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

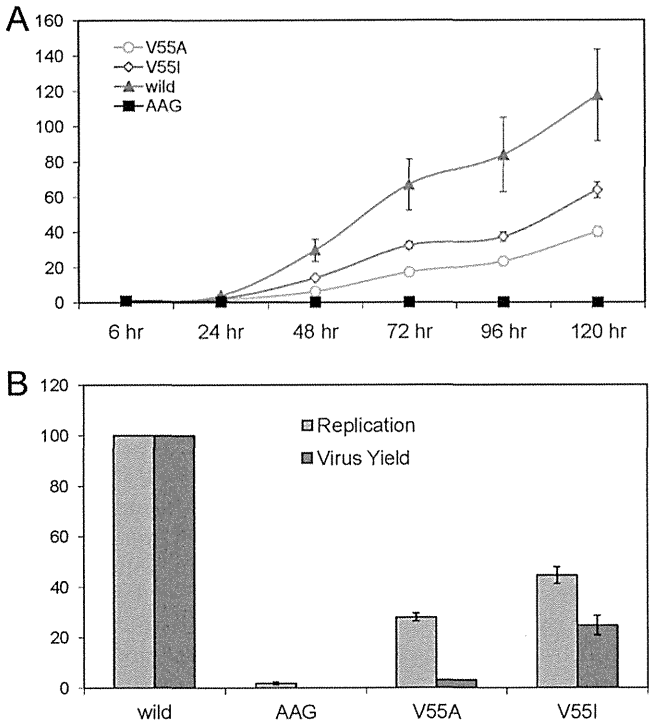

FIG 1 Replication capacity and infectious virus yield of H77S.3 RNAs with V55A and V55I substitutions in NS3. (A) RNA replication capacity reflecting the presence of H77S.3/GLuc2A V55A and V55I variants in Huh7.5 cells. Results are normalized to the 8-h GLuc activity and represent the mean of triplicate 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

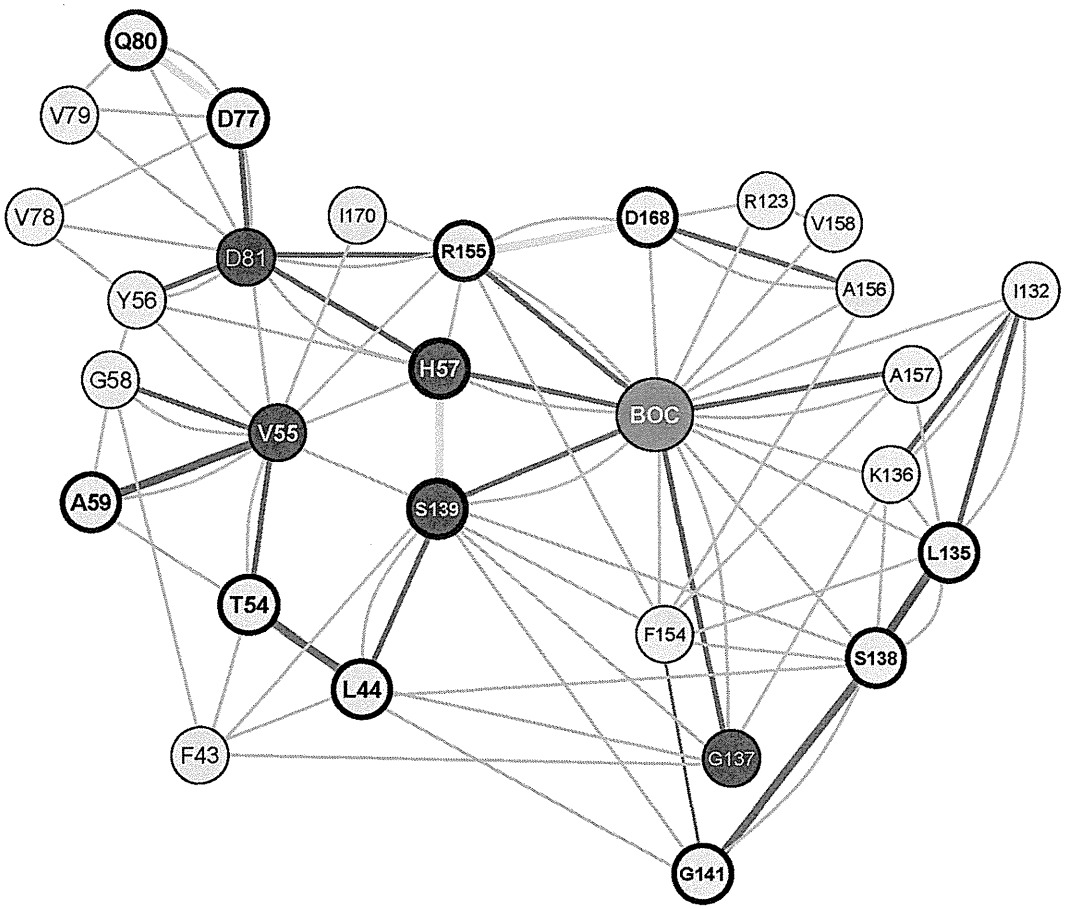

FIG 2 Residue-interaction network of the NS3 protease from PDB structure 2OC8, with nodes and edges representing protease residues and noncovalent interactions, respectively. The ketoamide compound boceprevir is represented by a single green node (BOC). The residue Val55 is given as a red node. Catalytic 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

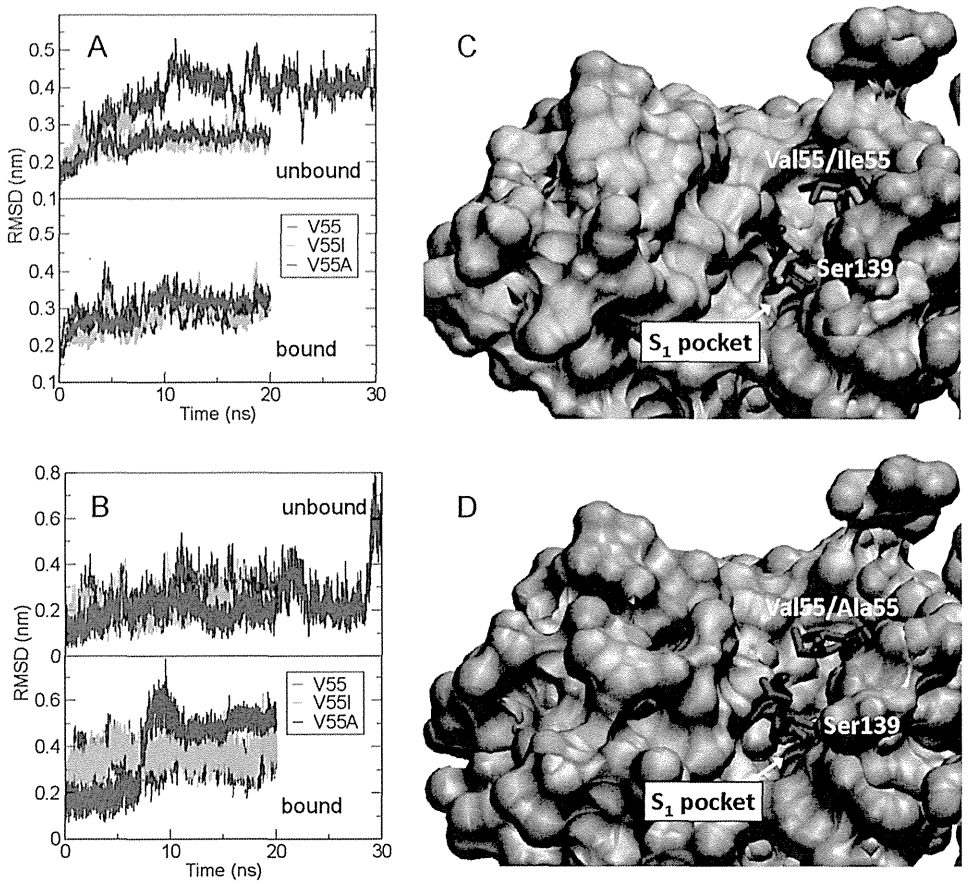

FIG 3 Backbone and side-chain root mean square deviations (RMSDs) and molecular dynamics simulations of the NS3 protease. (A) Backbone RMSDs of "unbound" and "bound" protease wild type and V55A and V55I variants. (B) RMSDs of protease residue His57 in "unbound" and "bound" wild-type and V55A and V55I 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.

JVI
Journals.ASM.org

Base Pairing between Hepatitis C Virus RNA and MicroRNA 122 3′ of Its Seed Sequence Is Essential for Genome Stabilization and Production of Infectious Virus

Tetsuro Shimakami,[a]* Daisuke Yamane,[a] Christoph Welsch,[a]* Lucinda Hensley,[a] Rohit K. Jangra,[b] and Stanley M. Lemon[a]

Division of Infectious Diseases, Department of Medicine, Inflammatory Diseases Institute and Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA,[a] and Department of Microbiology, Mount Sinai School of Medicine, New York, New York, USA[b]

MicroRNA 122 (miR-122) facilitates hepatitis C virus (HCV) replication by recruiting an RNA-induced silencing complex (RISC)-like complex containing argonaute 2 (Ago2) to the 5′ end of the HCV genome, thereby stabilizing the viral RNA. This requires base pairing between the miR-122 "seed sequence" (nucleotides [nt] 2 to 8) and two sequences near the 5′ end of the HCV RNA: S1 (nt 22 to 28) and S2 (nt 38 to 43). However, recent reports suggest that additional base pair interactions occur between HCV RNA and miR-122. We searched 606 sequences from a public database (genotypes 1 to 6) and identified two conserved, putatively single-stranded RNA segments, upstream of S1 (nt 2 and 3) and S2 (nt 30 to 34), with potential for base pairing to miR-122 (nt 15 and 16 and nt 13 to 16, respectively). Mutagenesis and genetic complementation experiments confirmed that HCV nt 2 and 3 pair with nt 15 and 16 of miR-122 bound to S1, while HCV nt 30 to 33 pair with nt 13 to 16 of miR-122 at S2. In genotype 1 and 6 HCV, nt 4 also base pairs with nt 14 of miR-122. These 3′ supplementary base pair interactions of miR-122 are functionally important and are required for Ago2 recruitment to HCV RNA by miR-122, miR-122-mediated stabilization of HCV RNA, and production of infectious virus. However, while complementary mutations at HCV nt 30 and 31 efficiently rescued the activity of a 15C,16C miR-122 mutant targeting S2, similar mutations at nt 2 and 3 failed to rescue Ago2 recruitment at S1. These data add to the current understanding of miR-122 interactions with HCV RNA but indicate that base pairing between miR-122 and the 5′ 43 nt of the HCV genome is more complex than suggested by existing models.

Hepatitis C virus (HCV) is a positive-strand RNA virus classified in the genus *Hepacivirus* of the family *Flaviviridae*. It is strongly hepatotropic and has an unparalleled ability among RNA viruses to establish lifelong persistent infections in the majority of persons it infects (for a review, see reference 18). About 4 million Americans are persistently infected with HCV, resulting in approximately 12,000 deaths annually due to cirrhosis and liver cancer, with particularly high rates of disease among those coinfected with human immunodeficiency virus (HIV) (6). Worldwide, more than 130 million people are infected, with an estimated 350,000 HCV-related deaths each year (23). HCV is thus important from a medical and public health perspective. It is also important as an example of an RNA virus that has achieved a uniquely close evolutionary relationship with its human host, upon which its replication depends. This relationship is exemplified by the novel role played by a particular, liver-specific microRNA (miRNA), miR-122, in the HCV life cycle.

miRNAs are ~22-nucleotide (nt) RNAs that originate from lengthier transcripts and that regulate gene expression, typically posttranscriptionally, by mediating mRNA deadenylation and degradation and/or translational repression (reviewed in reference 8). This requires loading of the "guide strand" of a miRNA duplex into an argonaute (Ago) protein-containing miRNA-induced silencing complex (miRISC) and, typically, the binding of its "seed sequence" (nt 2 to 8 of the miRNA) to the 3′-untranslated region (3′UTR) of target mRNAs. miRNAs are evolutionarily conserved, and as such, they play a controlling role in cellular metabolism by regulating a majority of all genes. At present, miRBASE (http://www.mirbase.org/) lists over 1,400 distinct human miRNAs, each of which targets a number of mRNAs. Of these, miR-122 (hsa-mir-122 or MI000442) is uniquely expressed at high abundance in the adult liver, accounting for over 50% of the mature miRNAs in hepatocytes (2). miR-122 is also expressed in Huh-7 human hepatoma cells, which are widely used to propagate laboratory strains of HCV (2).

Jopling et al. (17) demonstrated that the ability of synthetic HCV RNA to replicate in transfected Huh-7 cells is dependent upon a direct interaction of miR-122 with the 5′UTR segment of the positive-strand RNA. miR-122 binds to two sites, near the 5′ end of the HCV genome (S1 and S2), at which conserved viral sequences are complementary to the seed sequence of miR-122 (nt 2 to 7) (16) (Fig. 1A). Binding of the miR-122 seed sequence to S1 and S2 promotes protein expression from transfected synthetic viral RNA and has been suggested to enhance the efficiency of translation initiated by the HCV internal ribosome entry site (IRES) (14, 15). However, a comparison of the yields of infectious virus produced by cells transfected with HCV RNA with mutations in the miR-122 binding sites versus mutations in the IRES led us to conclude that promotion of translation is insufficient to

Received 28 February 2012 Accepted 13 April 2012
Published ahead of print 24 April 2012
Address correspondence to Stanley M. Lemon, smlemon@med.unc.edu.
* Present address: Tetsuro Shimakami, Department of Gastroenterology, Kanazawa University Graduate School of Medicine, Kanazawa, Japan; Christoph Welsch, J. W. Goethe University Hospital, Department of Internal Medicine I, Frankfurt am Main, Germany.
T.S. and D.Y. contributed equally to this article.

doi:10.1128/JVI.00513-12

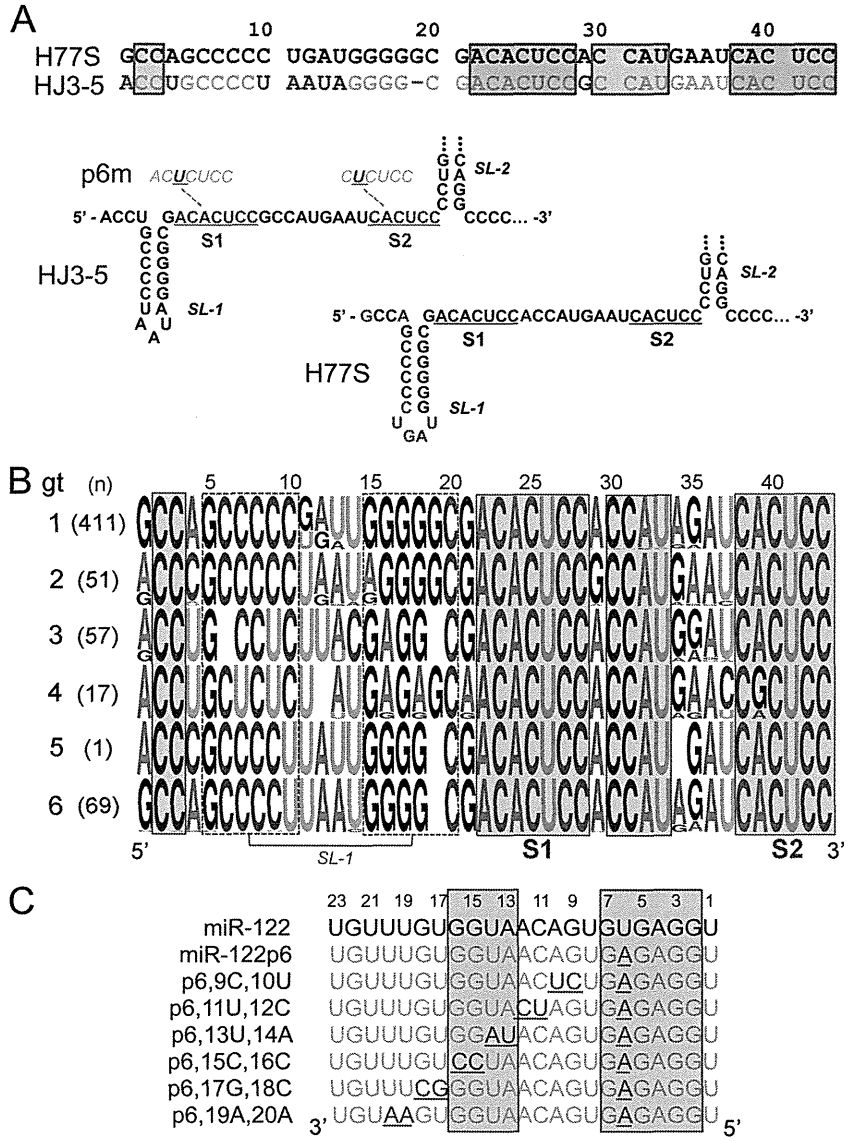

FIG 1 Conserved complementary sequences near the 5′ end of the HCV genome suggest supplementary binding of miR-122 to HCV RNA 3′ of the miR-122 seed sequences at both S1 and S2. (A) (Top) Alignment of the terminal 5′UTR sequences in H77S (genotype 1a) and HJ3-5 (genotype 2a; derived from JFH-1) viruses. Conserved sequences known to base pair with the seed sequence of miR-122 are highlighted in red boxes. Additional conserved sequences complementary to miR-122 are highlighted in green. Nucleotides in HJ3-5 that differ from those in H77S are shown in red font. (Bottom) Putative secondary structure of the RNA segments, showing stem-loop 1 (SL-1) and the base of stem-loop 2 (SL-2). The S1 and S2 seed match sequences are highlighted in red in the HJ3-5 sequence, with the single base changes in the p6m mutants of S1 and S2 shown above. Bases that differ between the genotype 1a and 2a sequences are highlighted in red in the H77S structure. (B) Logo depiction of sequence conservation in the 5′UTRs of different HCV genotypes. The number and height of each character reflect the diversity in bases and conservation of sequence at each position. Highly conserved sequences are highlighted in red and green as in panel A. Base-paired segments in stem-loop SL-1 are shown in open boxes. The number of sequences of each genotype studied is shown to the left. (C) Guide-strand sequences of miR-122 mutants and related mutants. Point mutations are shown in red. Passenger-strand sequences (Table 1) were modified to maintain the melting temperature of each duplex between 54.3 and 54.9°C. Bases are numbered in a 5′-to-3′ fashion at the top but shown in a 3′-to-5′ manner to facilitate visualization of base pairing with positive-strand HCV RNA. Nucleotides involved in base pair interactions are shaded in red and green as in panels A and B.

fully explain the requirement for miR-122 (15). These data suggest that miR-122 promotes HCV replication by a mechanism other than translation, stabilizing the viral RNA or possibly promoting the initiation of new viral RNA synthesis (15, 28). Consistent with this prediction, we recently confirmed that binding of miR-122 results in the recruitment of an Ago2-containing RISC-like complex that physically stabilizes the viral RNA genome (26). This miR-122–Ago2 complex slows the decay of positive-strand HCV RNA both in RNA-transfected cells and in cells persistently infected with HCV, most likely by protecting the RNA from 5′-exonuclease attack (26).

In addition to base pairing involving the seed sequence, the recognition of cellular mRNA targets by miRNAs may also entail supplementary or accessory interactions of bases 3′ of the seed sequence (1, 12). Such additional base pairing optimally involves nt 13 to 16 and spares nt 9 to 12 of the miRNA, presumably be-

cause such base pairing does not require helix formation throughout the length of the guide strand (as occurs with perfectly complementary small interfering RNA [siRNA]) and thus may not perturb its association with Ago in the RISC (1, 12). Several recent reports provide evidence that the interaction of miR-122 with HCV RNA similarly involves such 3′ supplementary interactions (20, 22, 24). Although details differ between these studies, these interactions appear to contribute to the ability of miR-122 to promote the amplification of HCV RNA in transfected cells or protein expression from reporter RNAs containing the HCV IRES in their 5′UTR. Here we confirm and extend these observations, using a genetic approach to demonstrate 3′ supplementary base pair interactions between miR-122 and the 5′ termini of both genotype 1a and genotype 2a viral RNAs. We show that 3′ supplementary base pairing contributes to the recruitment of Ago2 and the stabilization of HCV RNA by miR-122 and that it is also required for optimal production of infectious virus in cell culture.

MATERIALS AND METHODS

Cells. Huh-7, Huh-7.5, and FT3-7 cells (a clonal derivative of Huh-7 cells) were maintained as described previously (15, 25). Murine embryo fibroblasts (MEFs) were the kind gift of Alexander Tarakhovsky, Rockefeller University, and were maintained in cell culture as described previously (21).

Plasmids. pHJ3-5 (19), the related S1, S2, and S1-S2 p6m mutants (15), and pH77S/GLuc2A-AAG (25, 30) have been described previously. The *Gaussia princeps* luciferase (GLuc) coding sequence, followed by the foot-and-mouth disease virus 2A sequence, was inserted between p7 and NS2 in pHJ3-5 or pH77S.3 and the related S1, S2, and S1-S2 p6m, S1p3m, and S1p23m miR-122 binding site mutants, using the strategy adopted previously for H77S (25). Single- and multiple-base substitutions were created within pHJ3-5 and pHJ3-5/GLuc2A by site-directed mutagenesis of the segment spanning the EcoRI and AgeI sites. Base substitutions were similarly introduced into pH77S.3/GLuc2A and pH77S/GLuc2A-AAG by site-directed mutagenesis of the sequence spanning the NotI and AgeI sites. Base changes and the integrity of the surrounding sequences were confirmed by DNA sequencing.

RNA transcription. RNA transcripts were synthesized *in vitro* as described previously (25).

RNA oligonucleotides. RNA oligonucleotides have been described previously (15) and were synthesized by Dharmacon. Additional mutant miRNAs used in this study are shown in Table 1. All miRNAs were transfected as duplexes (15).

miRNA supplementation and HCV replication. FT3-7 cells seeded previously into a 6-well culture plate were transfected with a miRNA duplex (50 nM) by use of Lipofectamine 2000 (Invitrogen) as recommended by the manufacturer. At 24 h, the cells were retransfected with replication-competent *in vitro*-transcribed HCV RNA (1.25 μg/well) for 6 h, using the TransIT mRNA transfection reagent (Mirus Bio) according to the manufacturer's suggested protocol. The cells were retransfected with miRNA 24 h later and refed fresh medium every 24 h thereafter. Supernatant fluid samples were collected for virus titration or GLuc assay at 24-h intervals, unless noted otherwise. Total cellular RNA was extracted for Northern blots.

miRNA supplementation and HCV RNA stability and translation. Synthetic HCV RNA containing a replication-lethal mutation in the NS5B coding region (20 μg) and duplex miRNA (1 μM) were mixed with 1×10^7 wild-type (wt) MEF cells in a 4-mm cuvette and pulsed once at 400 V, 250 μF, and infinite Ω in a Gene Pulser Xcell Total system (Bio-Rad). Synthetic capped and polyadenylated RNA encoding *Cypridina* luciferase (CLuc) was cotransfected to monitor transfection efficiency. The cells were plated into three wells of a six-well culture plate, and cell culture supernatant fluid was collected for GLuc assay and total cellular RNA extracted for Northern blotting at 3-h intervals.

TABLE 1 miR-122 mutants used in this study

miR-122 mutant	Guide (5′-3′) and passenger-strand (3′-5′) RNA sequences[a]
miR-122	UGGAGUGUGACAAUGGUGUUUGU AUAAAUCACACUAUUACCGCAAA
miR-122p6	UGGAGAGUGACAAUGGUGUUUGU AUAAAUCACACUAUUACCGCAAA
miR-122p6,9C,10U	UGGAGAGUCUCAAUGGUGUUUGU AUAAAUCACAGAAUUACCGCAAA
miR-122p6,11U,12C	UGGAGAGUGAUCAUGGUGUUUGU AUAAAUCACACUAUUACCGCAAA
miR-122p6,13U,14A	UGGAGAGUGACAUAGGUGUUUGU AUAAAUCACACUAUAUCCGCAAA
miR-122p6,15C,16C	UGGAGAGUGACAAUCCUGUUUGU AUAAAUCACACUAUUAGGGCAAA
miR-122p6,17G,18C	UGGAGAGUGACAAUGGGCUUUGU AUAAAUCACACUAUUACCUGAAA
miR-122p6,19A,20A	UGGAGAGUGACAAUGGUGAAUGU AUAAAUCACACUAUUACCGCUUA
miR-122m15C,16C	UGGAGAGUGACAAUCCUGUUUGU AUAAAUCACACUAUUAGGGCAAA

[a] The top sequence for each mutant represents the guide strand, and the lower sequence represents the passenger strand. Underlined bases have been mutated from the wild-type sequence.

Infectious virus titration. Huh-7.5 cells were transfected with duplex miR122p6 (50 nM) and then seeded into 48-well plates at a density of 5×10^4 cells/well and maintained in a 37°C 5% CO_2 environment. At 24 h, the cells were inoculated with serial dilutions of virus-containing medium (100 μl). Cells were fed 300 μl fresh medium 24 h later. Following 48 h of additional incubation, cells were fixed, labeled with HCV core-specific antibody, and examined under an inverted UV fluorescence microscope as described previously (25, 30). Clusters of infected cells identified by specific staining for core antigen were considered to constitute a single infectious focus-forming unit (FFU); virus titers are reported as FFU/ml.

Ago2-RNA immunoprecipitation (IP). MEFs were electroporated with 10 μg HCV RNA and 1 μM duplex miRNAs. Six hours later, cells were harvested in lysis buffer (150 mM KCl, 25 mM Tris-HCl, pH 7.4, 5 mM EDTA, 1% Triton X-100, 5 mM dithiothreitol [DTT], Complete protease inhibitor cocktail [Roche], 100 U/ml RNaseOUT [Invitrogen]). Lysates were centrifuged for 30 min at 17,000 × *g* at 4°C, filtered through a 0.45-μm syringe filter, and incubated with anti-Ago2 monoclonal antibody (MAb) (2D4; Wako Chemicals) or isotype control IgG at 4°C for 2 h, followed by the addition of 30 μl of protein G Sepharose (GE Healthcare) for 1 h. The Sepharose beads were washed 3 times in lysis buffer, and RNA was extracted using an RNeasy minikit (Qiagen). HCV RNA associated with Ago2 protein was detected by reverse transcription-PCR (RT-PCR), using a SuperScript One-Step RT-PCR kit with Platinum *Taq* DNA polymerase (Invitrogen) and primers specific for the HCV core sequence as described previously (25).

Luciferase assay. Cell culture supernatant fluids were collected at 24-h intervals following RNA transfection and replaced completely with fresh medium. Secreted GLuc and CLuc activities were measured as described previously (25).

Northern blotting. Northern blotting for HCV RNA and actin mRNA (loading control) and quantitation of these results were carried out as described previously (15, 25).

Genotype-specific HCV 5′UTR sequence logos. HCV 5′UTR sequences of at least 300 bases were retrieved from the European HCV database (euHCVdb [http://euhcvdb.ibcp.fr/euHCVdb/]) (4). Genotype-specific (27) sequence alignments were computed using ClustalW (3) and MUSCLE (7), with minor manual modifications in the SeaView alignment editor (11). Sequence logos were computed using WebLogo (http://weblogo.berkeley.edu/logo.cgi) (5).

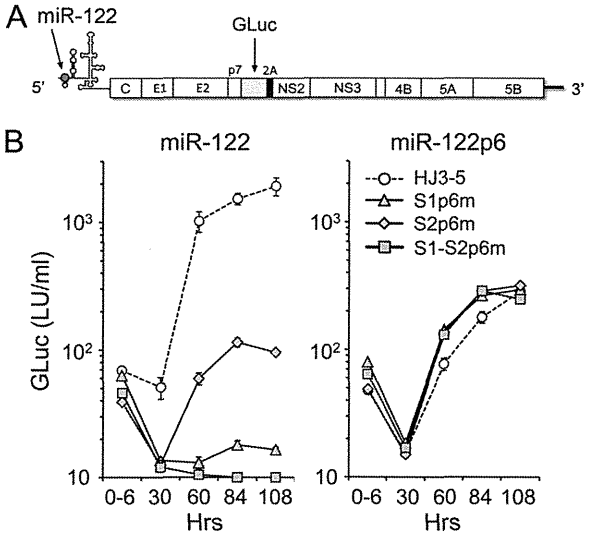

FIG 2 Replication of HJ3-5 RNAs with or without point mutations in the S1 and S2 seed match sites. (A) Organization of synthetic GLuc-expressing HCV genomic RNAs. The GLuc sequence was inserted downstream of p7, followed by the 2A sequence of foot-and-mouth disease virus. (B) GLuc activities in supernatant fluids of cultures transfected with HJ3-5 or related S1 and S2 mutant RNAs and either wt miR-122 (left) or the miR122p6 mutant (right). FT3-7 cells were transfected with the oligoribonucleotides 24 h before and after transfection of the viral RNAs. Supernatant fluids were collected to assay GLuc secreted into the medium between 0 and 6 h following HCV RNA transfection and over successive 24-h periods ending at 30, 60, 84, and 108 h posttransfection. Results shown represent the mean ratios of GLuc activities in replicate cultures ± ranges and are representative of multiple independent experiments. LU, light units.

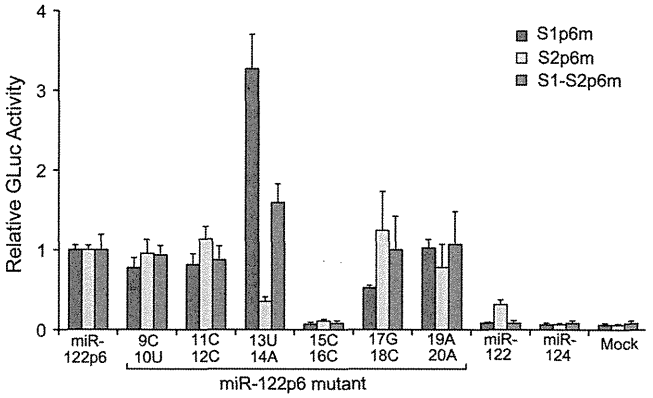

FIG 3 Capacity of various miR-122 mutants to support HCV RNA replication. HJ3-5 RNAs bearing either single or double S1 and S2 p6m mutations, as indicated, were cotransfected into cells with the indicated wt or mutant duplex miR-122s, using the same protocol as that described in the legend to Fig. 2B. Results shown represent the mean ratios of GLuc activities in triplicate cultures at 84 h ± standard deviations (SD), normalized for each viral RNA to that obtained with miR-122p6. The data shown are representative of multiple independent experiments. Mock, no oligonucleotide.

RESULTS

miR-122 bases involved in productive interactions with HCV RNA. A conserved CCAU sequence (HCV nt 30 to 33) upstream of the S2 site in HCV RNA (Fig. 1A, top panel) is complementary to nt 13 to 16 of miR-122 and is thus available for base pairing with the miRNA 3′ of its seed sequence (miR-122 nt 2 to 7) (12). In addition, a conserved CC sequence (HCV nt 2 and 3) upstream of stem-loop 1 represents a similar potential site of 3′ supplementary base pairing at S1. An analysis of over 600 individual viral sequences deposited in the European HCV database (euHCVdb) indicated that these miR-122-complementary bases are highly conserved across all 6 HCV genotypes (Fig. 1B). As indicated above, Machlin et al. (20) recently showed that 3′ supplementary interactions involving nt 15 and 16 of miR-122 and nt 2 and 3 and nt 30 and 31 of the viral RNA contribute to the ability of miR-122 to promote the accumulation of viral RNA following its transfection into permissive cells (20). To further characterize supplementary base pairing of miR-122 outside its seed sequence to HCV RNA, we synthesized a panel of miR-122 mutants in which successive pairs of bases, starting at nt 9 and 10 and going through nt 19 and 20, were replaced with their Watson-Crick complement (Fig. 1C) (the analysis of miR-122 mutants with base substitutions at nt 21 to 23 will be reported elsewhere). Each of these mutated miR-122s had a U-to-A substitution at nt 6 ("p6" mutation). We tested the ability of the mutants, transfected as duplex miRNAs into HCV-permissive FT3-7 cells, to promote the replication of HCV RNA carrying complementary substitutions ("p6m") (Fig. 1A, bottom panel) within the S1, S2, or both S1 and S2 seed sequence binding sites (15). HCV RNAs with a p6m base substitution in either S1 or S2 are severely handicapped in replication, especially when the mutation is at the S1 site, but can be rescued by transfection of the complementary miR-122 p6 mutant (15). This strategy functionally isolates the S1 and S2 miR-122 binding sites, allowing requirements for miR-122 sequence to be determined individually at each seed match site, and is similar to the strategy we used previously to dissect the contributions of the two sites to miR-122-mediated increases in viral translation versus replication (15). Where necessary, the passenger-strand RNA in the duplex was modified to maintain a melting temperature for each duplex that was similar to that of the parental miR-122p6 duplex (Table 1).

The genome-length HCV RNA used in this series of experiments was HJ3-5 (19), a chimera that contains the genotype 2a 5′UTR sequence rather than the genotype 1a H77 5′UTR sequence studied by Machlin et al. (20) (see Fig. 1A for a comparison of the relevant sequences). It was modified by the insertion of the GLuc (*Gaussia princeps* luciferase) sequence within the polyprotein, which allows replication to be monitored by measuring GLuc activity secreted into supernatant fluids of transfected cell cultures (Fig. 2A). As expected (15), viral RNAs with p6m mutations in S1 or S2 produced much less GLuc than RNAs containing the wt seed match sequences when cotransfected into cells with wt miR-122 (Fig. 2B, left panel). The GLuc expression pattern suggested a lack of replication of the double S1-S2p6m mutant. In contrast, the GLuc activity produced by each of these RNAs was substantially increased when the RNAs were cotransfected with miR-122p6 (Fig. 2B, right panel), with the S1-S2p6m mutant replicating with an efficiency similar to that of the other RNAs.

Each of the modified miR-122p6 duplexes (Fig. 1C) was cotransfected into cells with either the S1p6m, S2p6m, or S1-S2p6m HJ3-5 HCV RNA, and GLuc activity secreted into the culture supernatant fluids was monitored over the ensuing 84 h as a measure of genome amplification. The results are presented in Fig. 3, which shows the GLuc activity at 84 h posttransfection, normalized for each of the viral RNAs to that produced by cotransfection with miR-122p6 (which lacks base substitutions outside the seed

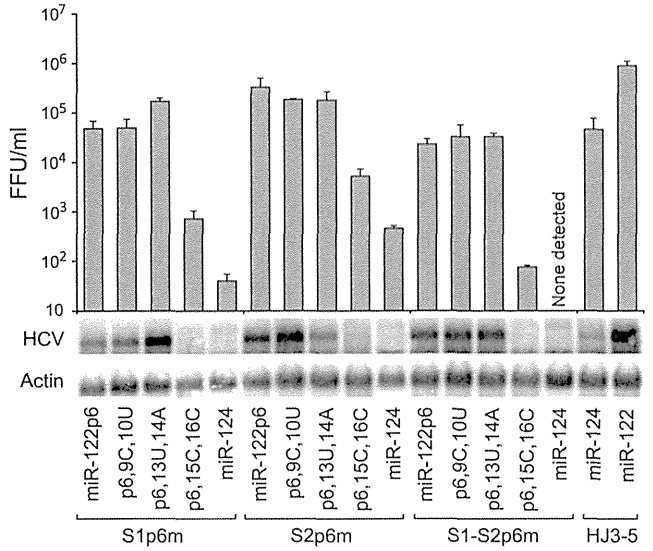

FIG 4 miR-122 nt 13 to 16 are essential for support of HCV RNA replication and production of infectious virus in cell culture. (Top) FT3-7 cells were transfected with HJ3-5 RNA and related p6m mutants 24 h after transfection with the indicated duplex oligoribonucleotides and then retransfected 24 h later with the same oligoribonucleotides. After an additional 48 h of incubation, supernatant fluids were harvested for titration of HCV infectivity in a focus-forming assay using Huh-7.5 cells transfected with miR-122p6 24 h prior to inoculation. (Bottom) RNA extracted from the HCV RNA-transfected cells was subjected to Northern analysis for HCV RNA. β-Actin mRNA served as a loading control.

sequence). As additional controls, the HCV RNA-transfected cells were also supplemented with wt miR-122 or miR-124, an unrelated brain-specific miRNA. miR-122p6 mutants with base substitutions in nt 9 to 12 and 17 to 20 functioned as well as miR-122p6 in promoting genome amplification, indicating that these bases are not important for the interaction with HCV RNA. However, very different results (either decreased or increased GLuc expression) were obtained with miR-122p6 mutants with base substitutions at nt 15 and 16 or nt 13 and 14 (Fig. 3). miR-122p6,15C,16C was markedly impaired in the ability to rescue the amplification of HCV RNAs with p6m mutations at either the S1 or S2 site (Fig. 3). In contrast, while miR-122p6,13U,14A had a significantly reduced capacity to rescue amplification of the S2p6m mutant, it demonstrated 3-fold greater activity than miR-122p6 in promoting replication of the S1p6m RNA and was able to enhance replication of the double S1-S2p6m mutant (Fig. 3). Northern blotting of RNAs extracted from cells transfected with HJ3-5 virus RNAs containing only the miR-122 binding site mutations (no GLuc2A insertion) confirmed these results, as well as the absence of any effect of substitutions at miR-122 nt 9 and 10 (Fig. 4, bottom panels).

The marked decrease we observed in the ability of the miR-122p6,15C,16C mutant to promote amplification of HCV RNA with a genotype 2a 5′UTR confirms the loss of activity of a similar miR-122 mutant against genotype 1a RNA reported by Machlin et al. (20) and is consistent with 3′ supplementary base pair interactions between miR-122 and bases upstream of both the S1 (nt 2 and 3) and S2 (nt 30 and 31) sites in the viral RNA (as suggested in Fig. 1). However, the results we obtained with miR-122 mutated at nt 13 and 14 were quite different from those described previously for a genotype 1 virus. Machlin et al. (20) reported an ~70% decrease in the ability of a similar miR-122 mutant to promote HCV RNA accumulation when bound at the S1 seed match site, while we observed a >3-fold increase in activity (Fig. 3). We show below that these differences, which are large and very significant for miRNA activities, reflect genotype-specific variation in the 5′UTR sequences used in the two studies.

Base substitutions in miR-122 and promotion of infectious virus yield. Since prior studies have not determined the relevance of 3′ supplementary miR-122 mutations to the production of infectious virus, we measured the yields of infectious virus produced from cells cotransfected with the HJ3-5 S1p6m, S2p6m, or S1-S2p6m RNA and key miR-122p6 mutants. Although the dynamic range of the FFU assay used for titration of infectious virus is far greater (extending over >4 orders of magnitude) than that of Northern blotting for HCV RNA, the results generally mirrored each other (Fig. 4). There was no detectable virus produced by the double S1-S2 mutant RNA in the absence of cotransfected miR-122p6. Virus yields from the single S1 and S2 mutants were also very low, but production of infectious virus from each of these HCV RNAs was boosted 100- to 1,000-fold by cotransfection with miR-122p6. In contrast, miR-122p6,15C,16C was severely impaired in the ability to support production of infectious virus by each of the three HJ3-5 mutants. Nonetheless, it did cause an approximately 10-fold increase in infectious virus yield compared with miR-124, indicating that 3′ supplementary base substitutions involving nt 15 and 16 of miR-122 are not absolutely essential for the interaction with HCV RNA and for miR-122 boosting of infectious virus production. The base substitutions at nt 13 and 14 had less impact on infectious virus yield. However, cotransfection of miR122p6,13U,14A with the S1 mutant RNA resulted in a significant increase in virus yield compared with that with miR122p6 and in a slight reduction of yield compared to that with the S2 mutant. These results are consistent with the GLuc results obtained in the transient-transfection experiments shown in Fig. 3. Although disruption of miR-122 binding at either the S1 or S2 binding site has a strong negative effect on viral RNA amplification and infectious virus yields, the results shown in Fig. 4 confirm a greater dependency on the S1 versus S2 seed match site for virus replication, as we suggested previously (15).

Complementary mutations at HCV nt 2 and 3 and nt 30 to 33 rescue miR-122 mutants with substitutions at nt 13 to 16. The strong negative impact of cytosine-for-guanine substitutions at nt 15 and 16 of miR-122 in the experiments described above is consistent with miR-122 base pairing to the conserved cytosines at positions 2 and 3 and positions 30 and 31 in HCV RNA (Fig. 1). This was confirmed by Machlin et al. (20), who demonstrated rescue of miR-122 promotion of viral RNA accumulation when mutations complementing base substitutions at nt 15 and 16 of miR-122 were constructed within the viral RNA. However, the negative impact of the 13U and 14A substitutions observed with miR-122 interactions at the S2 site (Fig. 3) suggested that the supplementary base pairing upstream of S2 extends in a 3′ direction on the viral RNA, to include nt 32 and 33. To confirm this, we constructed HJ3-5/GLuc2A-S2p6m RNAs with base substitutions at nt 32 and 33 that were complementary to those in miR-122p6,13U,14A. We also constructed viral RNAs with guanine substitutions at nt 2 and 3 and nt 31 and 32 to complement the base changes in the miR-122p6,15C,16C mutant. The replication of each of these viral RNA mutants was impaired compared with

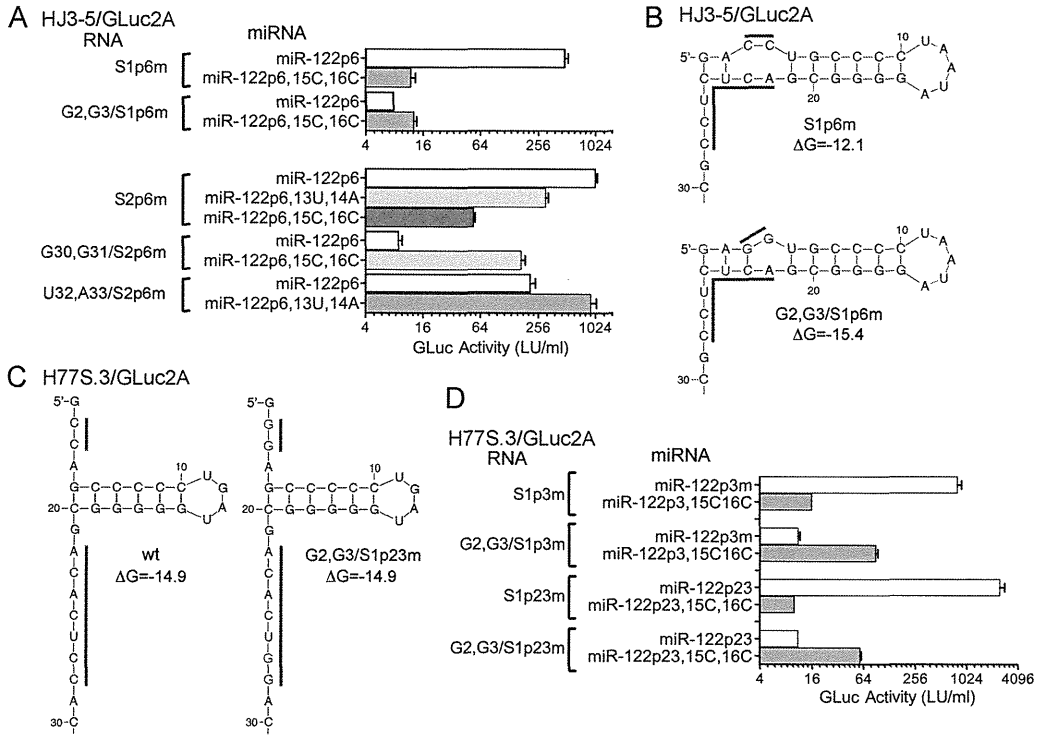

FIG 5 Functional rescue of miR-122 mutants by complementary substitutions in HCV RNA. (A) Complementary substitutions at nt 2 and 3 and nt 30 to 33 of HCV RNA rescue promotion of its amplification by miR-122p6 mutants with substitutions at nt 13 and 14 and/or nt 15 and 16. Replication-competent HJ3-5/GLuc2A RNAs with p6m mutations in S1 (top) or S2 (bottom) and additional base substitutions at nt 2 and 3, 30 and 31, or 32 and 33 were transfected into FT3-7 cells with the indicated oligoribonucleotides, as described in the legend to Fig. 2. Results shown represent mean GLuc activities in supernatant fluids 96 h following transfection ± ranges. Minimal but statistically significant rescue was observed with S1p6m RNA, while rescue was relatively efficient with S2p6m RNA. (B) mfold prediction of secondary structure at the 5′ end of the HJ3-5/GLuc2A-S1p6m transcripts with and without complementary G2 and G3 substitutions. Note that the first guanine base is generated from the T7 promoter sequence in the plasmid, whereas the first nucleotide of H77S is naturally guanine (see Fig. 1). (C) mfold predictions of secondary RNA structures of H77S.3/GLuc2A and H77S.3/GLuc2A-G2,G3-S1p23m. (D) Ability of complementary mutations in H77S.3/GLuc2A-S1p3m and -S1-p23m RNAs to rescue activity of cognate miR-122 mutants with 15C and 16C substitutions.

that of the cognate HJ3-5/GLuc2A-S1p6m or -S2p6m parent (as determined by GLuc expression in transient-transfection assays) (Fig. 5A), with the G2,G3 and G30,G31 mutants affected more than the U32,A33 mutant. With each mutant, the reduction in replication capacity was rescued at least partially by cotransfection with miR-122p6,15C,16C (G2,G3 and G30,G31 mutants) or miR-122p6,13U,14A (U32,A33 mutant). These results provide genetic evidence for the suspected supplementary base pairing, including its extension to nt 32 and 33 as suggested in Fig. 1B.

While we observed a reproducible, ~2-fold enhancement in GLuc expression from HJ3-5/GLuc2A-S1p6m,G2,G3 when it was cotransfected with its cognate miR-122p6,15C,16C mutant, its replication remained very inefficient. One possible explanation for this is an aberrant RNA secondary structure when guanines are inserted in lieu of the cytosines at positions 2 and 3 in the genotype 2a 5′UTR. mfold predictions (http://mfold.rna.albany.edu) of secondary structure in the HJ3-5/GLuc2a-S1p6m,G2,G3 mutant RNA suggest that it forms a stable stem-loop ($\Delta G = -15.4$ kcal) at its extreme 5′ end (Fig. 5B). We considered this a likely explanation for the lack of complete rescue of replication with the complementary miR-122 mutant, since Machlin et al. (20) did not report any difficulty in rescuing a similar mutant genotype 1a RNA. To confirm this, we constructed two G2,G3 mutants constructed in the background of H77S.3/GLuc2A, in which the viral sequence is entirely genotype 1a. In addition to the G2,G3 substitutions, we made base changes at position 3 or both positions 2 and 3 in the S1 seed match site (p3m and p23m mutants). None of these mutations alter the predicted secondary structure at the 5′ end of the genotype 1a genome (Fig. 5C). Despite this, cotransfection of the complementary miR-122p3,15C,16C and miR-122p23,15C,16C mutants with these HCV mutants only partially rescued the large defect in replication capacity imposed by the G2,G3 substitutions (Fig. 5D). We concluded that while these bases are engaged in supplementary 3′ base pairing with miR-122, there is an additional requirement for cytosines at positions 2 and 3. One possibility is that these bases are involved in recognition of the 5′ end of the genome by the HCV replicase (9).

3′ supplementary base pair interactions are required for miR-122 stabilization of HCV RNA. We have shown recently that binding of a miR-122–Ago2 complex stabilizes HCV RNA and results in decreased rates of decay of both synthetic viral RNA transfected into cells and replicating viral RNA in infected cells (26). To determine whether 3′ supplementary base pair interactions of miR-122 contribute to this stabilizing action, we monitored RNA decay in cells cotransfected with a replication-incompetent HCV RNA and various miR-122 mutants. We used MEFs for these experiments because they do not express miR-122. This ensures that results are not confounded by expression of endogenous miR-122. Cells were transfected with genotype 1a H77S RNA modified to express GLuc and containing a replication-lethal mu-

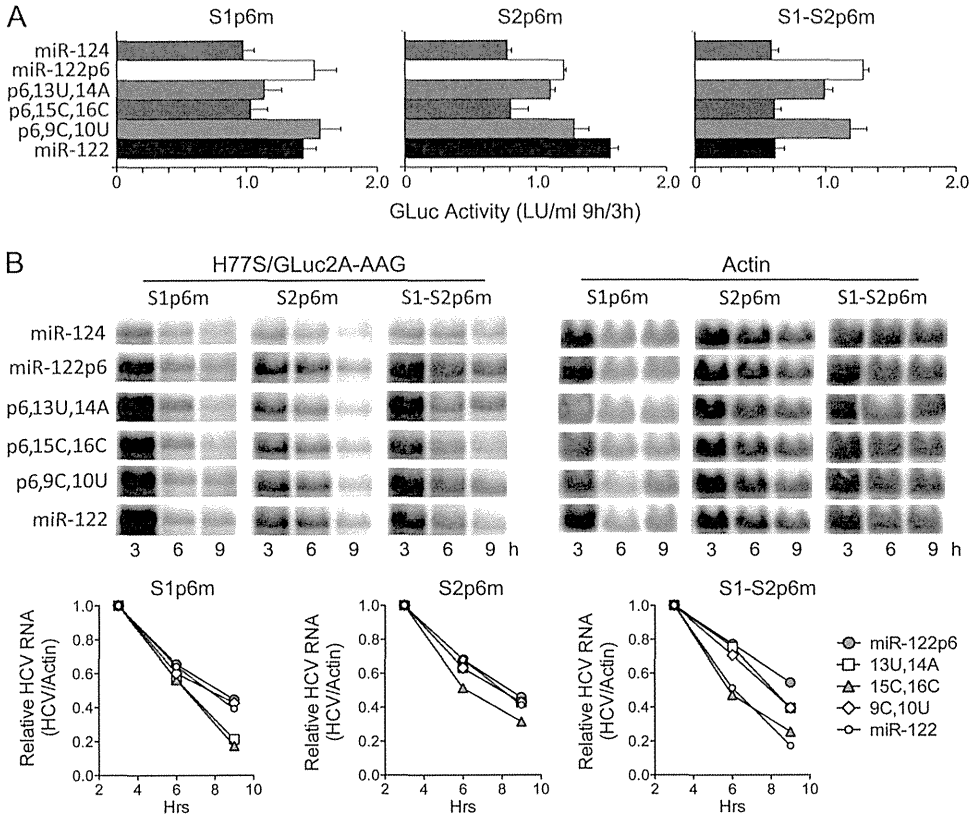

FIG 6 3′ supplementary interactions involving nt 13 to 16 contribute to stabilization of HCV RNA by miR-122. (A) GLuc expression from replication-incompetent, genotype 1a H77S/GLuc2A-AAG and related p6m S1 and S2 mutant RNAs that were coelectroporated into MEFs with the indicated miR-122 duplexes. Results shown represent the means ± ranges of the ratios of GLuc activities in supernatant fluids from duplicate transfected cultures at 9 h versus 3 h and are representative of replicate experiments. (B) Northern analysis of HCV RNA and β-actin mRNA (loading control) in RNAs extracted from cells 3, 6, and 9 h after electroporation of the indicated viral RNAs and duplex miRNAs. Quantitation of HCV RNA (HCV/actin mRNA) in representative Northern blots is shown below, normalized to that present at 3 h (arbitrarily set to 1.0), to compare the stability of the RNA cotransfected with different miR-122 mutants.

tation within the NS5B coding sequence (H77S/GLuc2A-AAG). As we have shown previously (25), GLuc expression from this RNA results exclusively from translation of the input RNA, peaking at around 3 h posttransfection. Thus, the ratio of GLuc activity in culture supernatant fluids at 9 h versus 3 h posttransfection reflects the rate of decay of the transfected RNA (25). p6m mutations were placed within the S1 or S2 seed match sites, as described above (Fig. 1A). As expected, the ratio of GLuc activity expressed by cells transfected with the H77S/GLuc2A-AAG/S1-S2p6m double mutant at 9 h versus 3 h was enhanced when the mutant was cotransfected with miR-122p6 but not miR-122 or the control, brain-specific miR-124 (Fig. 6A, right panel). miR-122p6,9C,10U was as capable as miR-122p6 at enhancing the level of GLuc expression at 9 h versus 3 h, but miR-122p6,15C,16C was no more active than miR-124 or miR-122. These results are consistent with our earlier observations that base pairing involving nt 15 and 16 of miR-122 is required for efficient replication of HJ3-5 RNA (Fig. 3 to 5). Stabilization of the HCV RNA was confirmed by Northern blots of total cellular RNA extracted 3, 6, and 9 h after transfection. These results showed greater abundance of the viral RNA at 9 h when it was cotransfected with miR-122p6 or miR-122p6,9C,10U than when it was cotransfected with miR-122p6,15C,16C, miR-122, or miR-124 (Fig. 6B).

In contrast to miR-122p6,15C,16C, miR-122p6,13U,14A demonstrated only a modest reduction in the ability to promote GLuc expression at 9 h versus 3 h (compared to that of miR-122p6) (Fig. 6A, right) or in the ability to stabilize the S1-S2p6m RNA in Northern blots (Fig. 6B). Thus, while base pairing at nt 13 and 14 of miR-122 contributes to the stabilization of viral RNA, it appears to be less important than the interactions involving nt 15 and 16.

We also examined the ability of the mutated miRNAs to stabilize H77S/GLuc2A-AAG RNA with mutations within either the S1 or S2 seed match sites. Unlike the case for the double S1-S2p6m mutant, GLuc assays (Fig. 6A, left and center panels) and Northern blots (Fig. 6B) revealed that these mutants were both stabilized by wild-type miR-122. This likely reflects interactions with the alternative, nonmutated binding site (e.g., the S2 site in the S1p6m mutant) and is consistent with both miR-122 binding sites contributing to stabilization of the RNA. As expected, the miR-122p6,15C,16C mutant had a smaller effect than that of the control (miR-122p6) on the stability of either the S1p6m or S2p6m mutant, based on GLuc expression (Fig. 6A) and HCV RNA abundance at 9 h (Fig. 6B). Base pairing of nt 15 and 16 at both binding sites is thus important for stabilization. However, in contrast to the substantial enhancement we observed with miR-122p6,13U,14A relative to miR-122p6 in promoting replication of the HJ3-5/GLuc2A-S1p6m mutant (Fig. 2 and 3), miR-122p6,13U,14A was less active

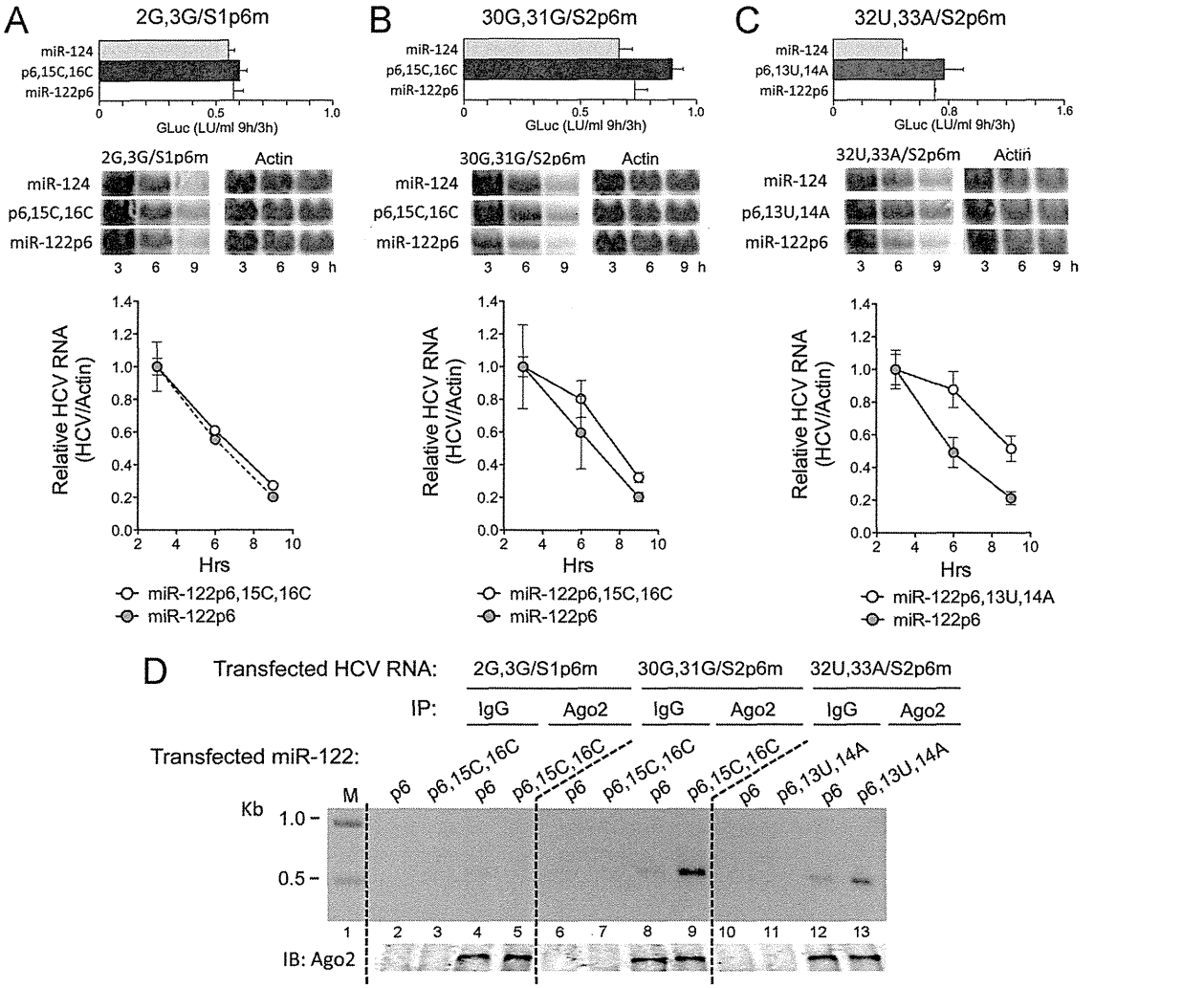

FIG 7 Complementary mutations at HCV nt 30 and 31 restore a stabilization phenotype to miR-122p6 mutants with base substitutions at positions 15 and 16. H77S/GLuc2A-AAG RNAs carrying an S1 or S2 p6m mutation and additional nucleotide substitutions at nt 2 and 3 (A), nt 30 and 31 (B), or nt 32 and 33 (C) were transfected into MEFs with the indicated miRNA duplexes. The top panel of each of these figures shows the ratio of GLuc activity at 9 h versus 3 h. The middle panels show Northern blots of the HCV RNA at 3, 6, and 9 h, with actin mRNA included as a loading control. At the bottom of each panel is shown the quantitation of Northern blotting results from duplicate transfections, comparing HCV RNA to actin mRNA at 3, 6, and 9 h (means ± ranges). Results were normalized to the HCV RNA abundance at 3 h (arbitrarily set to 1.0) in order to compare the stability of the viral RNA cotransfected with miR-122p6 versus the complementary mutant. (D) Ago2-RNA immunoprecipitation results obtained with lysates from cells cotransfected with the mutant HCV RNAs in panels A to C and either miR-122p6 or related mutants.

than miR-122p6 in stabilizing the H77S/GLuc2A-S1p6m mutant (Fig. 6A and B). Subsequent experiments (see below) demonstrated that this difference in the activity of the miR-122p6,13U,14A mutant with the H77S versus HJ3-5 RNA was due to differences in the nucleotide sequences of the two viral RNAs at nt 4.

Additional experiments were carried out to demonstrate whether the mutated miRNAs were capable of stabilizing H77S/GLuc2A-AAG RNAs carrying complementary mutations. These studies revealed that supplementation with miR-122p6,15C,16C resulted in no significant increase in the expression of GLuc 9 h versus 3 h following transfection of 2G,3G/S1p6m RNA compared with that with miR-122p6 (Fig. 7A, top panel). This was confirmed in quantitative Northern blots showing that miR-122p6,15C,16C did not enhance the stability of the 2G,3G/S1p6m RNA (Fig. 7A, middle and bottom panels). This is consistent with the minimal rescue of replication of the 2G,3G/S1p6m mutant by miR122p6,15C,16C (Fig. 5, top panel) and suggests that there may be other factors controlling miR-122 interactions at S1. In contrast, miR-122p6,15C,16C was capable of stabilizing the 30G,31G/S2p6m mutant, as evidenced by GLuc assays and in Northern blots (Fig. 7B), while miR-122p6,13U,14A similarly stabilized 32U,33A/S2p6m RNA (Fig. 7C). These last results provide strong support for the supplemental 3′ base pairing between nt 13 to 16 of miR-122 and the HCV sequence upstream of S2 proposed in Fig. 1, and they indicate its importance in stabilization of the viral genome.

3′ supplementary interactions of miR-122 with HCV RNA contribute to the recruitment of Ago2. miR-122-mediated stabilization of HCV RNA is dependent upon recruitment of Ago2 to the viral 5′UTR (26). To ascertain the involvement of the 3′ sup-

plementary base pair interactions in this process, we used RT-PCR to determine whether HCV RNA was physically associated with Ago2 immunoprecipitated from lysates of cells cotransfected with mutated viral RNAs and their cognate miR-122 mutants, as in the experiments shown in Fig. 7A to C. The technical details of these Ago2-RNA co-IP experiments are provided in Materials and Methods and have been published previously (26). The results of these experiments confirmed that base pairing of nt 13 and 14 and nt 15 and 16 of miR-122 with nt 32 and 33 and nt 30 and 31 in HCV, upstream of the S2 seed match site, is important for Ago2 recruitment (Fig. 7D). HCV RNA was enriched in Ago2 precipitates from cells transfected with the 32U,33A/S2p6m and 30G,31G/S2p6m HCV RNAs and the complementary p6,13U,14A and p6,15C,16C miR-122 mutants, respectively, versus miR-122p6, which lacks the relevant complementary base changes (Fig. 7D, lane 13 versus lane 12 and lane 9 versus lane 8). In contrast, very little HCV RNA was coimmunoprecipitated with Ago2 from lysates of cells transfected with the 2G,3G/S1p6m viral RNA, and there was no enrichment of the RNA in cells cotransfected with the complementary p6,15C,16C miR-122 mutant (Fig. 7D, compare lanes 4 and 5). These results are consistent with the inability of complementary miR-122 mutants with base changes at positions 15 and 16 to fully rescue replication and/or stabilization of either genotype 1a or 2a HCV RNA mutants with cytosine-to-guanine mutations at nt 2 and 3 (Fig. 5A and D). Again, these results suggest that the interaction of miR-122 with HCV RNA upstream of the S1 seed match site may be more complex than simple supplementary base pairing between nt 15 and 16 of miR-122 and nt 2 and 3 of HCV RNA.

Nucleotide 4 of genotype 1 HCV RNA base pairs with nt 14 of miR-122. The results described above demonstrate a surprising difference in the impact of the miR-122p6,13U,14A mutant on replication of HJ3-5 RNA when the miRNA is bound at the S1 seed match site versus its effect on protein expression and stability of the replication-incompetent H77S RNA. While it was significantly more potent than miR-122p6 in promoting the replication of HJ3-5/GLuc2A-S1p6m RNA (Fig. 3), it had little ability to stimulate translation or to stabilize H77S/GLuc2A-S1p6m RNA (Fig. 6A and B). A comparison of the nucleotide sequences of the H77S (genotype 1a) and HJ3-5 (genotype 2a) 5′UTRs (Fig. 1A) suggested a possible explanation for these contrasting effects. Among other differences, base position 4 is adenine in H77S (and most other genotype 1 viruses) but uracil in HJ3-5 (Fig. 1A). (Note, however, that this base is cytosine in most genotype 2 viruses [Fig. 1B].) The presence of A4 in H77S would allow an extension of base pairing at C2 and C3 to include A4, as it could base pair with U14 of miR-122 (Fig. 8A, top panel). Since this pairing is not possible with the miR-122p6,13U,14A mutant, the loss of this pairing could explain the reduced ability of the latter to promote translation or stabilization of the H77S RNA, as shown in Fig. 6A and B. On the other hand, while the genotype 2a HJ3-5 sequence does not allow base pairing with wild-type miR-122 at this position, it does allow it with the miR-122p6,13U,14A mutant (Fig. 8A, bottom panel). Thus, functionally important base pairing between nt 4 of the HCV genome and nt 14 of miR-122 could explain the increase in replication of the HJ3-5/S1p6m RNA cotransfected with miR-122p6,13U,14A, as shown in Fig. 3.

To determine whether loss of base pairing at nt 4 of the H77S RNA is responsible for the inability of miR-122p6,13U,14A to stabilize this RNA, we replaced the adenine at position 4 of the H77S RNA with a uracil. This mutation restores the potential for base pairing with nt 14 in miR-122p6,13U,14A. As anticipated, it conferred upon miR-122p6,13U,14A the ability to enhance GLuc expression and to stabilize the H77S RNA to a greater extent than that with miR-122p6 (Fig. 8B). These results are supportive of base pairing between nt 14 of miR-122 and nt 4 of the genotype 1 HCV RNA.

To further assess how miR-122 base pairs with the first 4 nucleotides of the genotype 2 5′UTR, we examined the impact of G2,G3 and G2,G3,A4 substitutions in HJ3-5/GLuc2A RNA cotransfected with either wt miR-122 or miR-122m15C,16C. There were no mutations in either seed match sequence in these experiments. The G2 and G3 substitutions resulted in a sharp reduction in replication of the RNA, as determined by GLuc expression 72 h after transfection, and in the absence of any stimulation when the RNA was cotransfected with wt miR-122 (Fig. 8C). However, GLuc expression was approximately doubled when the G2,G3 RNA was cotransfected with miR-122m15C,16C yet still remained much less than that observed with wt HJ3-5 (Fig. 8C). This provides additional support for base pair interactions between nt 2 and 3 of HCV and nt 15 and 16 of the miRNA but again shows that complementary mutations in miR-122 only partially rescue guanine substitutions at nt 2 and 3 of the genome (Fig. 5, top panel). The additional introduction of an adenine substitution at nt 4 resulted in a marked boost in replication when the RNA was cotransfected with miR-122m15C,16C but had no apparent impact when the RNA was cotransfected with wt miR-122 (Fig. 8C). There was also no increase in replication after cotransfection of miR-122m15C,16C when only the A4 substitution was introduced into HJ3-5 RNA, although this led to a 4-fold increase in GLuc expression over that in wt HJ3-5 in the absence of miR-122 supplementation (Fig. 8C). This likely reflects formation of the additional base pair at nt 4 with nt 14 of endogenous miR-122 (Fig. 8D). Consistent with this, replication was further increased when the A4 RNA was cotransfected with wt miR-122. Collectively, these data provide strong support for functionally important base pair interactions between nt 4 of HCV and nt 14 of miR-122. However, this interaction is possible only when the base composition of HCV includes an adenine at nt 4 and is permissive for it (i.e., in genotype 1 and genotype 6 viruses, as shown in Fig. 1B).

DISCUSSION

The data we present here both confirm and extend previous studies suggesting that bases 3′ of the seed sequence in miR-122 form accessory Watson-Crick interactions with HCV RNA at base positions located upstream of both the S1 and S2 seed match sequences. Previous studies by Machlin et al. (20) demonstrated that the introduction of complementary mutations at HCV nt 2 and 3 and nt 30 and 31 rescued the ability of miR-122 with base substitutions at nt 15 and 16 to promote amplification of a subgenomic genotype 1a HCV RNA in transfected cells. We confirmed this in studies of genome-length RNA possessing a genotype 2a 5′UTR (Fig. 5). However, we provide new genetic complementation data suggesting that functionally important base pairing extends to nt 14 of miR-122 and HCV nt 4 (Fig. 8), upstream of the S1 seed match site, and to nt 13 and 14 of miR-122 and HCV nt 32 and 33, upstream of S2 (Fig. 5). We also extend the understanding of the functional significance of these 3′ supplementary base pair interactions of miR-122 by showing that they are important for the

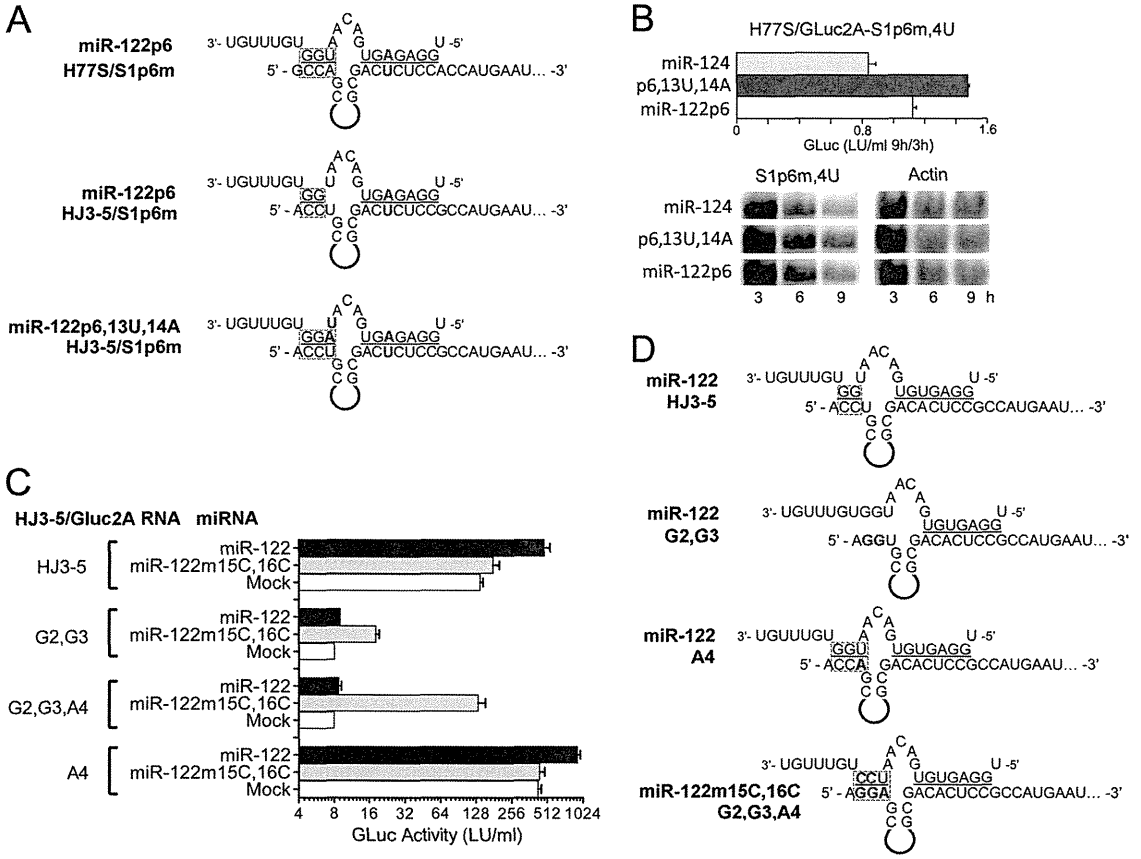

FIG 8 Influence of HCV nt 4 on the interaction with miR-122. (A) The adenine present at nt 4 of H77S (genotype 1a) (top) RNA provides an opportunity for an additional base pair to form with U14 of miR-122. This is not possible with HJ3-5 (genotype 2a) (middle) RNA, in which nt 4 is uracil. However, U4 of HJ3-5 can pair with A14 of miR-122p6,13U,14A (middle), potentially explaining why this mutant has more activity than wt miR-122 at the S1 site in HJ3-5 (see Fig. 3). (B) Like the case with HJ3-5, miR-122p6,13U,14A has greater activity than wt miR-122 in stabilizing the H77S/GLuc2A-S1p6m,4U mutant. Data shown represent the ratios of GLuc activities at 9 h versus 3 h (top) and Northern blots of the RNA at 3, 6, and 9 h posttransfection. (C) GLuc expressed by FT3-7 cells cotransfected with HJ3-5/GLuc2A or related mutant RNAs with wt miR-122 versus miR-122m15C,16C (no p6m mutation in either). "Mock" indicates no oligoribonucleotide. See the legend to Fig. 6 for additional details. (D) Putative base pair interactions involving nt 2 to 4 of HCV and wt miR-122 or miR-122m15C,16C.

production of infectious virus progeny from RNA-transfected cells (Fig. 4) and that they contribute significantly to the ability of miR-122 to both recruit Ago2 to the viral genome (Fig. 7D) and physically stabilize the viral RNA and lower its rate of decay in cells (26) (Fig. 6 and 7A to C).

For at least a subset of miRNAs, the specificity with which cellular mRNAs are targeted depends partly upon sequences 3′ of the seed sequence (nt 2 to 8). The base pairing that occurs between these miRNAs and their target mRNAs outside the seed sequence typically spares nt 9 to 12 of the miRNA and involves primarily nt 13 to 16 (1, 12). This is likely driven by the architecture of the miRISC and, more specifically, by the need to maintain proper orientation of the miRNA guide strand loaded into an argonaute protein (1). The seed sequence (~nt 2 to 8) is positioned to the exterior of this complex, where it is preoriented in a manner facilitating its base pairing with the mRNA target. The extension of contiguous base pairs beyond nt 8 of the miRNA would result in a helical structure that is likely to disrupt its association with Ago (1). Thus, the lack of involvement of nt 9 to 12 of miR-122 in base pair interactions with the HCV genome, as indicated by our studies, is likely to facilitate continued association of miR-122 with Ago2 yet permit additional base pair interactions, involving nt 13 to 16, that enhance the affinity and specificity of the target interaction. In contrast to our findings, recent studies using selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) to analyze the impact of miR-122 on the structure of the 5′UTR concluded that base pairing outside the seed sequence at S2 involves nt 9 to 12 of miR-122 (22). This likely reflects the fact that these studies were done *in vitro*, in the absence of Ago2. Supplemental base pairing of miRNAs with their targets 3′ of the seed sequence may not be common (13). However, the pattern of accessory base pair interactions revealed here for miR-122 and HCV RNA at S2 is consistent with what is known of the interaction of miRNAs with their mRNA targets in the context of a loaded miRISC, and thus consistent with miR-122 recruiting Ago2 to the HCV 5′UTR (26).

In addition to genome-length HCV RNAs such as those used in this study, short reporter RNAs containing the authentic 5′- and 3′UTR sequences of HCV have been used to demonstrate a role for miR-122 in promoting translation directed by the HCV IRES, as well as potential 3′ supplementary base pair interactions of miR-122 (14, 24). However, it remains uncertain how well these

reporter RNAs model the interactions of miR-122 with replication-competent viral RNAs. While the potential for base pairing between nt 2 to 4 of a reporter RNA and 3′ miR-122 sequences was demonstrated in a recent study by the loss of miR-122 regulation of reporter activity when substitutions were made at nt 3 and 4 of the reporter, no effect on miR-122 regulation was apparent with substitutions at nt 30 to 34, upstream of the S2 seed match site (24). These results contrast sharply with those shown in Fig. 5A, where the regulation of viral RNA replication by miR-122 was significantly ablated by G30,G31 and U32,A33 substitutions in HJ3-5/GLuc2A RNA yet rescued by complementary mutations in miR-122. These differences demonstrate that studies with short reporter RNAs need to be interpreted with caution and, where possible, confirmed with replication-competent RNAs.

Although our finding that miR-122 nt 15 and 16 are important for replication of genotype 1 HCV RNA (Fig. 3) is in agreement with the work of Machlin et al. (20), we observed substantial differences in the impact of other nucleotide substitutions, particularly at the most 5′ S1 binding site. We found that miR-122p6,13U,14A was >3-fold more potent than miR-122p6 in promoting the replication of HJ3-5/GLuc2A-S1p6m RNA (Fig. 3 and 4), while Machlin et al. (20) reported that a very similar miR-122 mutant had a reduced ability to support HCV RNA replication. This difference is related to variation in the 5′UTR sequences studied. In contrast to its increased ability to promote replication of HJ3-5 RNA, which contains a genotype 2a 5′UTR, the 13U,14A miR-122 mutant was unable to stimulate translation of or stabilize H77S/GLuc2A-S1p6m RNA (Fig. 6A and B), which, like the subgenomic HCV RNA studied by Machlin et al. (20), contains a genotype 1a 5′UTR. This was reversed by replacing the adenine normally present at position 4 of the genotype 1a 5′UTR with the uracil present in genotype 2a viruses (Fig. 8B). These results thus demonstrate that there are genotype-specific differences in base pairing of HCV RNA with miR-122. They also show that there is a potential for productive base pair interactions at nt 4 of genotype 1 and 6 HCVs that does not exist in other HCV genotypes (Fig. 1B).

There is also a potential for base pair formation between nt 1 of HCV (either guanine or adenine) and nt 17 of miR-122 (uracil) (Fig. 1B). Although we did not investigate this in further detail, the decreased ability of miR-122p6,17G,18C to promote HJ3-5/S1p6m replication (Fig. 3) suggests that it may be important. It is striking, however, that the potential loss of base pairing at HCV nt 1 has less impact than the loss of pairing at nt 4 (compare Fig. 3 and 8C). A model of the seed sequence and 3′ supplementary base pair interactions of miR-122 with the genotype 1a H77S virus, including the recruitment of two miR-122 molecules to the 5′ end of the RNA, is shown in Fig. 9. Although the base pair interactions upstream of the S1 seed match site shown in this model are supported by data presented in this communication, it is important that we were able to achieve only partial rescue of the negative effects of guanine substitutions at nt 2 and 3 of the HCV genome by cotransfection of miR-122 mutants with potentially complementary substitutions at nt 15 and 16 (Fig. 5A and D and 8C). We have no explanation for why this was not observed by Machlin et al. (20), as we encountered this issue with HCV RNAs containing either the genotype 1a or genotype 2a 5′UTR (Fig. 5A and D). mfold predictions suggest that the lack of complete rescue was not due to aberrant folding of the 5′-terminal HCV sequence (Fig. 5C), although this cannot be ruled out entirely.

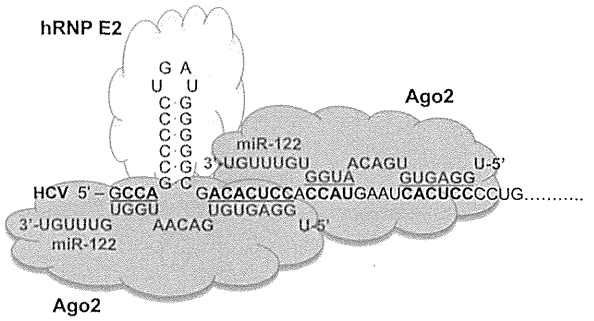

FIG 9 Base pair interactions between miR-122 and the 5′UTR terminal sequences of genotype 1a H77S RNA demonstrated by mutational studies. The model shown includes two distinct Ago2 molecules recruited by miR-122 molecules bound to the S1 and S2 seed match sequences. PCBP2 (hRNP E2) is shown bound to stem-loop 1 as described by Fukushi et al. (10). Interactions of miR-122 with the 5′ 30 nt of the HCV sequence are likely to be more complex than that shown here, since complementary substitutions at nt 2 and 3 of the HCV genome fail to rescue Ago2 recruitment by miR-122p6,15C,16C, targeting the S1 seed match site, and only partially rescue the ability of miR-122p6,15C,16C to promote HCV RNA accumulation in transfected cells.

One possibility, mentioned above, is that the native cytosines are required for recognition by the replicase complex. However, the residual defect in replication of the HJ3-5/GLuc2A-G2,G3 mutant cotransfected with miR-122 containing complementary 15C and 16C substitutions was overcome to a significant extent by an adenine substitution at nt 4 of HCV that is predicted to extend the interaction of miR-122 with nt 2 and 3 of the HCV RNA by an additional base pair (Fig. 8C and D). This suggests that the lack of complete rescue of the 2G,3G mutants may be due to a failure to recreate the native interactions of miR-122 with 5′ HCV RNA upstream of the S1 seed match site rather than a defect in replicase recognition of HCV RNA. This notion is reinforced by the failure of the complementary mutants to stabilize the viral RNA, as shown in Fig. 7A, or to efficiently recruit Ago2, as indicated in the Ago2-RNA co-IP experiment shown in Fig. 7D (left panel). Thus, the interactions of miR-122 with HCV RNA upstream of the S1 seed match site are likely to be more complex than those shown in Fig. 9. It is notable that SHAPE was not able to map the effect of miR-122 on RNA structure upstream of S1 (22), and it is possible that there are several alternative structures, only one of which is shown in Fig. 9. Several conserved cytosines in nt 6 to 10 are predicted to contribute to a stable stem-loop structure in the 5′UTR (SL-1 in Fig. 1B) but, alternatively, could pair with G15 and G16 of miR-122. However, this stem-loop has been suggested to bind poly(rC) binding protein 2 (PCBP2; also known as hRNP E2) (Fig. 9) and to be required for HCV replication (10, 29). Thus, although the data we present here provide a more detailed view of the interaction of miR-122 with HCV RNA than that available previously, it is clear that additional studies will be needed to resolve the structure(s) of the complex formed by miR-122 at the extreme 5′ end of the viral RNA.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Institutes of Health (RO1-AI095690 and P20-CA150343) and the University Cancer Research Fund.

REFERENCES

1. **Bartel DP.** 2009. MicroRNAs: target recognition and regulatory functions. Cell 136:215–233.