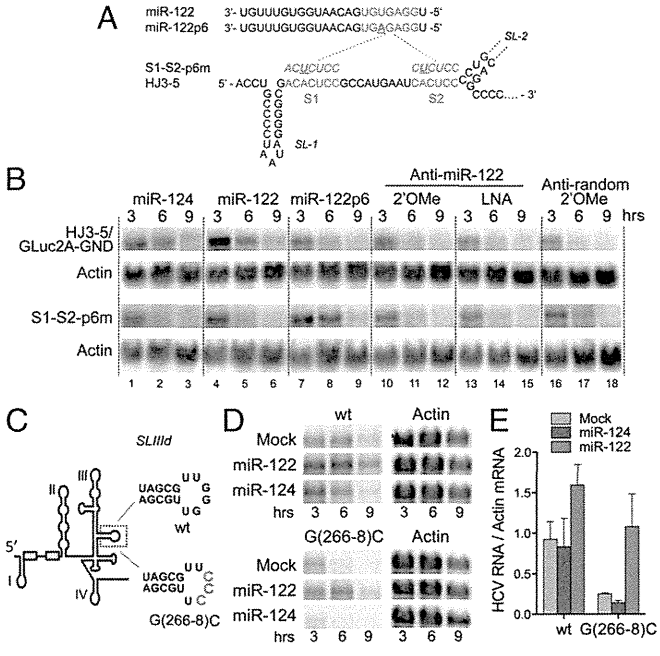

Fig. 3. Stabilization of HCV RNA requires miR-122 binding and is independent of translation. (*A*) *Upper*: miR-122 and mutant miR-122p6 guide-strand sequences. *Lower*: 5′ terminal sequence of HCV (HJ3-5/GLuc2A virus), with S1 and S2 binding sites shown in red. Point mutations (underlined) in the related S1-S2-p6m-GND mutant (11) are shown above. SL-1 and SL-2 are putative stem-loop structures in the 5′ UTR. (*B*) Northern blots of HJ3-5/GLuc2A-GND and the related S1-S2-p6m-GND mutant RNA after transfection into Huh-7.5 cells with RNA oligoribonucleotides as in Fig. 1*B*. (*C*) Putative secondary structure of the HCV 5′ UTR, showing the location of stem-loop IIId and the G(266-8)C IRES mutation that ablates translation. (*D*) Northern blot showing HCV RNA abundance in MEFs transfected with HCV RNA (H77S/GLuc2A-AAG) containing (*Upper*) the WT 5′ UTR vs. (*Lower*) the translationally inactive G(266-8)C mutant, and supplemented with the indicated miRNAs. (*E*) PhosphorImager quantitation of HCV RNA in Northern blots of 6-h cell lysates from two independent experiments carried out as shown in *D*.

inactive mutant was significantly slowed when the cells were supplemented with miR-122 (Fig. 3 *D* and *E*). These experiments were carried out in murine embryonic fibroblasts (MEFs) that do not express detectable endogenous miR-122. Similar results were obtained both in MEFs and hepatoma cells with another IRES mutant, G(267)C, that retains affinity for the 40S subunit but is also translationally dead (11, 16) (Fig. S3 *B–E*). Collectively these results show that miR-122 does not stabilize HCV RNA by promoting its translation or enhancing its association with ribosomes. Active engagement in translation and miR-122 promote stability of the RNA genome independently, and possibly additively.

Randall et al. (21) reported previously that RNAi-mediated depletion of any of the Ago proteins (particularly Ago4) or Dicer inhibited the ability of HCV to infect cells. In contrast, we found that only Ago2 depletion (Fig. S4*A*) inhibited HCV RNA replication in hepatoma cells with previously established infection (Fig. 4*A*). Ago2 depletion had no effect on cell growth (Fig. S4*B*) but reduced both HCV RNA abundance and protein expression (Fig. 4*A*). This was not observed with depletion of other Ago proteins or Dicer. Although we are uncertain why our results differ from those of Randall et al. (21), they suggest that Ago2 plays a special role in HCV replication. Two recent studies are consistent with this: they show Ago2 to be required for miR-122 to promote HCV genome amplification (15, 22). However, a third study found no impairment in miR-122 modulation of HCV in Ago2-depleted cells (14).

Consistent with the stabilization of HCV RNA requiring interaction with an miR-122–associated RISC complex, we observed no stabilizing or translation-enhancing effect in hepatoma cells transfected with single-stranded miR-122 (guide strand only, rather than the duplex miRNA transfected in previous experiments) (Fig. S4*C*). To determine whether miR-122 binds as a complex with Ago2, we lysed MEFs shortly after electroporation with HCV RNA together with duplex miR-122 and used RT-PCR to interrogate Ago2 immunoprecipitates for the presence of viral RNA. HCV RNA with WT S1 and S2 sequence coimmunoprecipitated with Ago2 (Fig. 4*B*). However, RNA with point mutations in S1 and S2 that ablate miR-122 binding (Fig. 3*A*) coimmunoprecipitated with Ago2 only when cells were supplemented with the complementary miR-122 mutant, miR-122p6 (Fig. 4*B*). Transfected HCV RNA was also enriched in anti-Flag immunoprecipitates from hepatoma cells ectopically expressing Flag-Ago2 vs. Flag-Ago1 (Fig. S5). Thus, miR-122 binds the HCV 5′ UTR in association with Ago2. Lesser amounts of Ago1 may also be present in the complex.

To determine whether Ago2 plays a functional role in stabilizing HCV RNA, we compared the ability of miR-122 to slow the decay of the RNA when electroporated with it into WT or Ago2-deficient ($Ago2^{-/-}$) MEFs (23). This revealed a striking dependence on Ago2, because miR-122 had no effect in $Ago2^{-/-}$ MEFs, whereas it significantly stabilized the viral RNA and enhanced viral translation in matched WT cells (Fig. 4 *C* and *D*). Ectopic expression of human Flag-Ago2 in the $Ago2^{-/-}$ MEFs restored the ability of miR-122 to positively regulate protein expression from HCV RNA, further confirming the requirement for Ago2 (Fig. S6). To ascertain whether it is possible for other Ago proteins to functionally substitute for Ago2 in stabilizing HCV RNA, we also overexpressed human Flag-Ago1 in the $Ago2^{-/-}$ MEFs. This only partially rescued the ability of miR-122 to promote HCV protein expression (Fig. S6*B*), although immunoblots with anti-Flag antibody indicated that Flag-Ago1 was overexpressed at almost threefold the abundance of Flag-Ago2 and well above physiologically relevant levels (Fig. S6*A*). Thus, Ago2 is not unique in its ability to support miR-122 enhancement of HCV protein expression, but it seems to do this more efficiently than Ago1. This is consistent with the greater enrichment of HCV RNA we observed in Flag-Ago2 compared with Flag-Ago1 immunoprecipitates (Fig. S5). Additional experiments confirmed that Ago2 is the dominant Ago protein involved in miR-122 stabilization of HCV RNA. RNAi-mediated depletion of Ago2 reduced the ability of miR-122 to stabilize HCV RNA and promote its translation in HeLa cells (Fig. S7 *A–D*). Importantly, the magnitude of this effect mirrored the reduction in miR-122–mediated suppression of a reporter mRNA containing the HCV miR-122 binding sites within its 3′ UTR (9) (Fig. S7*E*).

These results suggest a functional interaction of Ago2 with RNA decay machinery. Because HCV RNA lacks a 5′ m^7G cap, one possibility is that miR-122 recruits an Ago2 RISC complex to the 5′ end of viral RNA that protects it from 5′ exonuclease activity. If so, this protective action should be rendered redundant by providing the RNA with a 5′ cap. To test this, we synthesized HCV RNAs with or without a 5′ nonmethylated guanosine cap analog and compared their stabilities after transfection into MEFs with or without miR-122. As anticipated, miR-122 had no effect on the rate of decay of the 5′ capped RNA, whereas it substantially stabilized the uncapped HCV RNA (Fig. 5*A*). The magnitude of protein expression (GLuc) from capped RNA was comparable to that from the uncapped RNA in cells supplemented with miR-122 and was not further increased by miR-122 (Fig. 5*B*). Thus, the 5′ cap enhanced stability and protein expression from HCV RNA, functionally substituting for miR-122 and providing strong evidence that the miR-122–Ago2 complex protects HCV RNA from 5′ exonuclease. Importantly,

translation of the capped HCV RNA was IRES-dependent, because the nonmethylated 5′ cap lacked the ability to recruit eukaryotic initiation factors.

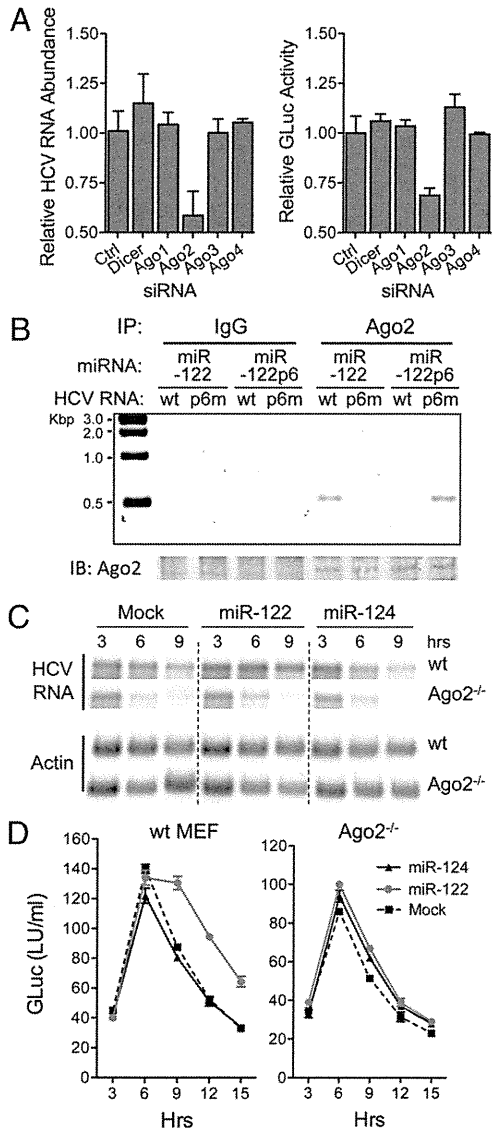

Fig. 4. Ago2 binds HCV RNA in association with miR-122 and is required for stabilization. (*A*) RNAi depletion of Ago2 impairs HCV genome amplification in persistently infected (HJ3-5/GLuc2A virus) cells. *Left*: HCV RNA abundance relative to that in cells transfected with nontargeting siRNA (Ctrl) 72 h after siRNA transfection. RNA was assayed by qRT-PCR; data are mean ± range of paired cultures. *Right*: GLuc activity secreted into media between 48 and 72 h, relative to si-Ctrl-transfected cultures (mean ± range). (*B*) miR-122 binds HCV RNA as a complex with Ago2. Lysates were prepared from WT MEFs 6 h after electroporation with HJ3-5/GND or S1-S2-p6m-GND RNA (Fig. 3*A*), mixed with miR-122 or miR-122p6, and immunoprecipitated with anti-Ago2 antibody. After extensive washing, RNA was extracted from the precipitates and subjected to a one-step HCV-specific RT-PCR (30 cycles). HCV RNA was enriched in precipitates from HJ3-5/GND–transfected cells supplemented with miR-122, or S1-S2-p6m-GND cells supplemented with miR-122p6. (*C*) Northern blots showing miR-122 does not stabilize HCV RNA in Ago2$^{-/-}$ MEFs. Cells were electroporated with HCV RNA (H77S/GLuc2A-AAG) together with miR-122, miR-124, or no miRNA (Mock), then lysed at 3-h intervals and assayed for HCV RNA abundance. miR-122 stabilized HCV RNA only in WT MEFs and was without effect in Ago2$^{-/-}$ cells. (*D*) GLuc activity in supernatant fluids from MEFs cotransfected with HCV RNA and the indicated miRNA (mean ± range). Data shown are representative of two or more independent experiments.

Discussion

Our results reveal a unique mechanism by which a miRNA in association with Ago2 regulates the expression of its target RNA, in this case the HCV genome. Although miR-10a enhances translation by binding the 5′ UTR of ribosomal protein mRNAs containing 5′ terminal oligopyrimidine (TOP) motifs (24), miRNAs have not been recognized to slow decay or up-regulate abundance of their RNA targets. miR-122 does this by recruiting an Ago2 RISC-like complex to the 5′ end of the HCV RNA genome. As evidenced by the IRES mutants (Fig. 3*D* and Fig. S3*D*), the stabilizing action of miR-122 does not require the target HCV RNA to be capable of translation and thus does not result from increased ribosomal loading. Stabilization also does not require the RNA to be replication competent (Fig. 1 *B* and *C*).

Because a nonmethylated 5′ cap analog functionally substitutes for miR-122 (Fig. 5), the RISC-like complex recruited by miR-122 is likely to act by protecting the RNA from 5′ exonuclease. Whether this occurs simply as a result of physically masking the 5′ end of the viral RNA from 5′ exonuclease attack, or whether Ago2 plays a more complex role by influencing the association of HCV RNA with P bodies (25), sites of mRNA degradation and storage, remains to be determined. Binding of a RISC-like complex could also limit recognition of the 5′ triphosphate of HCV RNA by retinoic acid-inducible gene I (RIG-I), a ubiquitous innate immune pathogen recognition receptor (26) that is capable of inducing interferons and interferon-stimulated genes, including RNase L, an endonuclease, and ISG20, a 3′-5′ exonuclease (27). However, subversion of RIG-I–dependent viral RNA degradation played no role in the stabilization of HCV RNA in our experiments because the Huh-7.5 cells used are deficient in RIG-I signaling (28).

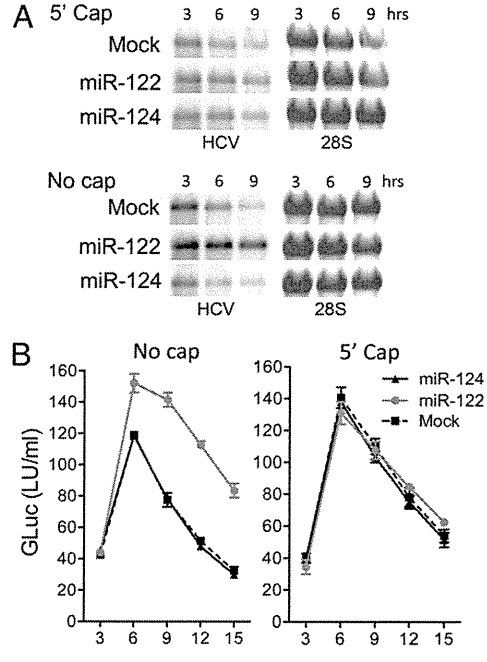

Fig. 5. A nonmethylated 5′ guanosine cap functionally substitutes for miR-122 in stabilizing HCV RNA. (*A*) Northern blots of HCV RNA in lysates of MEFs after electroporation with HCV RNA (H77S/GLuc2A-AAG) (Fig. 1*A*) synthesized with or without a nonmethlyated G[5′]ppp[5′]G-RNA cap. HCV RNAs were transfected together with miR-122, miR-124, or no miRNA (Mock). 28S rRNA is shown as a loading control. (*B*) GLuc activity in supernatant fluids of MEFs transfected with capped or uncapped HCV RNAs (mean ± range of two replicate cultures). Data shown are representative of two or more independent experiments.

The mechanism by which miR-122 stabilizes HCV RNA is distinct from the up-regulation of translation by miRNAs that target AU-rich elements within the 3′ UTRs of some mRNAs, because the latter occurs only in quiescent cells arrested at the G_0 phase of the cell cycle (29, 30). Moreover, miRNA-mediated stimulation of translation in quiescent cells has not been linked to stabilization of the target mRNA, as we show here for miR-122. Still to be determined is whether the protein composition of the RISC-like complex that is recruited to the 5′ UTR by miR-122 differs significantly from miRNA-induced RISC complexes involved in translational repression.

The stabilization of HCV RNA by miR-122 is likely to be responsible for the miR-122–induced enhancement of HCV translation reported previously (10, 11). Mutations in the viral RNA that prevent binding of miR-122 also ablate replication of infectious virus (11) and prevent stabilization of the genome by miR-122 (Fig. 3). Depletion of Ago2 has similar effects on replication, translation, and RNA stability (15, 22) (Fig. 4). This makes it difficult to distinguish between these consequences of miR-122 binding to the 5′ UTR or to determine the primary role played by miR-122 as a host factor for viral replication. However, stabilization of the genome is likely to be a key factor in the promotion of viral replication by miR-122. A recombinant HCV in which the U3 RNA sequence was inserted in lieu of the S1 binding site in the 5′ UTR was found recently to be less dependent upon miR-122 for replication (17), possibly because the U3 sequence stabilized the RNA, much as a 5′ cap did (Fig. 5). Nonetheless, it would not be surprising to find that miR-122 has other functions in the viral life cycle in addition to its role in stabilizing the viral RNA genome, perhaps in the initiation of viral RNA synthesis.

Methods

Viral RNA Stability in Transfected Cells. RNA was transcribed in vitro (11) from pH77S/GLuc2A-AAG, which contains the complete genotype 1a HCV sequence with GLuc2A placed in-frame within the polyprotein-coding region (31) and a lethal GDD to AAG mutation in NS5B. Where indicated, a non-methylated 5′ guanosine cap was added using the ScriptCap m7G Capping System (Epicentre Biotechnologies). Viral RNA (20 μg) and miRNA duplexes (11) or antisense oligoribonucleotides (1 μM) were mixed with 1×10^7 Huh-7.5 cells in a 4-mm cuvettete and pulsed once at 250 V, 950 μF, and 100 Ω in a Gene Pulser Xcell Total System (Bio-Rad). HeLa cells were electroporated at 300 V, 500 μF, and ∞ Ω, and MEFs at 400 V, 250 μF, and ∞ Ω. Cells were harvested at intervals and supernatant fluids assayed for GLuc activity (31) and HCV RNA abundance in cell lysates assessed by Northern blotting (11). Polyadenylated reporter RNAs encoding firefly or *Cypridina* luciferase were cotransfected to monitor transfection efficiency (11).

HCV-Infected Cells. Synthetic HJ3-5 or HJ3-5/GLuc2A RNA (31) was transfected into 1×10^7 FT3-7 cells, which were passaged until >90% positive for core antigen in an immunofluorescence assay (11). siRNA pools targeting Ago1-4 or Dicer and control siRNA pools (Dharmacon) were transfected using siLentfect Lipid Reagent (Bio-Rad). miRNA duplexes or single-stranded oligoribonucleotides (50 nM) were transfected using Lipofectamine 2000 (Invitrogen).

Additional methods and associated references can be found in *SI Methods*.

ACKNOWLEDGMENTS. We thank Lucinda Hensley for expert technical assistance; William F. Marzluff and Scott M. Hammond for helpful discussions; Alexander Tarakhovsky and Angela Santana, Rockefeller University, for mouse embryonic fibroblasts; Charles Rice, Rockefeller University, for Huh-7.5 cells; Thorleif Møller, Mirrx Therapeutics, for psi-Check2/Luc-3′ HCV DNA; and Angela Lam and Phil Furman, Pharmasset, Inc., for PSI-6130. This study was supported in part by the University Cancer Research Fund and National Institutes of Health Grants RO1-AI095690, U19-AI040035, P20-CA150343 (to S.M.L.), and AI042189 (to D.J.B.). R.K.J. was supported by a James W. McLaughlin Fellowship.

1. Fabian MR, Sonenberg N, Filipowicz W (2010) Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* 79:351–379.
2. Balagopal V, Parker R (2009) Polysomes, P bodies and stress granules: States and fates of eukaryotic mRNAs. *Curr Opin Cell Biol* 21:403–408.
3. Djuranovic S, Nahvi A, Green R (2011) A parsimonious model for gene regulation by miRNAs. *Science* 331:550–553.
4. Guo H, Ingolia NT, Weissman JS, Bartel DP (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466:835–840.
5. Chang J, et al. (2004) miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1. *RNA Biol* 1:106–113.
6. Lanford RE, et al. (2010) Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* 327:198–201.
7. Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P (2005) Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* 309:1577–1581.
8. Chisari FV (2005) Unscrambling hepatitis C virus-host interactions. *Nature* 436: 930–932.
9. Jopling CL, Schütz S, Sarnow P (2008) Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome. *Cell Host Microbe* 4:77–85.
10. Henke JI, et al. (2008) microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* 27:3300–3310.
11. Jangra RK, Yi M, Lemon SM (2010) miR-122 regulation of hepatitis C virus translation and infectious virus production. *J Virol* 84:6615–6625.
12. Norman KL, Sarnow P (2010) Modulation of hepatitis C virus RNA abundance and the isoprenoid biosynthesis pathway by microRNA miR-122 involves distinct mechanisms. *J Virol* 84:666–670.
13. Villanueva RA, et al. (2010) miR-122 does not modulate the elongation phase of hepatitis C virus RNA synthesis in isolated replicase complexes. *Antiviral Res* 88: 119–123.
14. Machlin ES, Sarnow P, Sagan SM (2011) Masking the 5′ terminal nucleotides of the hepatitis C virus genome by an unconventional microRNA-target RNA complex. *Proc Natl Acad Sci USA* 108:3193–3198.
15. Roberts AP, Lewis AP, Jopling CL (2011) miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Res* 39:7716–7729.
16. Kieft JS, Zhou K, Jubin R, Doudna JA (2001) Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* 7:194–206.
17. Li YP, Gottwein JM, Scheel TK, Jensen TB, Bukh J (2011) MicroRNA-122 antagonism against hepatitis C virus genotypes 1-6 and reduced efficacy by host RNA insertion or mutations in the HCV 5′ UTR. *Proc Natl Acad Sci USA* 108:4991–4996.
18. Ma H, et al. (2007) Characterization of the metabolic activation of hepatitis C virus nucleoside inhibitor beta-D-2′-Deoxy-2′-fluoro-2′-C-methylcytidine (PSI-6130) and identification of a novel active 5′-triphosphate species. *J Biol Chem* 282:29812–29820.
19. Lyons T, Murray KE, Roberts AW, Barton DJ (2001) Poliovirus 5′-terminal cloverleaf RNA is required in cis for VPg uridylylation and the initiation of negative-strand RNA synthesis. *J Virol* 75:10696–10708.
20. Murray KE, Roberts AW, Barton DJ (2001) Poly(rC) binding proteins mediate poliovirus mRNA stability. *RNA* 7:1126–1141.
21. Randall G, et al. (2007) Cellular cofactors affecting hepatitis C virus infection and replication. *Proc Natl Acad Sci USA* 104:12884–12889.
22. Wilson JA, Zhang C, Huys A, Richardson CD (2011) Human Ago2 is required for efficient miR-122 regulation of HCV RNA accumulation and translation. *J Virol* 85: 2342–2350.
23. O'Carroll D, et al. (2007) A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes Dev* 21:1999–2004.
24. Ørom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5′UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* 30:460–471.
25. Parker R, Sheth U (2007) P bodies and the control of mRNA translation and degradation. *Mol Cell* 25:635–646.
26. Hornung V, et al. (2006) 5′-Triphosphate RNA is the ligand for RIG-I. *Science* 314: 994–997.
27. Zhou Z, et al. (2011) Antiviral activities of ISG20 in positive-strand RNA virus infections. *Virology* 409:175–188.
28. Sumpter R, Jr., et al. (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.
29. Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: MicroRNAs can up-regulate translation. *Science* 318:1931–1934.
30. Vasudevan S, Tong Y, Steitz JA (2008) Cell-cycle control of microRNA-mediated translation regulation. *Cell Cycle* 7:1545–1549.
31. Shimakami T, et al. (2011) Protease inhibitor-resistant hepatitis C virus mutants with reduced fitness from impaired production of infectious virus. *Gastroenterology* 140: 667–675.

AAC
Journals.ASM.org

Ketoamide Resistance and Hepatitis C Virus Fitness in Val55 Variants of the NS3 Serine Protease

Christoph Welsch,[a,b,c] Sabine Schweizer,[d] Tetsuro Shimakami,[a,e] Francisco S. Domingues,[f] Seungtaek Kim,[a] Stanley M. Lemon,[a] and Iris Antes[d]

University of North Carolina at Chapel Hill, Division of Infectious Diseases, Department of Medicine and the Lineberger Comprehensive Cancer Center, Chapel Hill, North Carolina, USA[a]; J. W. Goethe-University Hospital, Department of Internal Medicine I, Frankfurt am Main, Germany[b]; Max Planck Institute for Informatics, Computational Biology & Applied Algorithmics, Saarbrücken, Germany[c]; Technical University Munich, Center for Integrated Protein Science (CIPS[M]), Department of Life Sciences, Freising, Germany[d]; First Department of Internal Medicine, School of Medicine, Kanazawa University, Takara-Machi, Kanazawa, Japan[e]; and Institute of Genetic Medicine, EURAC Research, Bolzano, Italy[f]

Drug-resistant viral variants are a major issue in the use of direct-acting antiviral agents in chronic hepatitis C. Ketoamides are potent inhibitors of the NS3 protease, with V55A identified as mutation associated with resistance to boceprevir. Underlying molecular mechanisms are only partially understood. We applied a comprehensive sequence analysis to characterize the natural variability at Val55 within dominant worldwide patient strains. A residue-interaction network and molecular dynamics simulation were applied to identify mechanisms for ketoamide resistance and viral fitness in Val55 variants. An infectious H77S.3 cell culture system was used for variant phenotype characterization. We measured antiviral 50% effective concentration (EC_{50}) and fold changes, as well as RNA replication and infectious virus yields from viral RNAs containing variants. Val55 was found highly conserved throughout all hepatitis C virus (HCV) genotypes. The conservative V55A and V55I variants were identified from HCV genotype 1a strains with no variants in genotype 1b. Topology measures from a residue-interaction network of the protease structure suggest a potential Val55 key role for modulation of molecular changes in the protease ligand-binding site. Molecular dynamics showed variants with constricted binding pockets and a loss of H-bonded interactions upon boceprevir binding to the variant proteases. These effects might explain low-level boceprevir resistance in the V55A variant, as well as the Val55 variant, reduced RNA replication capacity. Higher structural flexibility was found in the wild-type protease, whereas variants showed lower flexibility. Reduced structural flexibility could impact the Val55 variant's ability to adapt for NS3 domain-domain interaction and might explain the virus yield drop observed in variant strains.

Infection with hepatitis C virus (HCV) is a frequent cause of chronic hepatitis with liver cirrhosis and hepatocellular carcinoma as sequelae (1). Until recently, the standard of care (SOC) for patients with chronic hepatitis C virus infection (CHC) has consisted of a combination of pegylated interferon alpha plus ribavirin (Peg-IFN/RBV) administered for 24 to 48 weeks, depending on the HCV genotype. This is only partially effective, with about a 50% sustained viral response (SVR) in patients infected with HCV genotype 1, the most common genotype in Europe and North America (11, 21, 33, 44). The addition of a direct-acting antiviral agent (DAA) targeting the NS3/4A serine protease of HCV significantly improves the SVR rate, and two such drugs have recently been approved for clinical use by the European Medicines Agency (EMA) and the U.S. Food and Drug Administration (FDA). The ketoamide compounds boceprevir (Victrelis) and telaprevir (Incivek) were both designed to mimic the natural substrate of the NS3 protease (16, 19, 22, 23, 30). Clinical trials in treatment-naïve genotype 1-infected patients have revealed significant improvements in the kinetics of the virologic response with the addition of a DAA to the prior SOC, leading to improved SVR rates of up to 74% (16, 19, 22). Despite the progress, however, protease inhibitor (PI) resistance is a major challenge for future treatment. Resistant viral variants exist at low frequencies in untreated patients as part of the viral quasispecies population (29), reflecting the highly replicative nature of HCV infections and the error-prone character of its RNA-dependent RNA polymerase, NS5B.

The NS3 protease plays an essential role in the HCV life cycle by processing nonstructural (NS) proteins from the viral polyprotein downstream of the NS2-3 junction (24). The protease domain of NS3 comprises the amino-terminal third of the protein containing a catalytic triad, H57, D81, and S139, and an "oxyanion hole" at G137. It acts in concert with its cofactor, NS4A, which intercalates into its structure and is required for full enzymatic activity and proper subcellular localization. The carboxy-terminal two-thirds of NS3 consists of a DExD-box RNA helicase domain that is essential for productive viral infection (28). NS3 thus appears to be a critical component of the macromolecular viral RNA replicase that directs the synthesis of new viral RNAs. Genetic evidence indicates that NS3 has an additional distinct function in assembly of infectious virus particles (20, 34). Since viral RNA replication capacity and virus assembly are crucial determinants of viral fitness, mutations in NS3 that contribute to PI resistance can also profoundly influence virus fitness (42). The probability of a resistant variant emerging from the quasispecies population during treatment with a DAA is determined not only by its degree

Received 28 June 2011 **Returned for modification** 22 August 2011
Accepted 22 December 2011
Published ahead of print 17 January 2012
Address correspondence to Christoph Welsch, christophwelsch@gmx.net, or Iris Antes, antes@wzw.tum.de.

doi:10.1128/AAC.05184-11

of drug resistance, but also by its fitness. The NS3 domain interdependency might provide for novel molecular mechanisms in treatment escape from ketoamide compounds. Many mutations associated with PI resistance negatively impact the replication of genotype 1a HCV RNA in cell culture, while some have additional effects on the production of infectious virus (34). Compensatory second-site mutations may enhance the fitness of these resistant viruses (42), but current understanding about the underlying molecular mechanisms is rudimentary.

Previous studies identified the V55A variant as resistance-associated amino acid variant for the ketoamide compound boceprevir (27, 37). Furthermore, the V55A variant was found in the long-term follow up of patients up to 5.5 years upon boceprevir treatment completion. Moreover, the variant was dominant already at baseline in one of the patients before any PI exposure (36). Interestingly, Val55 is related to PI resistance, although it is buried in the NS3 protease domain structure, without direct ligand interaction (37). The V55A variant showed medium-level resistance against the ketoamide compounds boceprevir and telaprevir in enzymatic assays and had a negative impact on RNA replication in the Con1 replicon system (37). The viral fitness of the V55A variant has not been determined so far, since no data are available on its infectious virus yield.

In this article, we present a combined *in silico* and *in vitro* study on ketoamide resistance and viral fitness in Val55 variants of the NS3 protease. A comprehensive sequence analysis was performed using a public database with worldwide patient isolates to identify Val55 variants from dominant strains. Subsequently, we applied an HCV cell culture system of infection and molecular dynamics (MD) simulations to assess ketoamide resistance, viral variant fitness, and underlying molecular mechanisms in Val55 variants. The potential role of Val55 in the NS3 protease as structural and functional regulatory site is analyzed using a residue-interaction network approach.

MATERIALS AND METHODS

***In silico* sequence analysis.** We analyzed sequences of the hepatitis C virus NS3 protease deposited in the European HCV Database (euHCVdb) (7), which contains sequences of major variants from around the world. HCV genotypes and subtypes were differentiated according to a consensus proposal for a unified system of HCV genotype nomenclature (35). The sequences in this study were confirmed as HCV genotypes 1a, 1b, 2, 3, 4, 5, 6, and 7. Sequence alignments (not shown) were computed using ClustalW (6) and MUSCLE (10), with minor manual modifications, in the SEAVIEW alignment editor (12). Comprehensive sequence analysis on natural Val55 variants and correlated variants was performed in 676 HCV NS3 isolates from seven genotypes and their corresponding subtypes. We confirmed 202 strains as HCV genotype 1a and 335 strains as HCV genotype 1b. Subsequently, we use the genotype 1a H77 strain (UniProtKB reference no. P27958) for *in vitro* analysis and a genotype 1a protease from Protein Data Bank (PDB) structure 2OC8 for structure analysis. Both sequences are 98.9% identical, with only two variants between UniProtKB P27958 and PDB 2OC8, i.e., M76T and A149T, respectively.

Cells and reagents. Huh7 and Huh7.5 cells were provided from Apath and grown in Dulbecco's modified Eagle's medium (Gibco-BRL) supplemented with 10% fetal calf serum, penicillin, streptomycin, L-glutamine, and nonessential amino acids. Boceprevir (SCH503034) and telaprevir (VX-950) stock solutions were prepared in dimethyl sulfoxide (DMSO). All final dilutions contained 0.5% DMSO.

Plasmids. pH77S.3 and pH77S.3/GLuc2A are molecular clones of the genotype 1a H77 strain of HCV. Synthetic RNA transcribed from these plasmids replicates in transfected Huh7 cells and produces infectious virus. pH77S.3/GLuc2A also produces secreted *Gaussia* luciferase (GLuc) reporter protein. The V55A and V55I single-amino-acid variants and the T54S V55I double variant were created in these plasmids by site-directed mutagenesis, using AfeI and BsrGI restriction sites. Sequences of manipulated DNA segments were confirmed by DNA sequencing. pH77S/GLuc2A/AAG is a replication-defective pH77S/GLuc2A mutant. It was generated by inserting an AfeI/AscI restriction fragment containing the GLuc2A sequence between the corresponding sites of pH77S/AAG, in which the GDD motif of the polymerase, NS5B, is replaced with AAG (43).

RNA transcription and transfection. RNA was synthesized with T7 MEGAScript reagents (Ambion), after linearizing plasmids with XbaI. Following treatment with RNase-free DNase to remove template DNA, RNA was purified using the RNeasy minikit (Qiagen). RNA transfection was carried out with the Trans-IT mRNA transfection kit (Mirus) according to the manufacturer's suggested protocol. Transfection protocols were optimized for determination of antiviral susceptibility, RNA replication capacity, and infectious virus yield. Briefly, for cell-based antiviral activity assays, 150 ng RNA was transfected into 4×10^4 cells/well seeded in 48-well plates. To test the RNA replication capacity of GLuc-containing constructs, 250 ng RNA was transfected into 8×10^4 cells/well in 24-well plates. To test the capacity of RNA to produce infectious virus, 1.25 μg RNA was transfected into 6×10^5 cells/well in 6-well plates.

Luciferase activity assay. Replication of HCV genotype 1a RNA for the V55A and V55I variants was determined by measuring GLuc activity using the H77S.3/GLuc2A construct (34). Following RNA transfection, cell culture supernatant fluids were collected, and fresh medium was added at 24-h intervals. Secreted GLuc activity was measured in 25-μl aliquots of the supernatant fluids using the GLuc assay kit (New England BioLabs) according to the manufacturer's suggested protocol. The luminescent signal was measured on a Synergy 2 (Bio-Tek) Multi-Mode microplate reader.

Virus yield determination. Cells were split at a 1:1 ratio 24 h after RNA transfection, and the medium was replaced with fresh medium containing 10% HEPES every 24 h thereafter. Cell supernatant culture fluids collected at 72 h after transfection were assayed for infectious virus in a fluorescent focus virus titration assay described previously (43). Briefly, cells were seeded in 48-well plates at a density of 1×10^5 cells/well 24 h prior to inoculation with 50 to 100 μl of culture medium. Cells were maintained at 37°C in a 5% CO_2 environment and fed with 300 μl of medium at 24 h later. Following 48 h of additional incubation, cells were fixed in methanol-acetone (1:1) at room temperature for 9 min and stained with monoclonal antibody C7-50 (Affinity BioReagents, Golden, CO) to the HCV core protein (1:300). After extensive washing, cells were stained with fluorescein isothiocyanate-conjugated goat anti-mouse IgG antibody. Clusters of infected cells staining for core antigen were considered to constitute a single infectious focus-forming unit (FFU). Virus titers were reported as FFU/ml.

Antiviral activity assay. Wild-type, V55A and V55I variants, and the T54S V55I double variant viral RNAs were transfected as described above with serial dilutions of the ketoamide compounds boceprevir and telaprevir added separately to the growth medium. The medium was replaced with fresh medium containing boceprevir or telaprevir at 24 h, and at 24-h intervals thereafter, and secreted GLuc activity was determined 72 h later as described above. The concentration of boceprevir and telaprevir required to reduce the amount of secreted GLuc activity by 50% (antiviral 50% effective concentration [EC_{50}]) was determined using a three-parameter Hill equation (SigmaPlot 10.0).

MD simulation. Molecular dynamics (MD) simulation was performed for Val55 wild-type and V55A and V55I variants of the NS3 protease, each with ("bound") or without ("unbound") the covalently linked ketoamide inhibitor boceprevir (SCH503034). The program package GROMACS (15) and the GROMOS96 53a6 force field (25) were used for the simulations. The X-ray structure 2OC8 (26) from the Protein Data-

bank RCSB PDB (18) served as structural basis for MD. The structure contains an HCV genotype 1a NS3 protease cocrystallized with boceprevir. To obtain starting structures for the variants, the Val55 side chain from the wild-type structure was mutated using the tool IRECS (13). Boceprevir was removed from the corresponding NS3 ligand-binding sites to analyze "unbound" protease structures with empty binding pockets. Topology parameters for boceprevir were calculated by the PRODRG server (31). To describe the covalent linkage between the Ser139 residue (OG atom) and boceprevir (C34 atom), a new residue type (SER2, i.e., deprotonated SER at OG) was defined based on the predefined SER parameters of the GROMOS96 53a6 force field. Appropriate parameters for bond length and angles of the employed force field were added in the topology file to set the covalent bond. Throughout the simulations, periodic boundary conditions were applied. Long-range nonbonded interactions were treated by particle-mesh Ewald (PME) summation. The Berendsen scheme (5) was used to maintain temperature and pressure by weak coupling to an external bath with a temperature coupling relaxation time of 0.1 ps, a pressure coupling constant of 1.0 ps, and a compressibility of 4.5×10^{-5}. Bond lengths were constrained to ideal values using the LINCS procedure (14). After steepest-descent energy minimization, the systems were gradually heated from 0 to 300 K over 460 ps. A time step of 1 fs was chosen for the heat-up procedure, and position restraints were applied up to a temperature of 200 K. MD simulations were carried out at 300 K, using a time step of 2 fs and constant pressure of 1 atm for 20 ns. For the "unbound" wild-type structure, the simulation was extended by 10 ns to ensure a steady backbone root mean square deviation (RMSD). Tools of the GROMACS program were employed for analysis of trajectories. Properties (e.g., H bonds) were averaged over the last 6 ns of the trajectories unless otherwise noted. VMD (17) was applied for visualization of central cluster structures from simulated complexes.

RIN. We used PDB structure 2OC8 to generate a residue-interaction network (RIN) in two dimensions (2D). REDUCE (41) was applied for adding H-atoms to the original X-ray structure. PROBE (40) was used to identify noncovalent residue interactions. The RIN was visualized using Cytoscape (32) and the plugin RINalyzer (8). Protease residues were represented by nodes and corresponding noncovalent interactions by edges. The Cytoscape plugin NetworkAnalyzer (2) was used for calculation of Val55 topology measures within the RIN. We calculated "node degree," which is the number of edges linked to a specific node within an undirected network, as well as "node connectivity," which is the number of direct neighbors to a specified node. The "network heterogeneity" was computed to estimate the tendency of the RIN to contain highly connected nodes of putative functional importance (9). The length of a path in the RIN is given by the number of edges forming this path, whereby two nodes are possibly connected via multiple paths. We used the Cytoscape plugin Shortest Path, UCSF Resource for Biocomputing, Visualization, and Informatics (http://www.rbvi.ucsf.edu), to identify the shortest path or distance between two selected nodes. "Network diameter," which is the maximum length between two nodes in the RIN, was calculated as a reference value for the shortest path analysis. The RIN analysis considered different edge types for all possible interactions, van der Waals contacts, H bonds, overlaps, and main-chain and side-chain interactions. All measures considered weighted edges, if more than one interaction occurred between two nodes.

RESULTS

Genetic diversity of the NS3 protease at Val55. We performed a comprehensive analysis of NS3 protease sequences from 676 worldwide patient strains of the European HCV Database (euHCVdb) (7), for Val55 variants and correlated substitutions. Val55 was highly conserved throughout all genotypes and subtypes with no variants identified in HCV genotypes 1b, 2, 3, 4, and 6. Two conservative variants were observed in HCV genotype 1a, the V55A and V55I variants. A low variant frequency of four variant strains in 202 isolates was found. No covariant sequence polymorphisms in V55A variant strains were identified, whereas both V55I variant strains showed the covariant sequence polymorphism T54S (EU781818 and EF407443 strains). Two HCV genotype 5 strains with V55L and one genotype 7 strain with V55P were found. Assuming that no sequencing error occurred, the nonconservative change to Pro55 suggests considerable structural differences between the protease domains of HCV genotype 1a and genotype 7. Val55 is located buried at the tip of the β-strand C1 in an antiparallel β-sheet of the genotype 1a structure 2OC8 (37) with the main-chain NH interacting via an H bond with the side chain of Thr54. Different from the wild type, the Pro55 main-chain N atom is not available as an H-bond donor. Furthermore, Gly54 was found as neighboring residue in genotype 7 instead of Thr54, which was found in genotypes 1 to 6. Thus, the Thr54-Val55 H-bond interaction is not present in genotype 7 (38). Since ketoamide compounds are licensed for HCV genotype 1 and no variants were identified for genotype 1b, we subsequently focused on the V55A and V55I variants from HCV genotype 1a.

TABLE 1 Impact of the V55A and V55I variants, as well as the T54S V55I double variant on boceprevir resistance

Variant	IC_{50} (nM)[a]	Fold change[a]	EC_{50} (nM)[b]	Fold change[b]
Single				
Wild type	19		870 ± 48	1
V55A	80 ± 18	4.2 ± 0.96	1,360 ± 82	1.6
V55I	ND[c]		690 ± 14	0.8
Double				
T54S V55I	ND		450 ± 79	0.5

[a] IC_{50}s previously published by Susser et al. (37). The results shown represent the mean ± standard deviation and fold changes compared to the wild type.
[b] EC_{50}s determined from H77S.3/GLuc2A-transfected cell cultures. The results shown represent the mean ± standard deviation and fold changes compared to the wild type.
[c] ND, not determined.

Ketoamide resistance in Val55 variants. We introduced the V55A and V55I variants, as well as the V55I variant with its covariant T54S into separate backbones of H77S.3/GLuc2A and determined the resistance level of the HCV genotype 1a RNAs against the ketoamide compound boceprevir. We found boceprevir had antiviral activity in the genotype 1a wild type and both Val55 variants. The V55A variant showed low-level resistance, whereas the V55I variant showed no resistance against boceprevir. Average EC_{50}s were 1,360 nM and 690 nM for the V55A and V55I variants, respectively, corresponding to fold changes of 1.6- and 0.8-fold (Table 1). The V55I variant is more sensitive than the wild type for boceprevir. The T54S V55I double variant has an even lower EC_{50} of 450 nM, with a fold change of 0.5-fold. Thus, the T54S variant should not confer to ketoamide resistance in the T54S V55I double variant (Table 1).

Viral fitness of Val55 variants in HCV genotype 1a. To determine the impact of Val55 variants on viral fitness, we measured RNA replication capacity and infectious virus yields for V55A and V55I variants using the H77S.3 cell culture system.

(i) Replication capacity. The replication capacity of H77S.3/GLuc2A RNA for the V55A and V55I variants was determined by measuring GLuc activity in supernatant media collected at 24-h intervals following transfection of synthetic RNA. Results were normalized to the activity present at 8 h after transfection, as this represents GLuc expressed directly by the transfected input RNA. Compared to the parental H77S.3/GLuc2A RNA, both variants

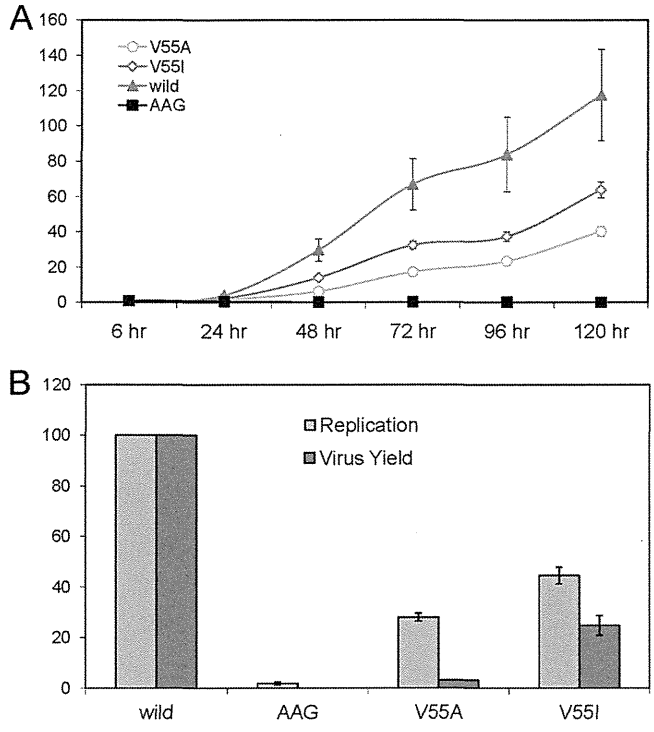

FIG 1 Replication capacity and infectious virus yield of H77S.3 RNAs with V55A and V55I substitutions in NS3. (A) RNA replication capacity reflecting the presence of H77S.3/GLuc2A V55A and V55I variants in Huh7.5 cells. Results are normalized to the 8-h GLuc activity and represent the mean of triplicate samples. (B) Comparison of RNA replication capacity (lightly shaded bars) and infectious virus yields (dark shaded bars). Both are normalized to those obtained with the relevant parental RNA (100%). The wild-type infectious titer is 2,720/ml (± standard deviation [SD] of 198/ml). Data represent the mean ± SD from at least 3 independent experiments. AAG is a replication-defective pH77S/GLuc2A mutant (43).

showed medium-level impairment of their RNA replication capacity, with the GLuc activity generally increasing after 24 h but consistently less than the parental RNA (Fig. 1). Variants showed similar courses in their RNA replication capacity, with slightly higher RNA replication in the V55I variant compared to the V55A variant. The V55I variant showed a steeper increase up to 64% compared to only 40% of parental RNA in the V55A variant. The negative impact on RNA replication correlated well with the fold changes observed in boceprevir resistance for both variants.

(ii) Infectious virus yield. We used H77S.3 RNA lacking the GLuc2A-coding sequence to assess the impact of V55A and V55I variants on the ability to produce infectious virus following transfection into Huh7.5 cells. Cell culture supernatant fluids were collected 72 h after transfection of respective variant RNAs, inoculated onto naïve Huh7.5 cells, and foci of infected cells were detected by immunofluorescence 72 h later. Both Val55 variants produced infectious virus yields lower than that expected from their RNA replication capacity. Thus, both variants directly impair infectious virus assembly or release, as described previously for other resistance-associated NS3 variants located in or near the protease ligand-binding site (34). Infectious virus yields were significantly different between V55I and V55A variants, with 25% and 3% relative infectivity, respectively (Fig. 1). Reductions in the fitness of these particular variants are confined due to defects in viral RNA replication and a significant drop in infectious virus production.

The discordances between RNA replication capacities and yields of infectious virus were higher than expected for conservative variants at a position buried in the NS3 protease domain. The large drop of infectious virus yield observed in the V55A variant resulted in a particularly low viral fitness of this variant compared to the V55I variant.

Network topology analysis for Val55 in the NS3 protease. We used a residue-interaction network (RIN) computed from an experimental 3D protease structure to analyze network topology measures and characterize the potential regulatory role of Val55 in the NS3 protease structure. The RIN comprises a total of 206 nodes, referring to residues of the NS3 protease and NS4A, as well as 1,752 edges, referring to noncovalent interactions between residues. The mean number of neighbors for a node in the NS3 protease RIN was calculated as 7.7, with a neighborhood connectivity range from 1 to 17. The majority of nodes showed only few connections with direct neighbors, whereas Val55 was highly connected with a node connectivity of 10. The node degree for Val55 was calculated with 24. The node degree distribution for the NS3 protease RIN identified 167 nodes below 24 (81%), six nodes (3%) with an equal node degree to Val55, and 33 nodes (16%) with a node degree above 24. The network diameter, which is the maximum distance between two connected nodes in the RIN, is 10 edges. Val55 showed direct contacts with 10 residues (Thr54, Tyr56, His57, Gly58, Ala59, Asp81, Ser139, Gly140, Arg155, and Ile170), including residues of the protease catalytic triad and a two-edge distance to the catalytic "oxyanion hole" (Gly137). In addition, Val55 showed van der Waals contacts with two residues of the protease ligand-binding site (Arg155 and Ile170) and indirect interaction via Thr54, Gly58, or Ser139 with Phe43, which participates in the formation of a narrow hydrophobic cavity at the S_1' pocket (38). Val55 resides in an H-bond network with Thr54, Gly58, and Ala59, without direct H-bond interaction to boceprevir. The distance to boceprevir in the RIN is two edges with a path via His57, Ser139, or Arg155 (Fig. 2). Overall, the topology measures suggest a putative structure key role for Val55 in the NS3 protease.

Molecular dynamics simulation of the Val55 wild type, as well as V55A and V55I variants. Molecular dynamics (MD) simulation was performed to further investigate the potential structure key role of Val55 in the protease structure. Simulations were carried out for wild-type and variant structures with and without the bound ketoamide boceprevir. Structures without boceprevir are referred as "unbound," whereas structures with boceprevir are referred as "bound." The variant structures were obtained by *in silico* mutation of the X-ray structure from PDB 2OC8 at Val55 with subsequent equilibration and MD simulation. Due to the covalent bond between boceprevir and Ser139, the ligand was kept in the mutated ligand-binding sites during the simulation. The final structures were analyzed for effects of Val55 variants on the H-bond pattern.

(i) Conformational stability. During simulation, the root mean square deviation (RMSD) serves as a measure for conformational stability and flexibility of the protease structure, with larger RMSDs indicating increasing structural flexibility.

(a) Backbone. The backbone RMSD of the "unbound" wild-type protease observed over time showed a longer equilibration time than was found in the variant structures. Final values and

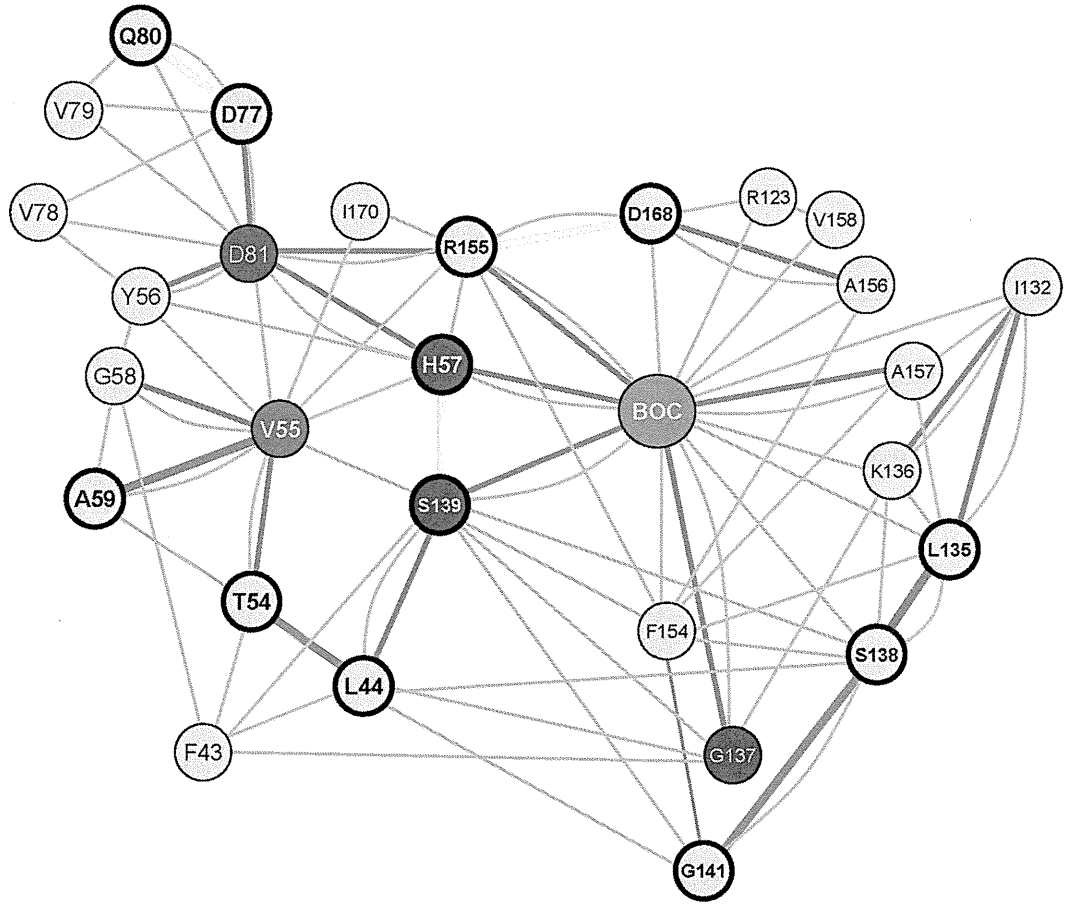

FIG 2 Residue-interaction network of the NS3 protease from PDB structure 2OC8, with nodes and edges representing protease residues and noncovalent interactions, respectively. The ketoamide compound boceprevir is represented by a single green node (BOC). The residue Val55 is given as a red node. Catalytic residues are given in blue (His57, Asp81, Gly137, and Ser139). H bonds and van der Waals contacts are represented by bold dark gray- or light gray shaded edges, respectively. To reduce complexity, multiple H-bonded or van der Waals contacts were represented by single edges between two nodes, irrespective of the number of pairwise atomic interactions. H bonds and nodes impacted due to V55A and V55I variants and/or boceprevir binding to the ligand-binding site are indicated in bold and color coded. The color coding denotes an H-bond loss upon boceprevir binding in wild-type and variant structures (cyan) and H-bond gain upon boceprevir binding in the wild-type structure (orange), as well as H-bond gain in V55A and V55I variants (magenta).

amplitudes were larger than in the corresponding "unbound" variant backbone RMSDs. In contrast, RMSDs of the "bound" protease were similar for the wild type and variants. RMSDs of "bound" variant structures were not significantly different from respective "unbound" variant structures (Fig. 3A).

(b) Single residues. The analysis of RMSDs from single residues near the protease ligand-binding site provided more detailed information on conformational changes. In the "bound" V55A variant structure, a rearrangement of His57 was observed (Fig. 3B). Similarly, a flip of His57 was found in the "unbound" wild type at the very end of the simulation for only few time steps. In contrast, the "bound" V55A variant structure showed the His57 flip at the beginning of the simulation and maintained this conformation until the end of the simulation. The larger RMSDs reflect higher structural flexibility of residues in the wild-type protease than the corresponding residues in the V55A and V55I variant structures.

(ii) H-bond network. We explored stabilizing and destabilizing effects on the H-bond pattern in the ligand-binding site for residue pairs selected from the protease residue-interaction network. We evaluated average numbers of H bonds per time frame during simulation.

(a) Wild type versus variants. The average number of H bonds differed only slightly for "bound" wild-type and variant structures. The total difference relating to all selected residue pairs was smaller than 1.5. This was different in the "unbound" protease structures (Table 2). The average number of H bonds between the residue pairs Leu135-Ser138, Arg155-Asp168, Ser138-Gly141, and Thr54-Leu44 was significantly larger for the "unbound" V55A and V55I variants than for the wild-type structure (Table 2). Additional H bonds in the V55A variant were identified between the catalytic residue pair His57-Ser139, as well as between the residue pair Gln80-Asp77 (Fig. 2). Taking all residue pairs in the ligand-binding site into account, the number of H bonds in the "unbound" variant structures compared to the wild type was increased by four for the V55I variant structure and six for the V55A variant structure.

(b) "Bound" versus "unbound" protease structures. The comparison of "bound" versus "unbound" protease structures reflects the propensity to bind boceprevir based on changes in the

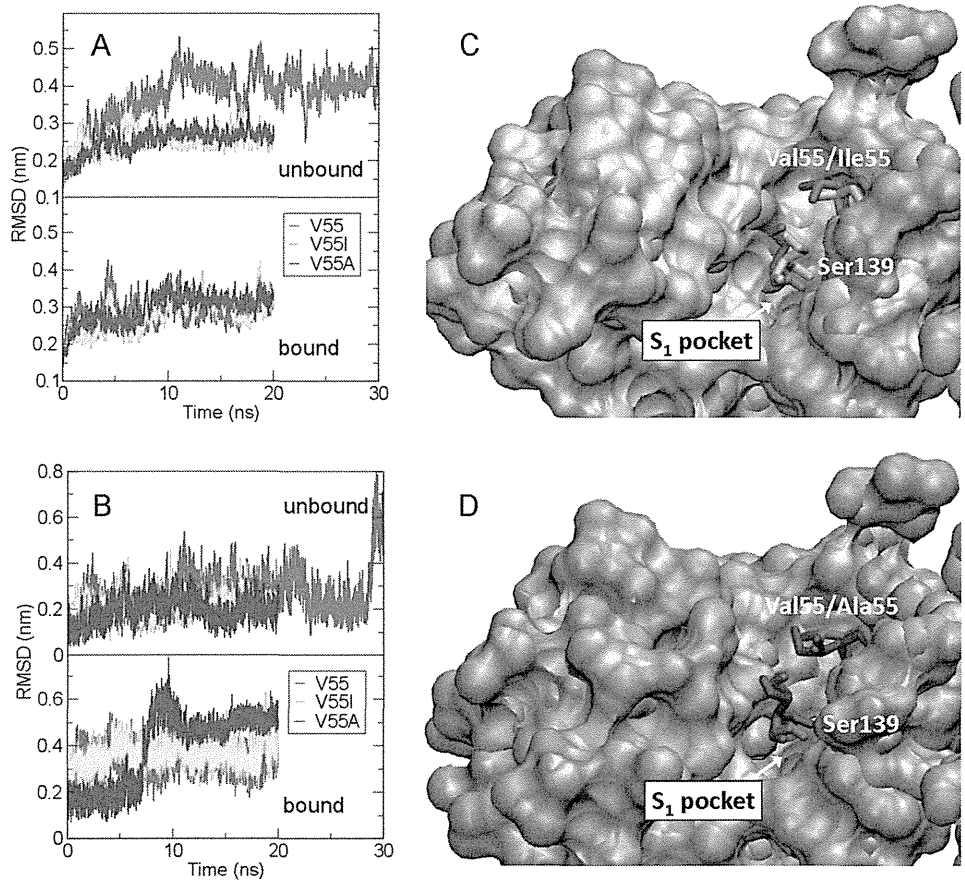

FIG 3 Backbone and side-chain root mean square deviations (RMSDs) and molecular dynamics simulations of the NS3 protease. (A) Backbone RMSDs of "unbound" and "bound" protease wild type and V55A and V55I variants. (B) RMSDs of protease residue His57 in "unbound" and "bound" wild-type and V55A and V55I variant structures. (C and D) Superposition of the "unbound" wild-type and V55I variant protease ligand-binding sites (C), as well as the "unbound" wild-type and V55A variant protease ligand-binding sites (D). Molecular changes in the S_1 pocket around Ser139 are shown in the surface representation (wild type in salmon, V55I variant in green, and V55A variant in blue). Wild-type and simulated side-chain orientations of Ser139 upon V55I and V55A mutation are depicted as stick models, with wild-type Val55, the Ile55 variant, and the Ala55 variant as red, green, and blue stick models, respectively.

H-bond network (Table 3). The comparison between the "bound" and "unbound" wild-type proteases showed an overall increase in H bonds upon boceprevir binding. The bond order for H bonds in the residue pairs Gln80-Asp77 and Arg155-Asp168 was slightly smaller in "bound" than "unbound" wild-type structures. Additional H bonds in the "bound" wild type were observed for the residue pairs Ser138-Gly141 and Thr54-Leu44. The V55A and V55I variant structures showed almost no change in H bonds for the residue pairs Ser138-Gly141 and Thr54-Leu44, while a loss of H bonds was observed for the residue pairs Gln80-Asp77 and

TABLE 2 Difference in the average number of H bonds per time frame for "unbound" variants (V55A and V55I) and wild-type (Val55) structures[a]

Residue pair	Difference in avg no. of H bonds: V55I − Val55	V55A − Val55
X55-Ala59[b]	0.2	0.6
His57-Ser139	−0.1	0.7
Leu135-Ser138	0.7	0.7
Gln80-Asp77	0.1	0.7
Arg155-Asp168	0.8	0.8
Ser138-Gly141	0.9	0.8
Thr54-Leu44	0.9	0.9
Sum[c]	3.5	5.3

[a] A negative sign denotes a loss in H bonds in the variant structure.
[b] X stands for Val (wild type), Ile (V55I variant), or Ala (V55A variant).
[c] Corresponding differences when taking all established residue pairs within the ligand-binding site into account are as follows: V55I − Val55, 4.1; and V55A − Val55, 6.1.

TABLE 3 Difference in the average number of H bonds per time frame between the "bound" and "unbound" structures for the wild-type (Val55) and variant (V55A and V55I) complexes[a]

Residue pair	Difference in avg no. of H bonds: Val55	V55I	V55A
His57-Ser139	−0.1	0.0	−0.8
Gln80-Asp77	−0.4	−1.0	−1.7
Arg155-Asp168	−0.4	−1.4	−0.6
Ser138-Gly141	0.9	0.0	0.1
Thr54-Leu44	0.9	−0.1	0.0
Sum[b]	0.9	−2.5	−3.0

[a] A negative sign denotes a loss of the H bonds upon binding of boceprevir.
[b] Corresponding differences when taking all established residue pairs within the ligand-binding site into account are as follows: Val55, 1.7; V55I, −3.8; and V55A, −4.4.

Arg155-Asp168. Furthermore, an H bond in the "unbound" but not "bound" V55A variant structure was observed between the catalytic residue pair His57-Ser139 (Fig. 2). Thus, "unbound" variant structures possessed a larger H-bond network than the wild type. Boceprevir binding was shown to disturb rather than stabilize this H-bond network. The "unbound" V55A variant structure showed more H bonds than the V55I variant structure with greater destabilizing impact due to lost H bonds upon boceprevir binding.

(iii) Shape of the NS3 protease ligand-binding site. To estimate the facility for ketoamides approaching the protease ligand-binding site, we analyzed structural changes during MD simulation for a representative pocket in the "unbound" protease structures (S_1 pocket). The minimum distances between Ser139 and its next neighbor residues in the S_1 pocket were calculated. We found most S_1 pocket residues in V55A and V55I variant structures closer to Ser139 compared to the wild-type structure. The overall S_1 pocket size was smaller in the variant structures with reduced accessibility to the ligand-binding site. Exceptions were found predominantly for residues involved in the formation of S_1' and S_2' pockets (i.e., Gln41, Thr42, and Lys136).

DISCUSSION

Drug-resistant viral variants are likely to preexist at a low frequency in the replicating viral quasispecies population of the typical HCV-infected patient (29, 39). The V55A variant was previously identified as resistance-associated amino acid variant against the ketoamide compound boceprevir, with Val55 suggested as a regulatory site in the NS3 protease structure (37) involved in resistance development and variant fitness. A recent study found the V55A variant in the long-term follow up of several patients previously treated with boceprevir. Moreover, this variant was identified as dominant baseline strain in one of the patients before treatment. Thus, variations at Val55 are likely to be clinically relevant in patients treated with ketoamide compounds (36). In the present study, we analyzed the natural variation present among 676 sequences from genotypes 1 to 7, collected from geographically diverse sites and deposited in a public database. They likely represent variants present within the dominant quasispecies of the patients from which these sequences were derived (7). We found Val55 highly conserved in all major HCV genotypes and subtypes; however, few conservative variants and one nonconservative variant were identified. Clinically most important, HCV genotype 1a showed four strains from 202 sequences with variations at Val55—two strains with V55A and two strains with V55I—whereas, no variants were found in 335 sequences of genotype 1b HCV.

We used a residue-interaction network approach to characterize the potential key role of Val55 in the NS3 protease structure related to resistance development and viral fitness in Val55 variants. The protease residue-interaction network identified only a few nodes in the network highly connected with their neighboring nodes, while the majority show only a few connections. Val55 is such a highly connected node with connections to protease catalytic residues and the ligand-binding site. The high connectivity points to a potential key role for Val55 with important function in the respective network and protein structure (9), where variants could have an impact on viral fitness and drug resistance development.

To further investigate this role of Val55 and its variants in the NS3 protease structure, we applied molecular dynamics simulation, which is a proven approach to study molecular changes and mechanisms potentially related to drug resistance, already successfully applied to other NS3 protease resistance-associated amino acid variants (38) (S. Schweizer et al., submitted for publication). Analyzing protein backbone RMSDs from the molecular dynamics simulations, we found higher structural flexibility in the wild-type protease than the variant structures. This was further supported by the larger conformational mobility and higher RMSDs of single residues in the wild-type ligand-binding site. Corresponding with that observation, we found fewer H bonds in the local H-bond network for the "unbound" wild type than variant structures. The wild type showed a gain of H bonds upon boceprevir binding, while the variants showed a loss of H bonds, which was most prominent in the V55A variant, with the lowest structural flexibility in that variant. This explains that ketoamides potentially fit easier into the ligand-binding site of the wild type than the Val55 variant structures, particular in that of the V55A variant. Structure simulations are in agreement with resistance data from the cell culture. The V55A variant led to a 1.6-fold increase in EC_{50} for boceprevir and a 3-fold increase in the telaprevir EC_{50}, which is likely to be clinically significant. The V55I variant showed no fold increase in EC_{50} for boceprevir and only a 1.4-fold increase in EC_{50} for telaprevir. Interestingly, the T54S V55I double variant was found for all V55I variant strains but showed no resistance against boceprevir, whereas a clinically significant 7.9-fold increase in the telaprevir EC_{50} was found, which is much higher than that of the V55I single variant (unpublished data). Accessibility of the ligand-binding site seems to play a crucial role in ketoamide binding. Since it is problematic to calculate absolute volumes of open binding pockets, we used the minimum distance between Ser139 and its next neighbor residues to compute a measure for size and volume of a representative ligand-binding site pocket (S_1 pocket). Binding pockets of the "unbound" wild-type structure were superimposed and compared for differences in shape and volume with respective pockets of "unbound" V55I and V55A variant structures (Fig. 3C and D). The variant structures showed most neighboring residues closer to Ser139 with constricted pockets in their ligand-binding sites. Apart from a few exceptions, the minimum distances were smaller than or similar to those obtained for the wild-type pockets. Since Ala55 needs less space than Ile55, we observed overall smaller pockets in V55A than V55I. Molecular changes in the ligand-binding site and effects on the H-bond pattern, together, explain the resistance levels observed in Val55 variants for HCV genotype 1a.

Viral variant fitness is crucial to select resistant variants from the quasispecies population under drug pressure. RNA replication capacity is one measure of the fitness of the virus, and this is dependent on proper processing of the polyprotein by the NS3 protease. V55A and V55I variants showed RNA replication of 28% and 44.5%, respectively, compared to the wild type. Ketoamides mimic the natural substrate of the protease at the site of NS3-NS4A scission, and it is likely that the constricted Val55 variant binding pockets lead to difficulties in protease-substrate interaction and interfere with substrate recognition and cleavage. This is in agreement with the stronger negative impact on RNA replication observed in the V55A variant compared to the V55I variant. Although yields of infectious virus generally correlate well with the RNA replication capacity (34, 39), both Val55 variants were found

involved in a drop of infectious virus yield leading to relative infectivity (compared to wild type) of 24.8% for the V55I variant and only 3.1% for V55A. Variants leading to impairments in infectious virus yield are likely to affect NS3 domain-domain interactions between the protease and helicase (3, 4, 20, 34). Corresponding to that, we identified the "unbound" wild-type protease with larger conformational flexibility during molecular dynamics simulation than the Val55 variant structures. The higher structural flexibility might allow the wild type to adapt more efficiently to conformational changes needed for NS3 domain-domain interaction and could serve as a possible explanation for the virus yield drop in Val55 variants with the lowest structural flexibility and corresponding largest virus yield drop in the V55A variant. The steep decline in infectious virus yield is surprising given the fact that the V55A variant was found repeatedly in a public database and as the dominant baseline strain in a patient before ketoamide exposure (36). Furthermore, it was found during long-term follow-up of 5 of 12 patients upon boceprevir treatment (36). Loss of fitness is likely to negatively impact the persistence of NS3 resistance-associated variants and their ability to compete with wild-type virus upon discontinuation of antiviral therapy. Nonetheless, it is possible that they could become fixed in the viral quasispecies by compensatory second-site mutations. Such second-site changes may explain how the V55A variant that negatively impacts replication and infectious virus yield in H77S.3 cell culture could dominate in some treatment-naïve and -experienced patients. Such second-site substitutions could exist within NS3 or outside NS3, so identifying putative compensatory amino acid changes in the same viral strain is a difficult task and beyond the scope of this article.

In summary, we identified preexisting Val55 variants in the NS3 protease, providing *a priori* for resistance against ketoamides. The topology of Val55 in a residue-interaction network of the NS3 protease indicates a potential structure key role to modulate molecular changes in the ligand-binding site and protease-helicase interaction interface. Val55 variants showed compromised viral fitness due to reduced RNA replication capacity and a distinct drop in their infectious virus yields. Molecular dynamics simulations revealed structural changes in conjunction with fitness costs and drug resistance. Structural flexibility might be important to adapt for NS3 domain-domain interactions involved in particle assembly with lower structural flexibility as a possible explanation for reduced yields in variant viruses.

ACKNOWLEDGMENTS

The present study was supported by a DFG grant (Clinical Research Unit, CRU129, TP3, LE 491/16-2) and a DFG Research Fellowship (WE 4388/3-1) to C.W. S.M.L. was supported by grants and contracts from the National Institute for Allergy and Infectious Diseases and the NIH (U19-AI40035, R21-AI81058, and N01-AI25488). S.S. is grateful for financial support by the CIPS^M Gender Support Program. I.A. acknowledges financial support by the CIPS^M excellence cluster.

S.M.L. has served as a consultant for Abbott, Hoffmann-LaRoche, Juvaris Biotherapeutics, Merck, Novartis, and Pfizer; research in his laboratory is supported by Merck and Tibotec. The remaining authors disclose no conflicts.

REFERENCES

1. **Alter HJ, Seeff LB.** 2000. Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. Semin. Liver Dis. **20**:17–35.
2. **Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M.** 2008. Computing topological parameters of biological networks. Bioinformatics **24**:282–284.
3. **Beran RK, Pyle AM.** 2008. Hepatitis C viral NS3-4A protease activity is enhanced by the NS3 helicase. J. Biol. Chem. **283**:29929–29937.
4. **Beran RK, Serebrov V, Pyle AM.** 2007. The serine protease domain of hepatitis C viral NS3 activates RNA helicase activity by promoting the binding of RNA substrate. J. Biol. Chem. **282**:34913–34920.
5. **Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR.** 1984. Molecular dynamics with coupling to an external bath. J. Chem. Phys. **81**:3684.
6. **Chenna R, et al.** 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. **31**:3497–3500.
7. **Combet C, et al.** 2007. euHCVdb: the European Hepatitis C Virus Database. Nucleic Acids Res. **35**:D363–D366.
8. **Doncheva NT, Klein K, Domingues FS, Albrecht M.** 2011. Analyzing and visualizing residue networks of protein structures. Trends Biochem. Sci. **36**:179–182.
9. **Dong J, Horvath S.** 2007. Understanding network concepts in modules. BMC Syst. Biol. **1**:24.
10. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797.
11. **Fried MW, et al.** 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N. Engl. J. Med. **347**:975–982.
12. **Galtier N, Gouy M, Gautier C.** 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. **12**:543–548.
13. **Hartmann C, Antes I, Lengauer T.** 2007. IRECS: a new algorithm for the selection of most probable ensembles of side-chain conformations in protein models. Protein Sci. **16**:1294–1307.
14. **Hess B, Bekker H, Berendsen HJC, Fraaije JGEM.** 1997. LINCS: a linear constraint solver for molecular simulations. J. Comput. Chem. **18**:1463.
15. **Hess B, Kutzner C, van Der Spoel D, Lindahl E.** 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput. **4**:1549.
16. **Hezode C, et al.** 2009. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. N. Engl. J. Med. **360**:1839–1850.
17. **Humphrey W, Dalke A, Schulten K.** 1996. VMD: visual molecular dynamics. J. Mol. Graph. **14**:33–38.
18. **Kouranov A, et al.** 2006. The RCSB PDB information portal for structural genomics. Nucleic Acids Res. **34**:D302–D305.
19. **Kwo PY, et al.** 2010. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. Lancet **376**:705–716.
20. **Ma Y, Yates J, Liang Y, Lemon SM, Yi M.** 2008. NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. J. Virol. **82**:7624–7639.
21. **Manns MP, et al.** 2001. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet **358**:958–965.
22. **McHutchison JG, et al.** 2009. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. N. Engl. J. Med. **360**:1827–1838.
23. **McHutchison JG, et al.** 2010. Telaprevir for previously treated chronic HCV infection. N. Engl. J. Med. **362**:1292–1303.
24. **Moradpour D, Penin F, Rice CM.** 2007. Replication of hepatitis C virus. Nat. Rev. Microbiol. **5**:453–463.
25. **Oostenbrink C, Villa A, Mark AE, van Gunsteren WF.** 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J. Comput. Chem. **25**:1656–1676.
26. **Prongay AJ, et al.** 2007. Discovery of the HCV NS3/4A protease inhibitor (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(S)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(S)-carboxamide (Sch 503034). II. Key steps in structure-based optimization. J. Med. Chem. **50**:2310–2318.
27. **Qiu P, et al.** 2009. Identification of HCV protease inhibitor resistance mutations by selection pressure-based method. Nucleic Acids Res. **37**:e74.
28. **Raney KD, Sharma SD, Moustafa IM, Cameron CE.** 2010. Hepatitis C virus non-structural protein 3 (HCV NS3): a multifunctional antiviral target. J. Biol. Chem. **285**:22725–22731.
29. **Rong L, Dahari H, Ribeiro RM, Perelson AS.** 2010. Rapid emergence of

protease inhibitor resistance in hepatitis C virus. Sci. Transl. Med. 2:30ra32.
30. **Sarrazin C, et al.** 2007. SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon alpha-2b for genotype 1 nonresponders. Gastroenterology **132**:1270–1278.
31. **Schuttelkopf AW, van Aalten DM.** 2004. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. **60**:1355–1363.
32. **Shannon P, et al.** 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **13**:2498–2504.
33. **Shiffman ML, et al.** 2007. Peginterferon alfa-2a and ribavirin for 16 or 24 weeks in HCV genotype 2 or 3. N. Engl. J. Med. **357**:124–134.
34. **Shimakami T, et al.** 2011. Protease inhibitor-resistant hepatitis C virus mutants with reduced fitness from impaired production of infectious virus. Gastroenterology **140**:667–675.
35. **Simmonds P, et al.** 2005. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology **42**:962–973.
36. **Susser S, et al.** 2011. Analysis of long-term persistance of resistance mutations within the hepatitis C virus NS3 protease after treatment with telaprevir or boceprevir. J. Clin. Virol. **52**:321–327.
37. **Susser S, et al.** 2009. Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients. Hepatology **50**:1709–1718.
38. **Welsch C, et al.** 2008. Molecular basis of telaprevir resistance due to V36 and T54 mutations in the NS3-4A protease of the hepatitis C virus. Genome Biol. **9**:R16.
39. **Welsch C, et al. 7 December** 2011, posting date. Peptidomimetic escape mechanisms arise due to genetic diversity in the ligand-binding site of the HCV NS3/4A serine protease. Gastroenterology [Epub ahead of print.]
40. **Word JM, et al.** 1999. Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. J. Mol. Biol. **285**:1711–1733.
41. **Word JM, Lovell SC, Richardson JS, Richardson DC.** 1999. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. **285**:1735–1747.
42. **Yi M, et al.** 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.
43. **Yi M, Villanueva RA, Thomas DL, Wakita T, Lemon SM.** 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.
44. **Zeuzem S, et al.** 2004. Peginterferon alfa-2b plus ribavirin for treatment of chronic hepatitis C in previously untreated patients infected with HCV genotypes 2 or 3. J. Hepatol. **40**:993–999.

Peptidomimetic Escape Mechanisms Arise via Genetic Diversity in the Ligand-Binding Site of the Hepatitis C Virus NS3/4A Serine Protease

CHRISTOPH WELSCH,*,‡,§ TETSURO SHIMAKAMI,* CHRISTOPH HARTMANN,§ YAN YANG,*
FRANCISCO S. DOMINGUES,‖ THOMAS LENGAUER,§ STEFAN ZEUZEM,‡ and STANLEY M. LEMON*

Division of Infectious Diseases, Department of Medicine, Inflammatory Diseases Institute, and the Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; ‡Department of Internal Medicine I, J. W. Goethe-University Hospital, Frankfurt, Germany; §Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarbrücken, Germany; ‖Institute of Genetic Medicine, EURAC Research, Bolzano, Italy

BACKGROUND & AIMS: It is a challenge to develop direct-acting antiviral agents that target the nonstructural protein 3/4A protease of hepatitis C virus because resistant variants develop. Ketoamide compounds, designed to mimic the natural protease substrate, have been developed as inhibitors. However, clinical trials have revealed rapid selection of resistant mutants, most of which are considered to be pre-existing variants. **METHODS:** We identified residues near the ketoamide-binding site in x-ray structures of the genotype 1a protease, co-crystallized with boceprevir or a telaprevir-like ligand, and then identified variants at these positions in 219 genotype-1 sequences from a public database. We used side-chain modeling to assess the potential effects of these variants on the interaction between ketoamide and the protease, and compared these results with the phenotypic effects on ketoamide resistance, RNA replication capacity, and infectious virus yields in a cell culture model of infection. **RESULTS:** Thirteen natural binding-site variants with potential for ketoamide resistance were identified at 10 residues in the protease, near the ketoamide binding site. Rotamer analysis of amino acid side-chain conformations indicated that 2 variants (R155K and D168G) could affect binding of telaprevir more than boceprevir. Measurements of antiviral susceptibility in cell-culture studies were consistent with this observation. Four variants (ie, Q41H, I132V, R155K, and D168G) caused low-to-moderate levels of ketoamide resistance; 3 of these were highly fit (Q41H, I132V, and R155K). **CONCLUSIONS: Using a comprehensive sequence and structure-based analysis, we showed how natural variation in the hepatitis C virus protease nonstructural protein 3/4A sequences might affect susceptibility to first-generation direct-acting antiviral agents. These findings increase our understanding of the molecular basis of ketoamide resistance among naturally existing viral variants.**

Keywords: Virology; Genetic; Drug Resistance; Treatment.

Until recently, the standard of care (SOC) for patients with chronic hepatitis C virus (HCV) infection has consisted of a combination of pegylated interferon-alfa plus ribavirin (Peg-IFN/RBV) administered for 24 to 48 weeks, depending on HCV genotype.[1–3] The sustained virologic response rate for this SOC has been only about 50% in patients infected with genotype 1 HCV, the most prevalent genotype in Europe and North America. The addition of a direct-acting antiviral agent (DAA) targeting the nonstructural protein (NS) 3/4A serine protease of HCV significantly improves the sustained virologic response rate, and 2 such drugs have recently been approved for clinical use in Europe and the United States. The ketoamide compounds boceprevir (SCH503034) and telaprevir (VX-950) were both designed to mimic the natural substrate of the protease.[4–6] Clinical trials in treatment-naïve genotype 1–infected patients have revealed significant improvements in the kinetic of the virologic response with the addition of a DAA to the earlier SOC, leading to improved sustained virologic response rates of up to 74%.[7–9]

Unlike Peg-IFN/RBV, the selection of drug-resistant virus variants during treatment with protease inhibitors (PIs) is a major concern. According to recent calculations by Rong et al,[10] most, if not all, potential drug-resistant viral variants pre-exist at low frequencies within the viral quasi-species population in untreated patients. The highly replicative nature of HCV infection, with approximately 10^{12} new virions produced each day in the typical infected individual,[11] coupled with the lack of proofreading activity in the RNA-dependent RNA polymerase, NS5B, results in generation of every possible viral variant every day. Thus, each patient is infected with a viral quasi-species "cloud" comprised of genetically distinct but closely related viral genomes. In the absence of concomitant Peg-IFN/RBV therapy, drug-resistant viral variants are rapidly selected and can emerge at frequencies as high as 5% to 20% in the quasi-species of patients as early as the second day of treatment. Unless suppressed by concomitant PegIFN/RBV, these pre-existing resistant variants are likely to be selected with subsequent treatment failure.[10]

The NS3/4A protease plays an essential role in the HCV replication cycle by proteolytically processing nonstruc-

Abbreviations used in this paper: BSV, binding-site variant; DAA, direct-acting antiviral agent; EC_{50}, 50% effective concentration; GLuc, Gaussia luciferase; NS, nonstructural protein; PDB, Protein Databank; Peg-IFN/RBV, pegylated interferon-alfa plus ribavirin; PI, protease inhibitor; SOC, standard of care; TLL, telaprevir-like ligand.

0016-5085/$36.00
doi:10.1053/j.gastro.2011.11.035

tural proteins from the viral polyprotein downstream of the NS2-3 junction.[12] The protease domain of NS3, comprising the amino-terminal third of the protein contains a catalytic triad, H57, D81, and S139, and an "oxyanion hole" at G137. It acts in concert with its cofactor, NS4A, which intercalates into its structure and is required for full enzymatic activity and proper subcellular localization. The carboxy-terminal two thirds of NS3 comprises a DExD-box RNA helicase domain that is essential for productive viral infection.[13] NS3 appears to be a critical component of the macromolecular viral RNA replicase that directs the synthesis of new viral RNAs. Genetic evidence indicates that NS3 has an additional distinct function in the assembly of virus particles.[14,15] Because viral RNA replication capacity and virus assembly are crucial determinants of viral fitness, mutations in NS3 that contribute to PI resistance can also profoundly influence virus fitness.[16] The probability of a resistant variant emerging from the quasi-species population during treatment with a DAA is determined not only by its degree of resistance, but also by its fitness. Many mutations associated with PI resistance negatively impact the replication of genotype 1a HCV RNA in cell culture, while some have additional effects on the production of infectious virus.[14] Compensatory second-site mutations can enhance the fitness of resistant viruses,[16] but current understanding of these is rudimentary.

Given the genetic diversity that exists among different HCV strains, it is possible that naturally occurring polymorphisms in the NS3/4A sequence could provide a priori resistance to DAAs and negatively impact the success of future treatment regimens. Here, we have studied the variation in amino acid residues that neighbor ketoamide compounds in the ligand-binding site of the protease. We identified natural amino acid substitutions at these positions in NS3 among genotype 1a sequences deposited in a public database, and modeled their side-chain conformations to assess their potential impact on ketoamide binding. To corroborate these in silico predictions, we then introduced these amino acid substitutions into a cell culture-infectious genotype 1a virus (H77S.3)[14] and determined their impact on both susceptibility to ketoamide PIs and replication fitness in a cell culture system.

Materials and Methods

Details of the Materials and Methods can be found in the Supplementary Materials.

In Silico Analysis

We used x-ray structures of the genotype 1a HCV NS3/4A protease from the RCSB Protein Databank (PDB)[17] co-crystallized with boceprevir (PDB 2OC8) or a telaprevir-like ligand (TLL; PDB 2P59) to deduce sets of ketoamide-neighboring residues. We designated the P_4 to P_1 and P_1' groups for ligands and their corresponding specificity pockets within the ligand-binding site, S_4 to S_1 and S_1', according to the numbering scheme of Schechter and Berger.[18] We then analyzed 219 genotype 1a HCV NS3/4A sequences deposited in the European HCV database,[19] which contains sequences collected from around the world, to identify potential natural binding site variants (BSVs) at residues that neighbor the ketoamides within the structure of the protease. The side-chain conformations of these BSVs were modeled using IRECS (Iterative REduction of Conformational Space)[20] (details in Supplementary Material).

Cell Culture and Reagents

Details of the cells and reagents used in this study are provided in Supplementary Material.

Plasmids

The pH77S.3 and pH77S.3/Gaussia luciferase (GLuc)2A are molecular clones of the genotype 1a H77 strain of HCV. Synthetic RNA transcribed from these plasmids replicates in transfected Huh7 cells and produces infectious virus.[14] pH77S.3/GLuc2A RNA also produces secreted GLuc reporter protein. Amino acid substitutions in BSVs expected to impact ketoamide binding were created in these plasmids by site-directed mutagenesis.[14]

Virus Fitness and Antiviral Resistance

Genome-length RNA was transcribed from the mutated pH77S.3 and pH77S.3/GLuc2A plasmids in vitro, and studies to assess antiviral resistance and viral fitness carried out as described previously.[14]

Results

To identify amino acid residues in NS3 that are in close proximity to ketomides in the ligand-binding site of the protease, we analyzed PDB structure 2OC8 in which boceprevir is co-crystallized with the NS3/4A protease.[21] Because no similar co-crystallized structure is publically available for telaprevir, we used PDB structure 2P59 in which the protease is co-crystallized with a TLL that has 2 small differences from telaprevir.[22] Its P_2 group and the P_4 capping group are slightly modified, with the P_4 providing a pyrrole NH for H-bond interactions with the protease (Supplementary Figure 1).[23] We identified 20 residues interacting with or neighboring boceprevir and TLL in these x-ray structures (Table 1 and Supplementary Materials and Methods) and analyzed genotype 1a HCV sequences in the European HCV database to identify differences from the consensus sequence at these residues. The residues were relatively well conserved in 219 genotype 1a sequences from diverse geographic regions. However, we identified 13 different BSVs involving 10 ketoamide-neighboring residues (Supplementary Table 1 and Supplementary Figure 2), 8 of which have not been described previously as PI resistance-associated variants. None of the patients from whom these sequences were derived appear to have been treated previously with a DAA (see Supplementary Table 3). Most of these naturally occurring BSVs were single amino acid substitutions without changes in other ketoamide-neighboring residues. However, T42A and K136R were both present within a single sequence (EU677251). Two different BSVs were identified as substitutions at T42 (A/S), V55 (A/I), and D168 (G/E), while single BSVs were identified at the other 7 residues

Table 1. Ketoamide-Neighboring Residues in NS3[a]

Boceprevir		Telaprevir-like ligand	
RCSB PDB 2OC8 (chain B)	UniProtKB P27958	RCSB PDB 2P59 (chain B)	UniProtKB P27958
Q41	Q41	Q1067	Q41
T42	*T42*[b]	T1068	T42
F43	F43	F1069	F43
V55	*V55*	–	–
H57	H57	H1083	H57
D81	*D81*	*D1107*	*D81*
R123	R123	*R1149*	*R123*
I132	I132	I1158	I132
L135	L135	*L1161*	*L135*
K136	K136	K1162	K136
G137	G137	G1163	G137
S138	S138	S1164	S138
S139	S139	S1165	S139
F154	F154	F1180	F154
R155	R155	R1181	R155
A156	A156	A1182	A156
A157	A157	A1183	A157
V158	V158	V1184	V158
C159	C159	C1185	C159
	–	*T1186*	*T160*
D168	D168	D1194	D168

[a]Residues are numbered relative to the PDB residue sequence and H77c reference sequences (UniProtKB P27958).
[b]Residues shown in italics do not directly interact with boceprevir or TLL, but are within 5.0 Å distance of the ligand (see Supplementary Material).

(Supplementary Table 1 and Supplementary Figure 2). We examined the first-neighbor residues of these BSVs in an effort to identify potential second-site variants in the BSV strains (see Supplementary Material, Supplementary Figure 3, and Supplementary Table 2). Although there were no additional substitutions at these first-neighbor residues in most BSVs, T54S was found in both V55I strains (EF407443 and EU781818) and I170V was found in a R155K strain (EU781805).

Rotamer Analysis of BSVs and Expected Impact on Ketoamide Binding

We modeled the energetically most favorable side-chain conformations for the genotype 1a consensus and BSV sequences (Figure 1). Because V55 is buried in the protease domain,[24] it was not amenable to rotamer analysis. Molecular dynamics simulation of variants at this position will be reported separately (Welsch et al, Antimicrob Agents Chemother, accepted manuscript).

Expected impact of Q41H and K136R at the ketoamide P_1' position. Q41 is located adjacent to S_1'. The H41 variant represents a nonconservative change from the consensus sequence involving substitution of an uncharged polar side chain with a charged aromatic residue. Telaprevir possesses a cyclopropyl group at P_1' that is oriented away from the Q41 side chain without noncovalent interactions. The conformation of Q41 suggests a potential H-bond interaction of its carboxamide chain, including the H-bond donor NH_2 and acceptor C=O, with the backbone of boceprevir and TLL (OE1—HN distance for Q41-ketoamide: 2.7 Å). However, the H41 BSV has no side-chain group allowing an H-bond interaction with the ligand (Figure 2*A*). This suggests H41 will be associated with a minor decrease in binding affinity for boceprevir and telaprevir.

In contrast, K136R is a conservative change between polar residues, both with long, flexible, and charged side chains at S_1'. The K136 side chain is not defined in the electron density map of PDB structure 2P59. As modeled, it has close contacts with the TLL P_1' cyclopropyl group (Figure 2*B*). Boceprevir possesses no comparable P_1' group

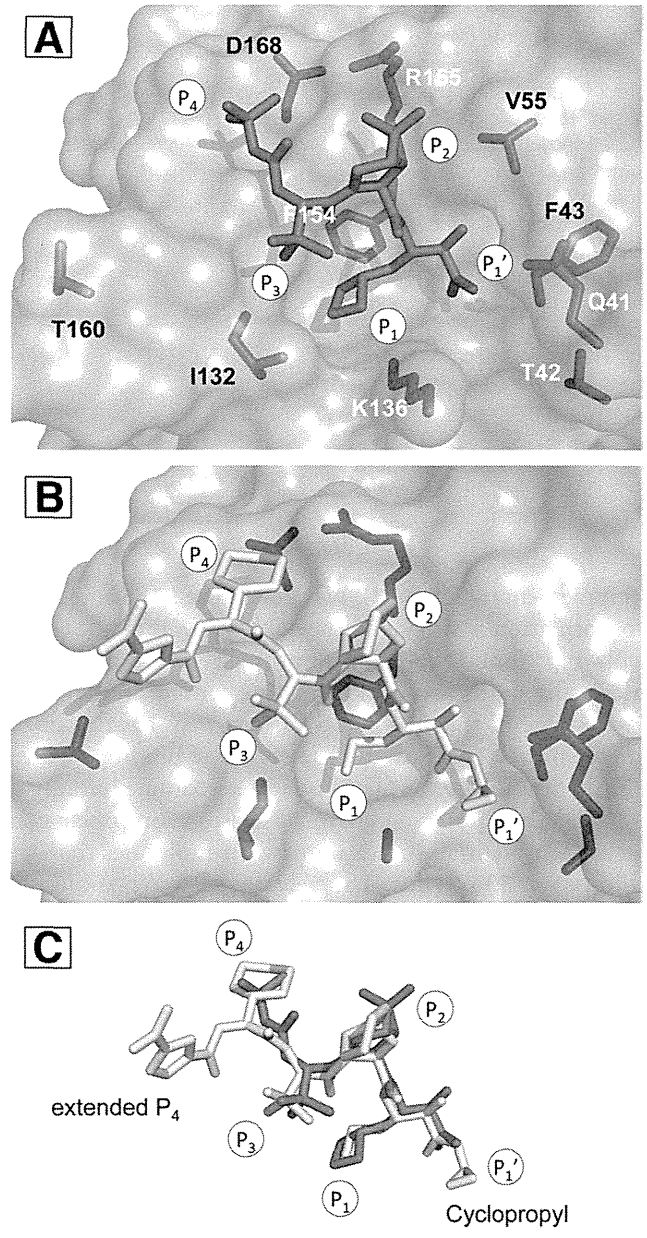

Figure 1. Ligand-binding site of NS3/4A with ketoamides and BSVs. (*A*) Surface representation from PDB structure 2OC8 with boceprevir shown in *magenta*. Side chains at amino acid residues with BSVs are shown in stick format (*light blue*). (*B*) Similar representation of TLL within the PDB structure 2P59 with side chains at amino acid residues with BSVs (*dark blue*). (*C*) Superposition of boceprevir and TLL.

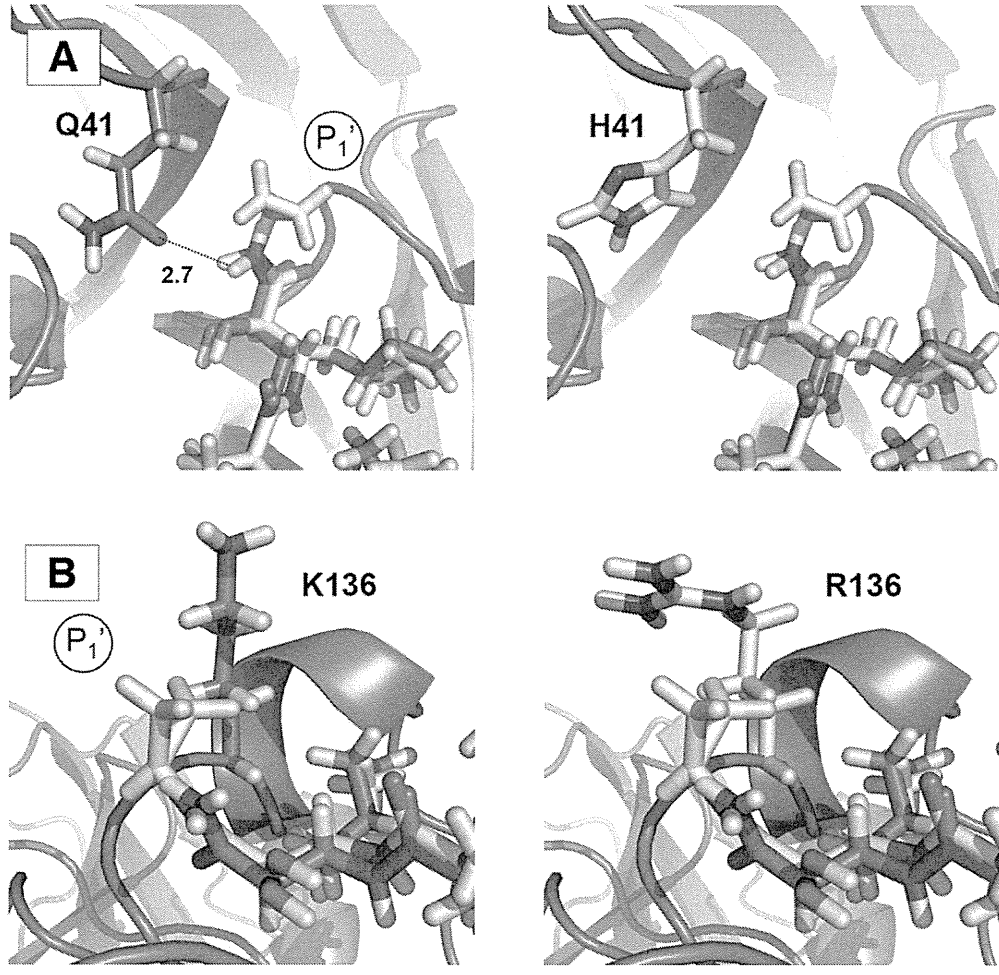

Figure 2. Rotamer analysis of Q41H and K136R. Detail of PDB structure 2OC8 showing boceprevir (*magenta*) and the superposed TLL (*yellow*). Panels on the *left* show the side-chain conformation of the consensus residue, and those on the *right* show the predicted BSV conformation. (*A*) *Left*: The Q41 side-chain OH-group with predicted H-bond interaction (*dotted line*) with ketoamide backbone Cα (*right*), this H-bond interaction is absent in H41. (*B*) Predicted R136 side-chain conformation and the TLL P_1' cyclopropyl showing closer van der Waals contacts in (*right*) R136 than in (*left*) the consensus K136.

(Supplementary Figure 1). We predict the R136 side-chain conformation is similar to the wild-type K136, but that it will wrap around the TLL cyclopropyl group, providing an increasing number of van der Waals contacts (Figure 2*B*). This suggests that the R136 variant may have stronger effects on telaprevir than on boceprevir binding. We expect hydrophilicity to play a role in ketoamide binding at S_1', and the R136 side chain possesses more polar contacts and potentially binds more water molecules than the wild-type K136. Thus, the release of water molecules bound to the R136 side chain could increase systemic entropy and potentially add to the free binding energy and affinity for telaprevir, although decreased enthalpy may compensate for the effects on entropy. Because the R136 side chain does not have a tight H-bond or salt bridge interaction at its end, electrostatic attraction between the nonpolar P_1' cyclopropyl group of telaprevir and the polar R136 side chain is unlikely. Overall, however, the impact of K136R on binding of telaprevir is not readily predicted by rotamer analysis alone.

Expected impact of I132V and F154Y on ketoamide P_1 and P_3 interactions. I132 is located adjacent to S_1 and S_3, and its substitution with valine in V132 represents a conservative change between hydrophobic residues. The consensus I132 side-chain conformation forms a hydrophobic $S_{1/3}$ interface with the ketoamide P_1 and P_3 groups. The I132 side-chain C_δ carbon points toward the ketoamide P_3 group (Figure 3*A*). The variant V132 shows a similar side-chain conformation but without a C_δ carbon. This suggests a loss of van der Waals contacts for V132, resulting in a minor reduction in binding affinity for boceprevir and telaprevir. F154Y is also a relatively conservative change, in this case from nonpolar toward polar, among aromatic residues. This residue is close to the ketoamide P_1 and P_3 groups at the bottom of the S_1 pocket. The wild-type F154 and the predicted variant Y154 side chains are oriented toward the ketoamide P_1 group, potentially influencing boceprevir and telaprevir binding (Figure 3*B*). The Y154 side chain has an OH group that is not present in the wild-type F154. Ketoamides do not offer an opportunity for H-bond interactions at P_1, but this OH group may provide for alternative binding of a water molecule. The polar, hydrophilic nature of the Y154 side chain reduces the hydrophobic properties of the S_1 pocket, and this variant is expected to cause a significant reduction in binding affinity for boceprevir and telaprevir.

Expected impact of R155K, D168E, and D168G on ketoamide P_2 and P_4 interactions. R155K is a conservative change between positively charged residues within the S_2 pocket. The R155 side chain is predicted to participate in a pattern of noncovalent interactions involving its neighboring residues, R123 and D168, at S_4 (the O-H

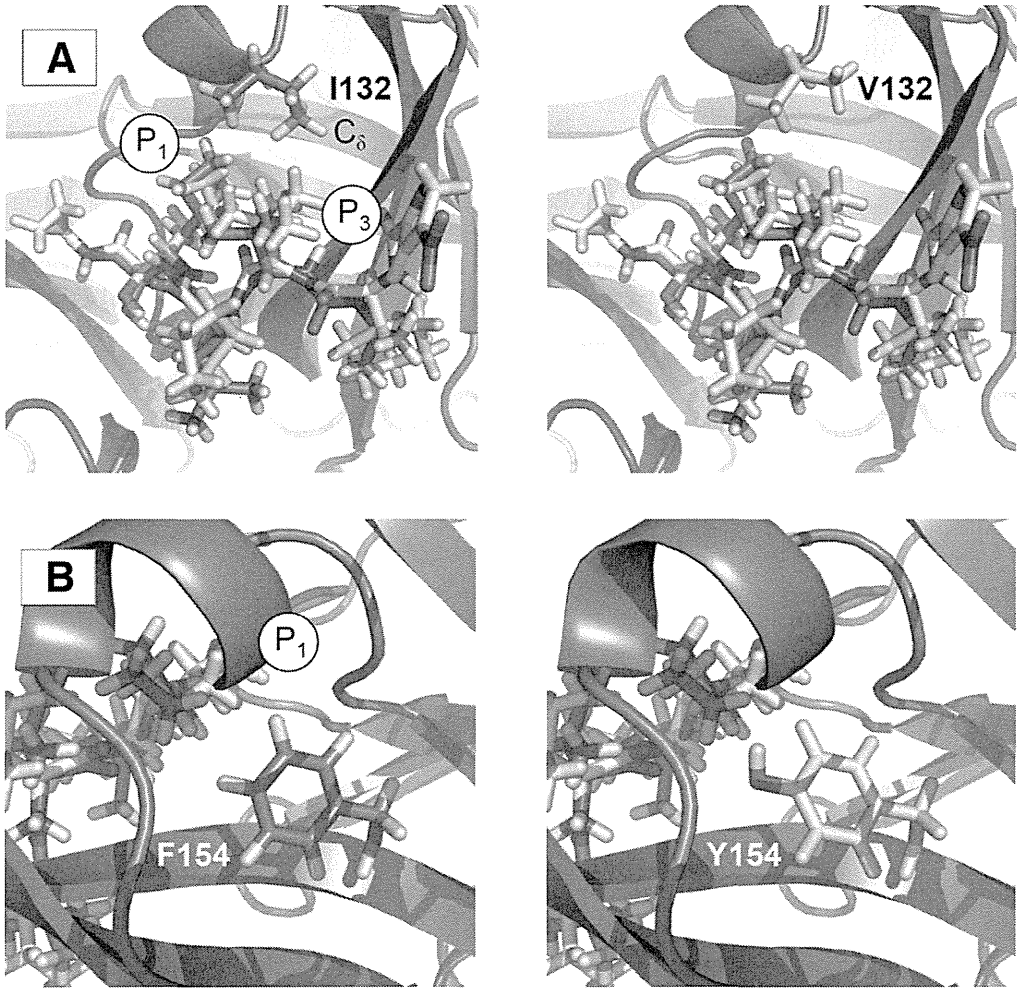

Figure 3. Rotamer analysis of I132V and F154Y. (*A*) *Left*: The I132 side chain provides a hydrophobic $S_{1/3}$ interface for the ketoamide P_1 and P_3 moieties. The I132 C_δ carbon is oriented toward the ketoamide P_3, while (*right*) V132 lacks a C_δ carbon. (*B*) *Left*: F154 neighbors S_1 underneath the ketoamide, while (*right*) Y154 is predicted to assume a similar side-chain conformation but with an additional OH-group.

distance for D168-R155 is 2.6 Å; for R123-D168, it is 2.5 Å) (Figure 4). The combination of H-bond and electrostatic interactions is predicted to result in a particularly strong noncovalent salt bridge interaction. Polar interactions between the D168, R123, and R155 side chains contribute to the strength of this noncovalent interaction network. The D168 side chain is predicted to be tightly fixed, allowing no other H-bond interaction and contributing to a nonpolar $S_{2/4}$ interaction interface for the P_2 and P_4 groups in boceprevir and telaprevir. The variant K155 disrupts this nonpolar $S_{2/4}$ interface by leaving the negatively charged D168 unbound, and is expected to reduce binding affinity for boceprevir and telaprevir. A stronger effect is expected for telaprevir than for boceprevir, because the telaprevir P_2 cyclopental-proline is larger than the respective boceprevir isopropyl-proline (Supple-

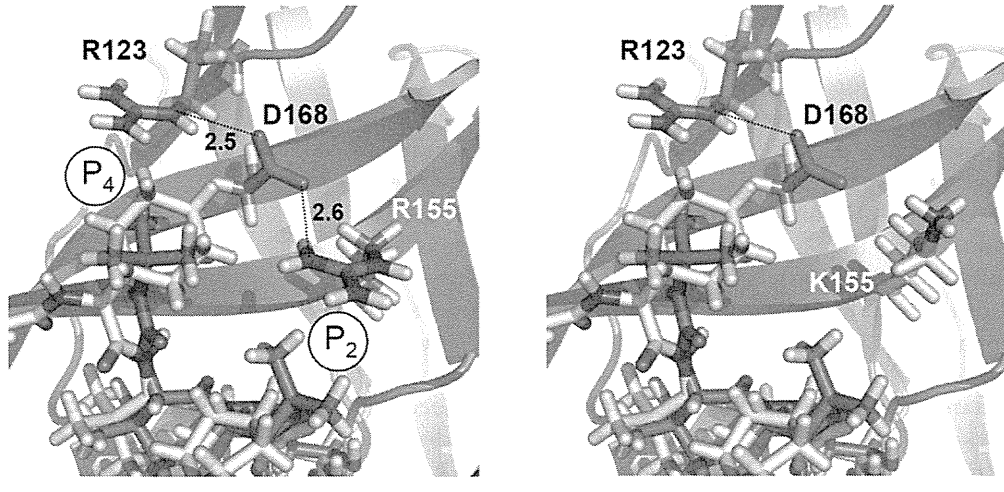

Figure 4. Rotamer analysis of R155K. *Left*: The H-bond pattern (*dotted lines*) in the R155 structure with D168 serving as a nonpolar S_4 contact interface for the ketoamide P_4. *Right*: K155 disables this H-bond pattern, leaving D168 unbound with a polar OH-group at S_4. The K155 side-chain conformation is also predicted to be shifted slightly away from the ketoamide P_2.

mentary Figure 1). Compared with R155, the K155 side chain is also predicted to be shifted slightly away from the ketoamide P_2 group. This shift is likely to cause a loss in van der Waals contacts with the ketoamide P_2 group, and might reduce binding affinity for telaprevir. The effects on polarity and loss of van der Waals contacts suggest that there will be a significant decrease in binding affinity for both ketoamides, but a larger impact on telaprevir.

Two variants were observed at the D168 position that contributes to the S_4 pocket in the protease. D168G is a nonconservative change from aliphatic and polar to a smaller, nonpolar side chain. The wild-type D168 side-chain conformation is oriented toward the ketoamide P_4 group, which is smaller in boceprevir than telaprevir. The extended P_4 group in telaprevir (Figure 1) points away from D168. Thus, the D168 variant is expected to have only minor effects on ketoamide binding. G168 is expected to have effects on the nonpolar $S_{2/4}$ interface similar to K155, because it offers no polar side-chain interaction partner for R155. It is predicted to change the polarity of the $S_{2/4}$ interface and expected to cause a minor decrease in binding affinity for both ketoamides. D168E is a conservative change as both side chains are negatively charged and aliphatic. The side-chain conformation is predicted to be similar to the wild-type D168 and to preserve $S_{2/4}$ interface polarity. D168E is not expected to impact ketoamide binding.

Additional BSVs without expected impact on ketoamide binding. T42A is located at the periphery of the S_1' pocket, and is a nonconservative change from aliphatic, polar, and negative toward aliphatic and hydrophobic. A side-chain OH group is lost in the A42 variant but is preserved in the S42 variant at this position, allowing for a potential H-bond interaction. The wild-type T42 and variant A42 and S42 side-chain conformations are predicted to be similar. No ketoamide H-bond donor or acceptor group is found in close proximity to T42 or S42. Because of its relative distance from the ketoamide, the polarity change in T42A is predicted not to have a significant effect on ketoamide binding. Thus, the T42A/S variants are expected to have no impact on ketoamide binding. F43S is a nonconservative change from aromatic and nonpolar toward aliphatic and polar, and is found at the bottom of S_1'. No change is predicted in the side-chain conformation. There is an additional OH group in the variant, but because there is no ketoamide H-bond donor or acceptor group in close proximity, this by itself is unlikely to have any impact on ketoamide binding. Nevertheless, F43S can cause ketoamide resistance because the F43 aromatic ring directly participates in the formation of the S_1' pocket, which impacts binding of the ketoamide P_1' group.[25] Substitutions at F43 have been shown previously to cause resistance to ketoamide compounds.[25] T160A is a nonconservative change from polar toward nonpolar and hydrophobic, located at a distance from S_4 near the boundary of an extended S_5 pocket. T160 interacts directly with the extended P_4 group in TLL in the PDB structure 2P59, but has no noncovalent interactions with boceprevir in PDB structure 2OC8. The variant A160 side chain is not expected to influence ketoamide binding.

Replication Capacity and Infectious Virus Yield From RNAs Containing BSVs

Those BSVs for which the rotamer analysis suggested a possible impact on ketoamide binding (ie, Q41H, I132V, F154Y, R155K, and D168G) were selected for phenotypic characterization. We similarly tested K136R, for which the rotamer analysis provided no clear predictions. 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 and production of infectious virus determined in RNA-transfected cells.

RNA replication capacity. The replication capacity of H77S.3/GLuc2A RNA mutants with BSV substitutions in NS3 was assessed by measuring GLuc activities in supernatant media collected at 24-h intervals after transfection of synthetic RNA, as described previously.[14] Results were normalized to the activity present at 6 h after transfection, as this represents GLuc expressed directly by the transfected input RNA. The RNA replication capacity of the Q41H variant was similar to the parental H77S.3/

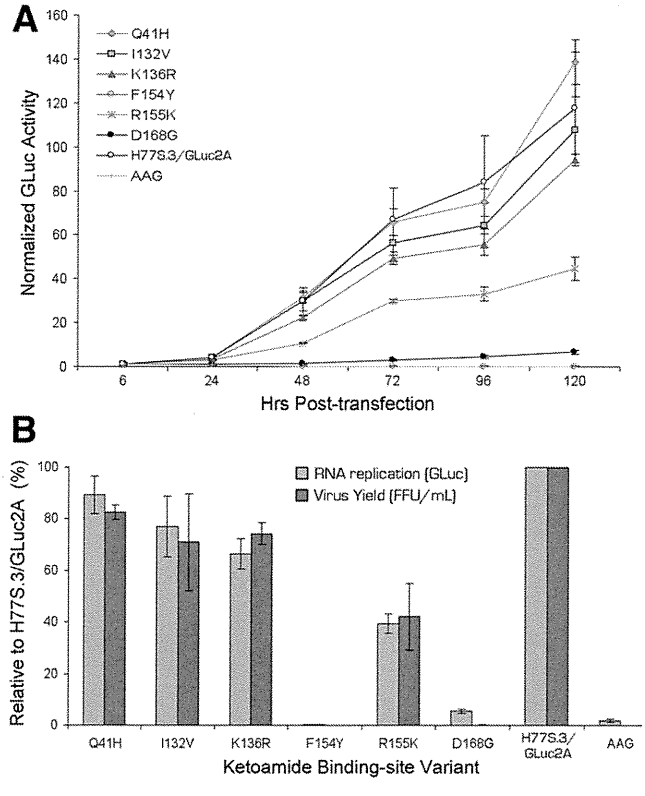

Figure 5. Replication capacity and infectious virus yield of H77S.3 RNAs with BSV substitutions in NS3. (*A*) RNA replication capacity reflected of H77S.3/GLuc2A BSV into Huh7.5 cells. Results are normalized to the 6-h GLuc activity, and represent the mean of triplicate samples. (*B*) Comparison of RNA replication capacity (*lightly shaded bars*) and infectious virus yields (*dark shaded bars*). Both are normalized to those obtained with the relevant parental RNA (100%). Data represent the mean ± SD from at least 3 independent experiments.

Table 2. Predicted and Measured Impact of BSVs on Antiviral Activity of Ketoamides[a]

	IRECS		EC_{50} (nM)		Fold-change in EC_{50}	
Genotype 1a HCV	Boceprevir	TLL	Boceprevir	Telaprevir	Boceprevir	Telaprevir
H77S.3/GLuc2A	–	–	870 (±48)	120 (±10)	1.0	1.0
Q41H	*MiN*[b]	*MiN*	1040 (±60)	410 (±10)	1.2	3.5
T42A	NoI	NoI	–	–	–	–
T42S	NoI	NoI	–	–	–	–
F43S	NoI	NoI	–	–	–	–
I132V	*MiN*	*MiN*	980 (±60)	280 (±40)	1.1	2.4
K136R	NoI	*NP*	440 (±19)	100 (±10)	0.5	0.9
F154Y	*SN*	*SN*	ND	ND	ND	ND
R155K	*MeN*	*SN*	1830 (±240)	1010 (±240)	2.1	8.8
T160A	NoI	NoI	–	–	–	–
D168G	*MiN*	*MeN*	490 (±50)	260 (±30)	0.6	2.2
D168E	NoI	NoI	–	–	–	–

ND, not determined.

[a]The table compares the in silico predictions of the impact of BSVs on ketoamide binding from the Iterative REduction of Conformational Space (IRECS)[20] analysis and corresponding EC_{50} values determined from H77S.3/GLuc2A-transfected cell cultures. Results shown represent the mean ± SD and fold-changes compared with wild-type.

[b]Predicted impact: NoI, no impact; MiN, minor negative impact, change only in a single ketoamide-BSV interaction (H-bond, van der Waals) or in polarity; MeN, moderate negative impact, combined changes in ketoamide-BSV interaction and polarity; SN, strong negative impact, same as for MeN but at particularly close ketoamide-BSV interaction sites; NP, not predictable.

GLuc2A RNA, while I132V and K136R were minimally impaired (Figure 5*A*). In contrast, as shown previously,[14] the replication of R155K was moderately impaired, and D168G was severely handicapped for RNA replication. The maximum RNA replication capacity observed was with Q41H (89% of parental RNA) and the lowest was with D168G (5.5%). The F154Y substitution had a lethal effect on RNA replication in the H77S.3 background, suggesting that viruses containing this sequence variant (and possibly also G168) possess one or more compensating changes at other amino acid positions (see Discussion). None of the BSVs demonstrated enhanced RNA replication capacity compared with the parental H77S.3/GLuc2A RNA (Figure 5*B*).

Infectious virus yield. We also assessed the impact of these BSVs when placed within the background of H77S.3 RNA, which lacks the GLuc2A-coding sequence and produces infectious virus as described previously.[14] Cell culture supernatant fluids were collected 72 h after transfection of the RNA, and subsequently inoculated onto naïve cells, with foci of infected cells detected by immunofluorescence 72 h later. Each of the BSVs tested produced infectious virus yields in the range of that expected from their RNA replication capacity, although the low replication capacity of the D168G variant precluded a careful analysis on production of infectious virus (Figure 5). Thus, none of these amino acid substitutions were documented to directly impair infectious virus assembly or release, as described previously for some resistance-associated NS3 variants.[14] Reductions in the fitness of these particular BSVs are confined primarily to defects in viral RNA replication.

Ketoamide resistance of BSVs. We measured the antiviral activities (50% effective concentration [EC_{50}]) of boceprevir and telaprevir against each of the H77S.3/GLuc2A mutants by determining the concentration of each compound required to cause a 50% reduction in RNA replication (GLuc expression by RNA-transfected cells). Resistance testing could not be performed for the F154Y variant because it was not competent for replication.

Boceprevir demonstrated antiviral activity against each of the BSVs (Table 2). The EC_{50} value against the parental H77S.3/GLuc2A was 870 ± 48 nM. The maximum fold-change in the EC_{50} for any of the BSVs was 2.1 for R155K.[14] Only R155K showed significant, albeit low-level, resistance against boceprevir. Telaprevir showed greater molar activity than boceprevir, with an EC_{50} against the H77S.3/GLuc2A of 120 ± 10 nM. As with boceprevir, telaprevir was active against each of the BSVs, but low- to medium-level resistance was evident with Q41H, I132V, R155K, and D168G (Table 2). The maximum fold-change in the EC_{50} was 8.8 for R155K. As expected from the in silico analysis, the range of fold-changes in EC_{50} was broader for telaprevir than boceprevir. In general, these changes were in good agreement with the impact of these BSVs on ketoamide binding predicted from the rotamer analysis, except for K136R, which was difficult to predict and showed greater antiviral activity than anticipated against both ketoamide compounds (Table 2).

Discussion

Mathematical arguments suggest that every possible drug-resistant viral variant is likely to pre-exist at a low frequency in the replicating viral quasi-species population of the typical HCV-infected patient.[10] Whether this is actually the case, and at what frequency such variants

actually exist, may never be formally demonstrated due to technical difficulties. In this study, we analyzed the natural variation present among ketoamide-neighboring residues in 219 genotype 1a HCV sequences collected from geographically diverse sites and deposited in a public database. We cannot exclude the possibility that some of the BSVs we identified in this set of sequences might represent variants that were present at low frequency in their source patient, or even unrecognized sequencing errors. However, it is likely that they represent true variants present within the dominant quasi-species of the patients from whom these sequences were derived, because multiple BSVs were identified at some residues (T42, V55, and D168) (Supplementary Figure 2), while others (H41, A42, A55, I44, and K155) were present in more than one sequence. We identified BSVs at one or more ketoamide-neighboring residues in 17 of 219 (7.8%) genotype 1a sequences. Importantly, 8 of these variants (Q41H, T42A, T42S, V55I, I132V, K136R, F154Y, and T160A) have not been identified, to our knowledge, in previous in vivo or in vitro studies of ketoamide resistance.

Although they are both linear ketoamide compounds, boceprevir and telaprevir have distinct structural features (Supplementary Figure 1). Telaprevir possesses an extended P_4 capping group and a P_1' cyclopropyl group. The P_2 group is different from the isopropyl-proline in boceprevir, which is smaller than the telaprevir P_2 cyclopentalproline.[23] These structural differences likely contribute to the lower EC_{50} of telaprevir against genotype 1a HCV in cell culture (Table 2), but could also pose a higher risk for telaprevir resistance among BSVs. Rotamer analysis predicted that 2 BSVs (R155K and D168G) would exert a greater negative effect on the binding of telaprevir than boceprevir. The impact of the K136R substitution proved difficult to predict on the basis of rotamer analysis alone, however, and subsequent tests in cell culture demonstrated that it imposes no resistance against either ketoamide (Table 2).

On the other hand, the R155K and D168G substitutions led to a 2- to 4-fold greater increase in the EC_{50} of teleprevir compared with boceprevir, and an almost 9-fold increase in the teleprevir EC_{50}. This was consistent with predictions from the rotamer analysis, which also agreed with previous crystallographic studies.[26] Whether such changes in the EC_{50} result in clinical resistance would be dependent on the potency of a drug and the drug exposure achieved in a typical patient. A 2-fold change might not be clinically relevant, but a 9-fold increase such as that found with R155K is likely to be significant. Importantly, R155K has been associated previously with resistance to both ketoamide and macrocyclic compounds.[23,26]

Three BSVs (Q41H, I132V, and F154Y) were predicted to interact with ketoamide structural features common to both telaprevir and boceprevir, and thus to influence the binding of these compounds equally. Although Q41H was expected to cause only a minor decrease in affinity for both ketoamides, in vitro assays revealed a 3.5-fold increase in the EC_{50} for teleprevir vs 1.2 for boceprevir. Its greater impact on telaprevir potency is likely related to its P_1' cyclopropyl group and induced structural changes in the corresponding S_1' pocket. Such changes were not detectable using rotamer analysis, but probably influenced the binding of telaprevir. A similar difference in the magnitude of the change in EC_{50} was observed for the I132V variant (1.1-fold change for boceprevir vs 2.4-fold for teleprevir). This could result from a loss of van der Waals contacts with the ketoamide P_3 group, which differs slightly in its orientation in the co-crystallized boceprevir and TLL structures (see Figure 1). The F154Y substitution was lethal for RNA replication when placed in the background of the H77S.3 virus, and thus we could not measure antiviral activity against it.

Viral fitness coupled with the degree of resistance conferred by a BSV are likely to be the major determinants driving selection of a variant from within the viral quasi-species during therapy. RNA replication capacity is one measure of the fitness of the virus, and this is dependent on proper processing of the polyprotein by NS3/4A. Ketoamides mimic the natural substrate of the protease at the site of NS3-NS4A scission, and it is likely that the negative influence of BSVs on RNA replication (Figure 5*A*) reflects altered recognition of the polyprotein substrate related to structural changes similar to those leading to drug resistance. Despite this, there is no strong correlation between the degree of PI resistance and the impact on RNA replication.[14] This is reflected in the marked deficit in replication demonstrated by the D168G substitution (Figure 1*A*), which confers only a minimal increase in the teleprevir EC_{50} (Table 2). There is no obvious structural or molecular explanation for this difference, and it is not possible to exclude the possibility that G168 might provide for more robust RNA replication when placed in the context of a different virus sequence.

In contrast, F154Y was lethal for replication in the context of the H77S.3 virus. This substitution occurs at a central position within the ligand-binding site at the bottom of the S_1 pocket. A recent crystallographic study found the F154 aromatic ring to directly contact the substrate P_1 side chain.[27] Thus, the complete loss of RNA replication observed with F154Y could be due to either intrusion of the Y154 side chain into the space normally occupied by the polyprotein substrate, or a polarity shift within S_1 due to the Y154 OH group.

The presence of Y154 in the database suggests that it is capable of functioning in an alternative sequence context. This highlights a limitation of the technical approach we have taken here to study the phenotypic effect of BSVs, as second-site substitutions in the same strain might compensate for defects fitness. In fact, the sequence of the F154Y BSV (EU677253) contains 2 additional substitutions in NS3 that differ from the genotype 1a consensus: R11G and H110R (Supplemen-