate turnover. Nevertheless, epistatic secondary

substitutions that most effectively compensate protein stability select under DAA pressure in treatment-failure patients harboring NS3-Q80K (Fig. 1). This is in contrast to the prevalence of Q80K-A91S with only minor fold-stabilizing potential (Figs. 1 and 3 and Table S1) that decreased upon DAA failure.

Notably, NS3 has both protease and RNA helicase enzymatic activities that modulate their respective functions and play important roles in the propagation of HCV (19, 20). We considered the possibility that stabilizing the protease protein fold could potentially support domain–domain interactions and thereby contribute to the phylogenetic stability of the Q80K polymorphism. Our protease functional assay, however, lacks the NS3 helicase domain (27) and cannot capture mechanisms related to the domain interplay. However, we could investigate the effect of the mutations on the full-length protein in a cellular model of viral replication that requires the helicase functionality. Hence, we assessed related mechanisms of epistatic amino acid substitutions in NS3-Q80K in an infectious cell culture model. Surprisingly, only Q80K-S174N in cell culture showed replicative fitness that equals the NS3-Q80K level. Yet, none of the epistatic amino acid substitutions boosted the replicative fitness of NS3-Q80K variants. We reasoned that attaining increased stability by epistatic residue interactions likely impact the NS3-Q80K protease ability to adopt novel functions not directly related to the virus replication cycle.

As shown previously, the NS3-4A protease function not only plays an important role for escape from immune pressure (18, 32) but also is a prime epitope for cytotoxic T lymphocyte (CTL)–mediated immune responses (41). T-cell responses against the human leucocyte antigen class I–restricted immunodominant protease epitope NS3$_{1073–1081}$ have frequently been associated with clearance of acute HCV infection (42). Several studies have confirmed the immunodominant role of this epitope in human leucocyte antigen-A2+ patients and provided evidence for a role of CTLs to this epitope in the control of infection (41, 42). To escape CTL recognition and promote viral persistence, the virus can mutate within immunodominant epitopes. Although such mutations can completely abolish recognition by CTLs, variability within the T-cell receptor–binding domain is accompanied by severe viral variant fitness deficits that limit their selection in nature (42). The two epistatic NS3-Q80K interaction sites A91 and S174, however, are located outside the NS3$_{1073–1081}$ epitope (Fig. 7). Accordingly, the corresponding epistatic substitutions, in our cell culture model, only had a minor impact on the replicative fitness of NS3-Q80K variants, although they showed significant potential to rescue protease fold instability including the NS3$_{1073–1081}$ epitope. Interestingly, fold-stabilizing mutations can mediate escape from antiviral immunity, as previously reported (43). The underlying mechanism depends on the conformational stability of the target peptides. Protein unfolding is rendering the peptide backbone accessible to proteolytic degradation, thereby determining the availability of appropriate antigenic peptides that can trigger CTL responses (44, 45). To our knowledge,
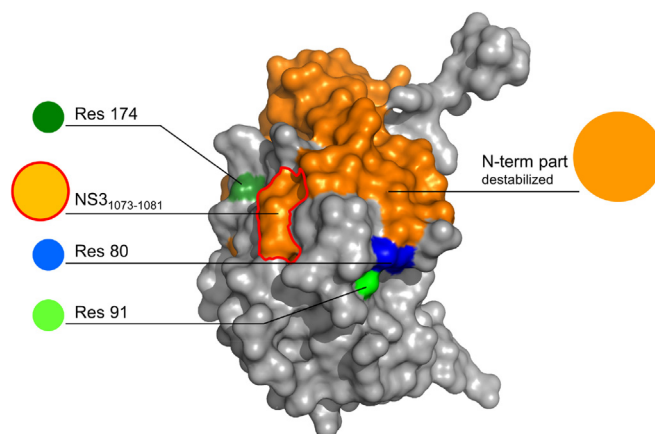
Figure 7. Localization of NS3-Q80K, epistatic interaction sites, and the T-cell epitope NS3$_{1073–1081}$. Surface representation showing the destabilized N-terminal part of the NS3-Q80K protease structure (*orange*; residues 7–78, based on RMSF from Fig. 5B) (from Protein Data Bank ID: 3M5O (62)) comprising the immunodominant T-cell epitope NS3$_{1073–1081}$ (*red circled*; corresponding to protease residues 47–55); residue 80 (*blue*); and epistatic interaction site 91 (*light green*) and 174 (*dark green*). RMSF, root mean square fluctuation.

thermal stability so far is unknown as an immune escape mechanism in HCV. A prominent example, however, is reported from the study by Klein *et al.* (43) in influenza A virus infection. They found that pandemic H1N1 influenza A strains split into two lineages that had different relative hemagglutinin protein stabilities with later variants descended from the higher-stability lineage. Analysis of the mutations that constrained the influenza A virus evolution associated with selective sweep showed early appearance of highly stabilizing mutations in the hemagglutinin protein (43). Similarly, CTL immune escape in HCV could explain the stable transmission of NS3-Q80K between hosts. Such a mechanism would also support the replicative fitness of NS3-Q80K variants in the presence of DAAs and could contribute to a general drug resistance phenotype (46) (Fig. 1).

In summary, our results suggest that protease protein stability plays a significant role in the persistence of NS3-Q80K variants. Analysis of the amino acid substitutions associated with NS3-Q80K found that they were characterized by rescue in protein stability that trade-off against protease function. The data lead us to conclude that epistatic interactions in NS3-Q80K contribute to the high prevalence of this polymorphism by mechanisms not directly related to RNA replication. By stabilizing part of the protein structure comprising a prime epitope for CTL-mediated immune responses, epistatic interactions potentially protect NS3-Q80K variants from detection by T cells.

Experimental procedures
Study populations, sequence retrieval, and analysis

Sequence analysis for secondary substitutions in the NS3-4A protease of genotype 1a with NS3-Q80K was based on sequence information deposited in the Frankfurt Resistance Database from serum of HCV patients with DAA failure sent from European study sites. At least 73% of these patients

received first-line DAA therapy. Twenty-four patients (58.5%) were treated with paritaprevir/ritonavir—ombitasvir and dasabuvir, the 3D regimen (PrOD), eight patients (19.5%) with simeprevir/sofosbuvir, and nine patients (22%) with grazoprevir/elbasvir (Table S1). The standard treatment duration was 12 weeks. The majority of patients experienced relapse (81%), whereas a nonresponse or breakthrough was detected in only approximately 10% of individuals (the remaining 9% of patients discontinued treatment, or the type of failure was not specified). Sequence information on DAA-naïve patients of genotype 1a harboring NS3-Q80K was retrieved from the European HCV database (24). Details are provided in the supporting information.

Protein biochemistry

The protein construct used for thermal shift and protease activity assays comprises the N-terminal protease domain of NS3 linked to the NS4A peptide of the viral polyprotein. For transformation and recombinant NS3-4A protein expression (27), the pET-15b vector incorporating the single-chain proteolytic domain ($NS4A_{21-32}$–GSGS-$NS3_{3-181}$) of the NS3-4A protease WT or mutants of NS3-4A was transformed into *Escherichia coli* Rosetta pLysS (DE3) competent cells (Novagen) according to the manufacturer's protocol. Recombinant protein production and purification was carried out as described previously (32).

Thermal shift assay

Stability of the protein fold of expressed WT and mutant NS3-4A was measured by fluorescence-based thermal shift using Sypro Orange, a temperature-stable fluorophore that exhibits enhanced fluorescence upon interacting with unfolded proteins (25, 32). The impact of NS3-Q80K and epistatic amino acid substitutions at protease residue A91 and/or S174 (in double and triple mutant patterns) on protease protein unfolding patterns was characterized by fluorescence emission curves. T_m from purified protease WT and mutants is determined by fitting the sigmoidal melt curve to the Boltzmann equation.

Protease activity assay

The enzymatic activity and Michaelis–Menten kinetics of the NS3-4A protease WT and mutant proteases were measured by a fluorescence-based enzyme activity assay (26, 32). To measure peptide cleavage kinetics, a viral polyprotein substrate derived from the NS4A/4B cleavage site was purchased from Anaspec (SensoLyte 520 HCV; Anaspec).

Computational studies

The models for the NS3-Q80K protease mutants were created with FoldX (http://foldxsuite.crg.eu) (47) based on the structure of the NS3 protease–helicase complex from genotype 1b (48) (Protein Data Bank ID: 1CU1), which in the protease domain is 92% identical to the genotype 1a enzyme. A structure of the protein monomer was extracted for further analysis. The structures were optimized, and MD simulations were performed with GROMACS, version 5.1.4 (https://www.gromacs.org) (49), using Amber99SB*-ILDN force field (50, 51). The molecules were centered in a cubic box with a 1.5 nm buffer, under periodic boundary conditions, and the systems were explicitly solvated with TIP3P water molecules. Counterions (52) ClJ and NaJ for Amber99SB*-ILDN force field simulations were added when necessary to neutralize the overall charge (0.15 mol/l concentration). Energy minimization for each of the molecular systems was performed using the steepest descent algorithm. A maximum of 50,000 steps was performed until a maximum force of 1000.0 kJ mol^{-1} nm^{-1} was achieved. Following the energy minimization, each of the molecular systems was subject to two consecutive steps of equilibration procedure. At first, each system was maintained at a temperature of 310 K during the canonical (NVT; particle number N, volume V and temperature T) ensemble for 100 ps with a time step of 2 fs, followed by a 100 ps simulation in the isothermal–isobaric (NPT; particle number N, pressure P and temperature T) ensemble with a time step of 2 fs maintaining the pressure at 1 bar to equilibrate the system.

After the NPT ensemble simulations, we performed 28 production simulations with 100 ns each and four replicas for each of the NS3-Q80K models (WT protease and NS3-Q80K mutants). The coordinates were recorded at every 100 ps. The temperatures of solute and solvent were coupled separately to the velocity rescale thermostat (modified Berendsen thermostat) (53) at 310 K with a relaxation time of 0.1 ps. The pressure was maintained at 1 atm by isotropic coordinate scaling with a relaxation time of 5 ps using Parrinello–Rahman barostat (54). A time step of 2 fs was used to integrate the equations of motion based on the leap-frog algorithm (55). Lennard–Jones interactions were set to a cutoff of 1.4 nm, and the Particle Mesh Ewald (56) method was used to treat long-range electrostatic interactions. All bonds were constrained using P-LINCS algorithm (57).

From the resulting trajectories, the first 10 ns were removed as part of the equilibration process, and the rest of the 90 ns from each of the simulated protease models were concatenated with its own replicas. From each concatenated trajectory, 181 frames were extracted at equal time steps, and residue interaction networks were created using RINerator (https://rinalyzer.de/rinerator.php) (58). The number of contacts for each position of interest was investigated with custom scripts in Python.

Infectious HCV cell culture

Details of the cells and reagents used for characterization of viral variant replicative fitness in H77S cell culture are reported elsewhere (23).

Plasmids

The pH77S.3 and pH77S.3/(GLuc)2A are molecular clones of the genotype 1a H77 strain of HCV (59). The H77 strain harbors the sequence pattern Q80-S91-N174, which differs from the consensus sequence pattern Q80-A91-S174 in HCV

genotype 1a–infected individuals ([60]). Epistatic amino acid substitutions, A91S/T and S174N, were created by site-directed mutagenesis in plasmids harboring NS3-Q80K. pH77S.3/GLuc2A RNA produces secreted GLuc reporter protein.

Virus fitness

Genome-length RNA was transcribed from the mutated pH77S.3 and pH77S.3/GLuc2A plasmids *in vitro*, and studies to assess viral fitness were carried out as described previously ([23]).

Statistical analyses

Comparisons of measurements from thermal shift and protease activity assay were analyzed by two-sided t tests with a level of significance set at a minimum of $p = 0.05$. Two-way ANOVA was used to compare the magnitude of changes in RNA replication after normalization to WT controls. All statistical tests were carried out using Prism 6.0 software (GraphPad Software, Inc).

Data availability

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

Abbreviations—The abbreviations used are: CTL, cytotoxic T lymphocyte; DAA, direct-acting antiviral; FC, fold change; GLuc, Gaussia luciferase; HCV, hepatitis C virus; MD, molecular dynamics; PI, protease inhibitor.

References

1. McCloskey, R. M., Liang, R. H., Joy, J. B., Krajden, M., Montaner, J. S., Harrigan, P. R., and Poon, A. F. (2015) Global origin and transmission of hepatitis C virus nonstructural protein 3 Q80K polymorphism. *J. Infect. Dis.* **211**, 1288–1295
2. Götte, M., Berghuis, A. A., Matlashewski, G., Sheppard, D., and Wainberg, M. A. (2017) *Handbook of Antimicrobial Resistance*, 1st Ed, Springer-Verlag, New York, NY
3. Pham, L. V., Jensen, S. B., Fahnøe, U., Pedersen, M. S., Tang, Q., Ghanem, L., Ramirez, S., Humes, D., Serre, S. B. N., Schønning, K., Bukh, J., and Gottwein, J. M. (2019) HCV genotype 1-6 NS3 residue 80 substitutions impact protease inhibitor activity and promote viral escape. *J. Hepatol.* **70**, 388–397
4. Harrington, P. R., Komatsu, T. E., Deming, D. J., Donaldson, E. F., O'Rear, J. J., and Naeger, L. K. (2018) Impact of hepatitis C virus polymorphisms on direct-acting antiviral treatment efficacy: Regulatory analyses and perspectives. *Hepatology* **67**, 2430–2448
5. Sarrazin, C., Lathouwers, E., Peeters, M., Daems, B., Buelens, A., Witek, J., Wyckmans, Y., Fevery, B., Verbinnen, T., Ghys, A., Schlag, M., Baldini, A., De Meyer, S., and Lenz, O. (2015) Prevalence of the hepatitis C virus NS3 polymorphism Q80K in genotype 1 patients in the European region. *Antiviral Res.* **116**, 10–16
6. Verbinnen, T., Fevery, B., Vijgen, L., Jacobs, T., De Meyer, S., and Lenz, O. (2015) *In vitro* activity of simeprevir against hepatitis C virus genotype 1 clinical isolates and its correlation with NS3 sequence and site-directed mutants. *Antimicrob. Agents Chemother.* **59**, 7548–7557
7. Lenz, O., Verbinnen, T., Fevery, B., Tambuyzer, L., Vijgen, L., Peeters, M., Buelens, A., Ceulemans, H., Beumont, M., Picchio, G., and De Meyer, S. (2015) Virology analyses of HCV isolates from genotype 1-infected patients treated with simeprevir plus peginterferon/ribavirin in phase IIb/III studies. *J. Hepatol.* **62**, 1008–1014
8. Dietz, J., Susser, S., Vermehren, J., Peiffer, K. H., Grammatikos, G., Berger, A., Ferenci, P., Buti, M., Müllhaupt, B., Hunyady, B., Hinrichsen, H., Mauss, S., Petersen, J., Buggisch, P., Felten, G., *et al.* (2018) Patterns of resistance-associated substitutions in patients with chronic HCV infection following treatment with direct-acting antivirals. *Gastroenterology* **154**, 976–988.e4
9. Zerihun, M. B., and Schug, A. (2017) Biomolecular coevolution and its applications: Going from structure prediction toward signaling, epistasis, and function. *Biochem. Soc. Trans.* **45**, 1253–1261
10. VanderWeele, T. J. (2010) Epistatic interactions. *Stat. Appl. Genet. Mol. Biol.* **9**. Article 1
11. Miton, C. M., Chen, J. Z., Ost, K., Anderson, D. W., and Tokuriki, N. (2020) Statistical analysis of mutational epistasis to reveal intramolecular interaction networks in proteins. *Methods Enzymol.* **643**, 243–280
12. Frick, D. N. (2007) The hepatitis C virus NS3 protein: A model RNA helicase and potential drug target. *Curr. Issues Mol. Biol.* **9**, 1–20
13. 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**, 482–496
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 protein 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 adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat. Immunol.* **6**, 981–988

18. 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 macro-cyclic NS3-4A inhibitors subvert IFN-β induction by efficient MAVS cleavage. *J. Hepatol.* **62**, 779–784

19. 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

20. 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

21. 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

22. Sampath, A., and Padmanabhan, R. (2009) Molecular targets for flavivirus drug discovery. *Antiviral Res.* **81**, 6–15

23. 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

24. Combet, C., Garnier, N., Charavay, C., Grando, D., Crisan, D., Lopez, J., Dehne-Garcia, A., Geourjon, C., Bettler, E., Hulo, C., Le Mercier, P., Bartenschlager, R., Diepolder, H., Moradpour, D., Pawlotsky, J. M., *et al.* (2007) euHCVdb: The European hepatitis C virus database. *Nucleic Acids Res.* **35**, D363–D366

25. 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

26. Copeland, R. A. (2013) *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*, 2nd Ed, Wiley-Interscience, Hoboken, NJ

27. 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

28. Bigman, L. S., and Levy, Y. (2018) Stability effects of protein mutations: The role of long-range contacts. *J. Phys. Chem. B* **122**, 11450–11459

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

30. 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

31. 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

32. Dultz, G., Shimakami, T., Schneider, M., Murai, K., Yamane, D., Marion, A., Zeitler, T. M., Stross, C., Grimm, C., Richter, R. M., Bäumer, K., Yi, M., Biondi, R. M., Zeuzem, S., Tampé, R., *et al.* (2020) Extended interaction networks with HCV protease NS3-4A substrates explain the lack of adaptive capability against protease inhibitors. *J. Biol. Chem.* **295**, 13862–13874

33. Yi, M., Tong, X., Skelton, A., Chase, R., Chen, T., Prongay, A., Bogen, S. L., Saksena, A. K., Njoroge, F. G., Veselenak, R. L., Pyles, R. B., Bourne, N., Malcolm, B. A., and Lemon, S. M. (2006) Mutations conferring resistance to SCH6, a novel hepatitis C virus NS3/4A protease inhibitor. Reduced RNA replication fitness and partial rescue by second-site mutations. *J. Biol. Chem.* **281**, 8205–8215

34. Welsch, C., Shimakami, T., Hartmann, C., Yang, Y., Domingues, F. S., Lengauer, T., Zeuzem, S., and Lemon, S. M. (2012) Peptidomimetic escape mechanisms arise via genetic diversity in the ligand-binding site of the hepatitis C virus NS3/4A serine protease. *Gastroenterology* **142**, 654–663

35. 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

36. Zeldovich, K. B., Chen, P., and Shakhnovich, E. I. (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16152–16157

37. Peres-da-Silva, A., Antunes, D., Quintanilha Torres, A. L., Caffarena, E. R., and Lampe, E. (2019) Effects of the Q80K polymorphism on the physicochemical properties of hepatitis C virus subtype 1a NS3 protease. *Viruses* **11**, 691

38. Raab, M., Sanhaji, M., Pietsch, L., Béquignon, I., Herbrand, A. K., Süß, E., Gande, S. L., Caspar, B., Kudlinzki, D., Saxena, K., Sreeramulu, S., Schwalbe, H., Strebhardt, K., and Biondi, R. M. (2018) Modulation of the allosteric communication between the polo-box domain and the catalytic domain in Plk1 by small compounds. *ACS Chem. Biol.* **13**, 1921–1931

39. Bigman, L. S., and Levy, Y. (2020) Proteins: Molecules defined by their trade-offs. *Curr. Opin. Struct. Biol.* **60**, 50–56

40. Murai, K., Shimakami, T., Welsch, C., Shirasaki, T., Liu, F., Kitabayashi, J., Tanaka, S., Funaki, M., Omura, H., Nishikawa, T., Sumiyadorj, A., Honda, M., and Kaneko, S. (2018) Unexpected replication boost by simeprevir for simeprevir-resistant variants in genotype 1a hepatitis C virus. *Antimicrob. Agents Chemother.* **62**, e02601–e02617

41. Engler, O. B., Schwendener, R. A., Dai, W. J., Wölk, B., Pichler, W., Moradpour, D., Brunner, T., and Cerny, A. (2004) A liposomal peptide vaccine inducing CD8+ T cells in HLA-A2.1 transgenic mice, which recognise human cells encoding hepatitis C virus (HCV) proteins. *Vaccine* **23**, 58–68

42. Söderholm, J., Ahlén, G., Kaul, A., Frelin, L., Alheim, M., Barnfield, C., Liljeström, P., Weiland, O., Milich, D. R., Bartenschlager, R., and Sällberg, M. (2006) Relation between viral fitness and immune escape within the hepatitis C virus protease. *Gut* **55**, 266–274

43. Klein, E. Y., Blumenkrantz, D., Serohijos, A., Shakhnovich, E., Choi, J. M., Rodrigues, J. V., Smith, B. D., Lane, A. P., Feldman, A., and Pekosz, A. (2018) Stability of the influenza virus hemagglutinin protein correlates with evolutionary dynamics. *mSphere* **3**, e00554-17

44. Scheiblhofer, S., Laimer, J., Machado, Y., Weiss, R., and Thalhamer, J. (2017) Influence of protein fold stability on immunogenicity and its implications for vaccine design. *Expert Rev. Vaccines* **16**, 479–489

45. Schmidt, J., Iversen, A. K., Tenzer, S., Gostick, E., Price, D. A., Lohmann, V., Distler, U., Bowness, P., Schild, H., Blum, H. E., Klenerman, P., Neumann-Haefelin, C., and Thimme, R. (2012) Rapid antigen processing and presentation of a protective and immunodominant HLA-B*27-restricted hepatitis C virus-specific CD8+ T-cell epitope. *PLoS Pathog.* **8**, e1003042

46. 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

47. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: An online force field. *Nucleic Acids Res.* **33**, W382–W388

48. Yao, N., Reichert, P., Taremi, S. S., Prosise, W. W., and Weber, P. C. (1999) Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* **7**, 1353–1363

49. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005) GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718

50. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725

51. Martín-García, F., Papaleo, E., Gomez-Puertas, P., Boomsma, W., and Lindorff-Larsen, K. (2015) Comparing molecular dynamics force fields in the essential subspace. *PLoS One* **10**, e0121114

52. Joung, I. S., and Cheatham, T. E. (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041

53. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101

54. Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190

55. Van Gunsteren, W. F., and Berendsen, H. J. C. (1988) A leap-frog algorithm for stochastic dynamics. *Mol. Simul.* **1**, 173–185

56. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092

57. Hess, B. (2008) P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122

58. Doncheva, N. T., Klein, K., Domingues, F. S., and Albrecht, M. (2011) Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci.* **36**, 179–182

59. 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

60. Vallet, S., Gouriou, S., Nousbaum, J. B., Legrand-Quillien, M. C., Goudeau, A., and Picard, B. (2005) Genetic heterogeneity of the NS3 protease gene in hepatitis C virus genotype 1 from untreated infected patients. *J. Med. Virol.* **75**, 528–537

61. Gouy, M., Guindon, S., and Gascuel, O. (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224

62. 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