Virus Adaptation and Treatment

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

# Conformations of the monomeric hepatitis C virus RNA-dependent RNA polymerase



This article was published in the following Dove Press journal:
Virus Adaptation and Treatment
9 March 2010
Number of times this article has been viewed

Sreedhar Chinnaswamy[1,2,3]
Ayaluru Murali[1,3]
Hui Cai[1]
Guanghui Yi[1]
Satheesh Palaninathan[2]
C Cheng Kao[1]

[1]Interdisciplinary Biochemistry Program, Indiana University, Bloomington, Indiana, USA; [2]Department of Biochemistry and Biophysics, Texas A and M University, College Station, Texas, USA; [3]These authors contributed equally to this work and can be considered co-equal first authors

**Abstract:** The hepatitis C virus (HCV) RNA-dependent RNA polymerase (RdRp) changes its conformation and oligomerization state in association with the steps in RNA synthesis. Using a human right hand as an analogy, the crystal structure of the HCV RdRp has extensive interactions between the "finger" and "thumb" domains which result in a closed conformation. However, the RdRp must form a partially open conformation to accommodate the nascent and template RNA duplex during RNA synthesis, and to interact with the retinoblastoma protein through residues in the "palm" domain. A motif named the Δ1 loop has been previously proposed to regulate the transition from the closed to the open conformation. We used negative-stain electron microscopy and single particle reconstruction to identify several conformations of the HCV RdRp monomer, from closed to open. An RdRp with five amino acids deleted in the tip of the Δ1 loop resulted in an open conformation, confirming the importance of this loop in regulating RdRp conformations. Bioinformatics analysis of HCV strains focusing on the Δ1 loop and its interacting surfaces further defined the requirements for this gating mechanism in the HCV RdRp. These results provide glimpses into the dynamic conformations of the HCV RdRp and should provide insights into alternative conformations of the HCV polymerase that could serve as targets for antiviral development against HCV.

**Keywords:** hepatitis C virus, RNA-dependent RNA polymerase, single particle reconstruction, conformational change

## Introduction

Up to 3% of the world's population is infected with HCV (hepatitis C virus), the outcome of which could be liver cirrhosis or hepatocellular carcinoma.[1] Vaccines for HCV are currently not available. Furthermore, the established treatment regime of interferon and ribavirin is not only poorly tolerated but often ineffective with some of the six HCV genotypes and a number of subtypes.[2–4] Clearly, more effective therapies are needed for HCV. A better understanding of the HCV infection process will provide new targets for the design of HCV treatments.

The nonstructural proteins required for HCV replication are important therapy targets in HCV. The HCV RNA-dependent RNA polymerase (RdRp), NS5B, is also one of the best characterized RdRps from positive-strand RNA viruses.[5–7] Recombinant NS5B is usually expressed without the C-terminal 21 residues to improve its solubility. We primarily work with the RdRp from the 1b genotype of HCV. The structure of the 2a RdRp has also been recently reported.[8] The vast majority of the 1b RdRps and the 2a RdRp form a structure that resembles a closed right hand, with "thumb", "palm", and "finger" subdomains. The metal coordinating residues in the RdRp active site are

Correspondence: C Cheng Kao
Interdisciplinary Biochemistry Program, Indiana University, Bloomington, 107 S Indiana Ave, IN 47405, USA
Tel +1 812 855 7583
Fax +1 812 856 5710
Email ckao@indiana.edu

in the palm subdomain that is usually shielded by extensive thumb-finger interaction.[9,10] The closed conformation in the absence of template is in contrast with the DNA-dependent RNA polymerases, which can transition from an open to a closed complex upon template recognition.[11]

A major influence in the models for viral RNA-dependent RNA synthesis comes from the ternary structure of the phage ϕ6 RdRp.[12] Like the HCV RdRp, the ϕ6 RdRp forms a closed structure with a defined template channel. Both template channels are too narrow to accommodate the partially duplexed RNA that must form during RNA synthesis, indicating that the RdRp needs to transition to a more open conformation after initiation of RNA synthesis.[13] Consistent with a need for a change in conformation, a partially open conformation was also identified by X-ray crystallography and revealed significant reconfiguration of the secondary structure in the region where the finger domain interacts with the thumb domain.[14]

A secondary structure named the Δ1 loop that extends from the finger to the thumb subdomains has been proposed to regulate the open and closed forms of the HCV RdRp.[9,13] Mutations that affect the interaction between the Δ1 loop and the hydrophobic residues that it contacts have resulted in polymerases that can perform primer extension but not *de novo* initiation from short RNA templates.[13]

The open and closed structure transition *in vivo* may allow or be influenced by cellular proteins. The HCV RdRp is reported to interact with the tumor suppressor retinoblastoma (Rb) protein, to manipulate the cell cycle, and to cause its degradation.[15] The L-X-C/N-X-D Rb binding motif overlaps with the metal-coordinating residues of the HCV RdRp.[15] Therefore unless the RdRp adopts a more open conformation the Rb binding motif will remain inaccessible. The cellular protein cyclophillin, could also interact with the HCV RdRp to affect its ability to bind RNA, perhaps through a change in NS5B conformation.[16]

We have sought to understand better how the HCV RdRp conformations relate to RNA synthesis. This type of analysis will contribute to understanding of the mechanism of action of viral RdRps and provide new insight into conformation-specific drug targets. X-ray crystallography is often not feasible for the analysis of the flexible conformations. We used electron microscopy and single particle reconstruction to capture semi-stable states of the HCV RdRp. In addition, reconstruction of a mutant version of the HCV RdRp lacking the tip of the Δ1 loop resulted in a primarily open conformation. These results confirm the role of the Δ1 loop in conformational change in the HCV RdRp. Analysis of the residues in the Δ1 loop and the thumb domain identified several conserved interacting residues and differences in the interactions that formed the closed conformation of the 1b and 2a RdRps.

# Materials and methods

## Production of the HCV RdRp

The wild-type HCV RdRp we used is named Δ21 because it lacks the C-terminal 21 residues that interact with the cell membrane. Δ21 and mutant derivatives were expressed in *Escherichia coli* as C-terminal 6X His tagged proteins. The culture was incubated at 37°C with shaking until the $OD_{600}$ reached 0.8–1.0, after which the temperature was lowered to 16°C and IPTG was added to a final concentration of 0.2 mM. Induction of recombinant protein production was for 12–16 hours. The induced cells were pelleted at 8000 g for 15 minutes, and the pellets were frozen at −70°C until needed.

Ni-NTA purification used cells suspended in lysis buffer 20 mM Tris [pH 7.5], 300 mM NaCl, 10% glycerol, 10 mM imidazole, a protease inhibitor cocktail at the concentration suggested by the manufacturer (Sigma Inc.), 5 mM beta-mercaptoethanol (BME), and 0.1 mg/mL lysozyme). After sonication the sample was centrifuged at 16,000 g for 30 minutes. The supernatant was loaded onto Ni-NTA columns equilibrated with lysis buffer. The columns were washed with lysis buffer containing 20 mM imidazole and proteins eluted in 500 mM imidazole-containing buffer.

Δ21 purified from the Ni-NTA column was exchanged into Mono S equilibration buffer (20 mM Tri (pH 7.5), 100 mM NaCl, 10% glycerol, 10 mM BME) using a desalting column and then loaded onto MonoS column (HR 100/10), which was pre-equilibrated with the same buffer buffer. A gradient of 100 to 600 mM NaCl in desalting buffer was employed for elution. The final concentrations of the proteins from more than six preparations of each were 1–3 mg/L for Δ21 and 0.1 to 0.3 mg/L for m26–30. These proteins were highly pure (~98%) without any other detectable bands on SDS-PAGE gels and stored at 3 mg/mL.

## HCV RdRp assays

RNA synthesis by the HCV RdRp used a 19-nt template named LE19, which measures several RNA synthesis activities in one reaction, ie, *de novo* initiation, primer extension, template switch, and non-templated nucleotide addition.[13,17] The RdRp reactions used have been described previously,[13] with the exception that the α-$^{32}$P-CTP was at 33 nM. The buffer in the RdRp assays includes 20 mM sodium glutamate

(pH 8.2), 12.5 mM dithiothreitol, 4 mM $MgCl_2$, 0.5% (v/v) Triton X-100, 20 mM NaCl, and 1 mM $MnCl_2$.

## Differential scanning fluorimetry

Differential scanning fluorimetry utilized a Stratagene MX3005P Real-time PCR machine as described previously.[16]

## Electron microscopy and image reconstruction

The proteins were adjusted to ~1 ng/μL in buffer containing 50 mM Tris pH 7.5, 100 mM NaCl and 1 mM BME, adsorbed to freshly glow-discharged carbon-coated copper grids (EMS), and stained with uranyl acetate (1% w/v aqueous solution). Electron microscopy was performed using either a JEOL 1200 transmission electron microscope operated at 100 kV or the JEOL 1010 at 80 kV. Micrographs containing the proper spread of the particles taken using the JEOL 1200 instrument were digitalized using an Epson Projection 3200 scanner at 1200 dpi, corresponding to 3.7 Å/pixel at the specimen level. Images from the JEOL 1010 were collected with a CCD Camera (4K × 4K) and the micrographs were taken at 50 K magnification with an effective 2.35 Å/pixel at specimen level. Three-dimensional reconstruction of the single particles used the EMAN software package[18] with a Linux operating system as described by Sun et al.[19]

Individual particles of the HCV RdRp were selected using the BOXER routine, filtered, and centered. Class averages were generated without imposing any symmetry. Centered particles (about 38,000) were subjected to the Multirefine program to sort the particles among five models. The initial models used for the multirefinement process were generated from the HCV RdRp crystal structure (pdb: 1quv) with different noise levels to decrease constraints in the conformations. Particles from each of the models were collected separately and subjected to additional single model refinement to improve the convergence of the data set to the model. The reconstructions were iteratively refined until the structure was stable, as judged by Fourier shell correlation (FSC). The convergence was examined by comparing the projections of the model with corresponding class averages. A molecular mass of 65 kDa was used for the surface-rendering threshold of the 3D structures of the monomers, and 130 kDa was used for the dimers. Three-dimensional reconstructions were visualized using the UCSF Chimera software package.[20]

## Molecular modeling of m26–30

A crystal structure of the wild-type HCV-RdRp (Protein Data Bank code 1QUV) and Modeler (Insight II; Accelrys) was used to construct a model of m26–30. Residues 26–30 of the Δ1 loop deletion were substituted with glycines because the program does not allow deletions. The model was then subjected to a medium to high level of simulated annealing optimization with respect to the variable target function of Modeler with an effective radius parameter of 10–20 Å. The final model was selected based on visual inspection and its best fit to the most open EM map using UCSF chimera.

## Analysis of HCV sequences

All the sequences in analysis were extracted from HCV Sequence Databases.[21] Then the residues of Δ1 loop involving in interaction with Δ2 loop and thumb subdomain were chosen by PIC webserver.

# Results

## HCV RdRp exists in multiple conformations or oligomeric states

Several models for HCV RNA-dependent RNA synthesis currently exist in the literature. Some researchers propose that a monomer of RdRp is the form that carries out RNA synthesis.[12,22,23] Others have shown that oligomeric RdRps can interact cooperatively with RNA and that there are two broad surfaces for interactions that can promote the formation of active scaffolds for RNA synthesis.[24–30] It is likely that both forms will contribute to aspects of HCV replication, either directly, or through interaction with other viral and cellular factors.

A highly purified preparation of HCV RdRp free from contaminating RNAs will be necessary to understand HCV RdRp conformation. Our preparation of the HCV Δ21 protein is more than 98% pure, soluble at concentrations in excess of 10 mg/mL, and highly active for RNA synthesis. When examined via a S-200 gel-filtration chromatography column at a 100 mM monovalent salt concentration, Δ21 was present in various mass/conformational states (Figure 1A), with the first peak that eluted at a position that corresponded to molecular mass markers of ~40 kDa (Δ21 is ~65 kDa) and other peaks being less than 30 kDa when compared with the molecular mass markers. This is likely due to interactions of the protein with column matrix, and similar observations were reported previously.[30] SDS-PAGE of the different column fractions confirmed presence of Δ21 in all fractions suggesting that there are likely to be proteins with different masses or conformations in the preparation.

Should the HCV RdRp assume alternative conformations/mass states, a technique called differential scanning fluorimetry that measures the denaturation profile of proteins should

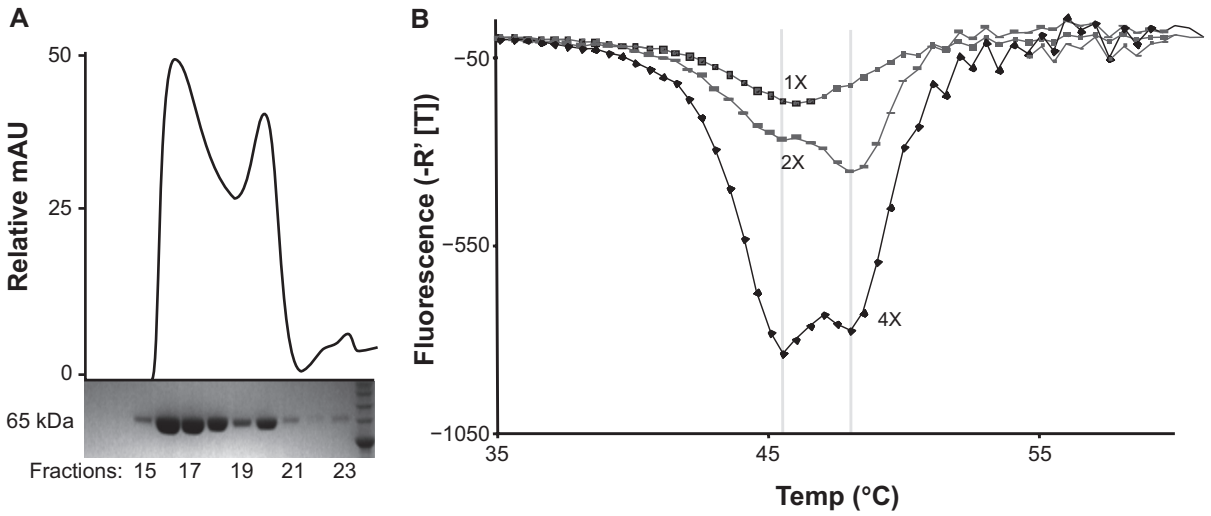



**Figure 1** Characterizations of the HCV RdRp used in this study. **A**) Elution of the 1b HCV RdRp from a S-200 gel filtration column. The column profile was generated using highly purified Δ21 in an AKTA Purifier. The elution volume was calibrated with native molecular weight standards purchased from Sigma Inc. Chromatography was performed at a flow rate of 0.25 mL/min while the elution was monitored by absorbance at 280 nm. The SDS-PAGE gel image below the elution profile of Δ21 shows the Coomassie Blue stained Δ21 within the eluted fractions. The gel image shows that the low molecular weight peaks in the elution profile likely represent Δ21 molecules that interacted with the resin, since the SDS-PAGE revealed that only Δ21 was present in these fractions. **B**) Results of differential scanning fluorimetry for Δ21. The first derivative of the fluorescence of SYPRO orange was determined and plotted against the temperature to allow elucidation of the apparent melting temperature of the protein. Δ21 was used at the final concentrations of 0.5 to 2 μM in this assay.

detect these changes.[31,32] We measure protein denaturation with a dye, SYPRO orange, which fluoresces upon binding exposed hydrophobic regions of proteins. The point of maximum increase in fluorescence is called the $Tm_{app}$. At low protein concentrations, Δ21 had a rather broad denaturation profile, consistent with the protein existing in multiple conformations. With two- or four-fold increases in the concentration of Δ21, at least two populations of proteins with distinct $Tm_{apps}$ can be discerned (Figure 1B). These results suggest that the oligomerization states of the HCV RdRp and/or significant changes can be manipulated by its concentration (Figure 1B). An analysis of the HCV RdRp will have to take into account the changes in the RdRp oligomerization state and conformations.

## RNA synthesis by HCV RdRp RNA

Synthesis assays which used a template named LE19 that can report on both *de novo* initiation and primer-extension in the same reaction was used to quantify the 19-nt product of *de novo* initiated and terminated RNA and the 32-nt product generated by elongation from a primed template. Other products in the reaction are ones with terminal additions to the template (20-nt and 21-nt) that are usually of low abundance and the more prominent template switch products that are multimers of 19-nt in length[33] (Figure 2A). These last two classes of enzymatic activities are likely to be minor in importance *in vivo* when compared with *de novo* initiation and primer extension.

We examined whether RdRp concentration and the associated changes in conformation/oligomerization states correlate with the mode of RNA synthesis. When a range of Δ21 concentrations for RNA synthesis was used, different requirements for *de novo* initiation and elongation were observed (Figure 2B). The primer-extended product was more prominent at the lower RdRp concentration and inhibited at the higher enzyme concentrations (Figure 2B). The opposite trend was observed with the *de novo* initiated product. These results suggest that lower Δ21 concentrations favored extension from a primed template (akin to elongation) while the higher order structures favored *de novo* initiation.

To confirm the effects seen with LE19, we designed two templates that can either direct *de novo* initiation or elongation without an abundance of the template switch products. The *de novo* initiation template is named LE19P, which has a puromycin covalently linked to the 3' terminus of LE19, blocking the ribose 3' hydroxyl from being extended. The primer extension template is PE46, which has a relatively stable hairpin structure with an extended 5' overhang. The 3' terminus of PE46 can be extended to form a complete hairpin of 46-nt. In a reaction with a range of Δ21 concentrations, we again observed that the primer extension product was produced in larger amounts at the lower Δ21 concentrations while *de novo* initiation increased with Δ21 concentration (Figure 2C). These results suggest that if the monomer of Δ21 participates in viral RNA synthesis, it is in elongative

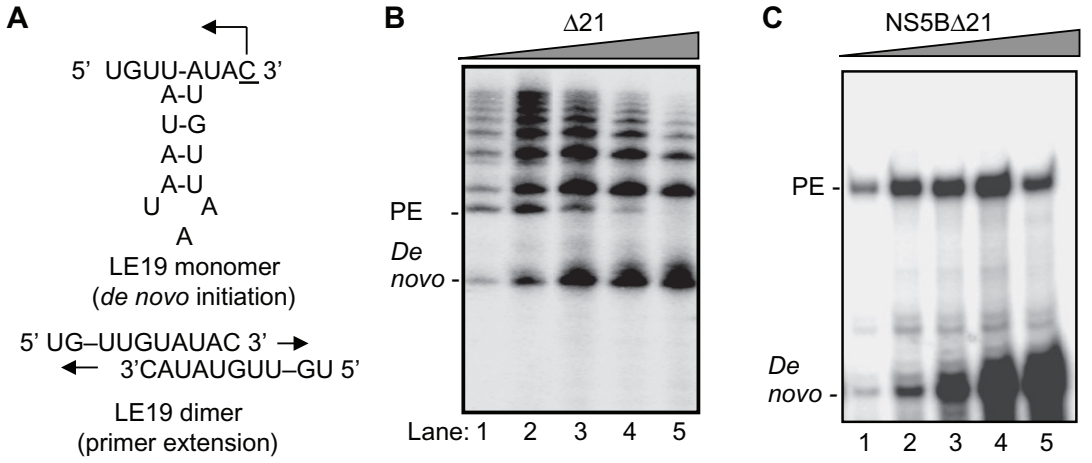



**Figure 2** Effects of Δ21 concentration on RNA synthesis. **A**) Schematic of secondary structure and depiction of the two primary modes of RNA synthesis directed by the RNA template LE19 by the HCV RdRp: *de novo* initiation and primer extension. **B**) RNA synthesis from template LE19 RNA by increasing concentrations of Δ21. Final concentrations of Δ21 were at 78, 96, 192, 384 and 780 nM in lanes 1–5. The enzyme was incubated with LE19 (50 nM) for 5 min. before initiating the RdRp reaction by adding rNTPs and radiolabeled CTP in a solution containing divalent metal ions as described in the Materials and Methods. After incubating the reaction components at 25°C for 1 h, the RNA was extracted with phenol-chloroform and ethanol precipitated before running the products in a 20% urea denaturing polyacrylamide gel. The gel was subjected to autoradiography, and the corresponding RNA band intensities were measured using Phosporimager analysis software (GE Healthcare). The 19-nt *de novo* initiated (*De novo*) and 32 mer primer-extension products (PE) generated from template LE19 are identified to the left of the gel image. **C**) Effect of increasing concentration of Δ21 in RNA synthesis with LE19P or PE46. Δ21 was used at 10, 20, 40, 80, and 160 nM in reactions 1–5. In panels B and C, the *De novo* initiated and primer-extension products are labeled to the left of the gel image to facilitate interpretation of the results and to identify the bands quantified.

synthesis from a nascent template-primer duplex. The oligomerization state for *de novo* initiation will be examined in a separate work. This manuscript focuses on the monomeric forms of Δ21.

## Electron microscopy and single particle reconstruction

We sought to visualize the Δ21 monomers, determine their shapes, and visualize any changes in conformation. A method appropriate for this analysis is the single particle reconstruction of proteins stained with uranyl acetate. Oligomeric states and conformational changes for innate immune receptors in the absence and presence of ligands were captured using this method.[34,35] We stained the peak of putative monomers that eluted within minutes from the S-200 column. A concentration of 1 ng/uL was empirically determined to give a good spread on the EM grids (Figure 3A). More than 200 micrographs were used to collect at least 38,000 well-separated particles. We noted that the particles represented different views of proteins as they landed on the grid as well as possible oligomers. The particles were then centered and sorted using the Multirefine function of the EMAN suite of programs. We initially designated eight

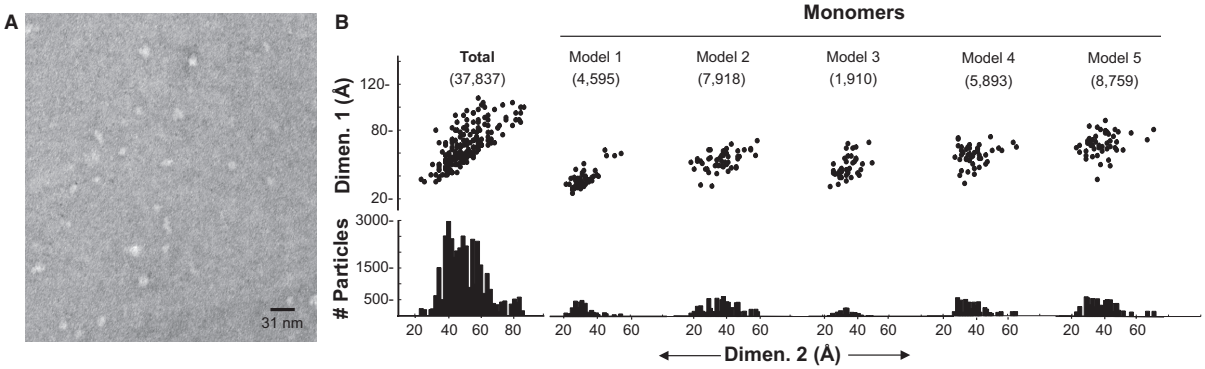



**Figure 3** Analysis of single particles of Δ21 by electron microscopy. **A**) A sample image of the micrographs used for the picking of single particles. Some of the particles may represent monomers, dimers, and higher order oligomers. However, it is also possible that the particles may land in various orientations on the carbon grid. The scale bar used in the imaging is shown at the lower right corner. **B**) Size distribution of all of the single particles for the HCV RdRp used for 3D image reconstruction and the distributions for the five models of the Δ21 monomer. More than 40,000 particles were individually picked for this analysis; 36,000 represent the five minimal classes of monomers as sorted by the multirefine program of EMAN. Each of the five sets of monomeric particles was used to generate class averages for image reconstruction. Each spot in the scatter plot depicts the longest and shortest dimensions (in angstroms) for each of the class averages. The bar graph shows the number of particles present in each class average. The total number of the particles used to generate each of the models is shown in parentheses.

subgroups for the sorting of the particles, but inspection of the results showed that three of the groups were nearly identical to three others. Therefore, we reduced the number of classes to five for 3D model building.

To examine whether the sorting was reasonable, we measured the shortest and longest dimensions for each of the subclasses from the five models and plotted the two dimensions in the graphs shown in Figure 3B. In the lower panel, a histogram shows the number of particles within each subgroup. While the entire collection of particles had a wide distribution (left-most plot in Figure 3B), it is clear that the subgroups within each of the five models are more closely clustered, providing some confidence for the sorting of the subclasses.

Three-dimensional reconstructions without presumed symmetry were performed for of all five models. After the standard two sets of eight refinements, all five models converged to a resolution of 20–23 Å (Figure 4). The final 3D models match well with the projections shown in the class averages, providing some confidence for the models. Interestingly, the five models ranged from those that are in a structure which is more closed related to the X-ray structure (models A and C) to ones which were similar to the X-ray structure (model B) and to two forms that appear to be more open (models D and E). The more closed form may explain the elution of Δ21 at a smaller molecular mass in the gel filtration column. These models offer a snapshot of the quasi-stable conformations that the HCV Δ21 can assume in solution.

## Conformation of m26–30

We hypothesize that the open and closed conformation of Δ21 is due to the loss of interaction between the Δ1 loop and the hydrophobic residues in the thumb domain. To gain evidence in support of this model, we sought to reconstruct the structure of the m26–30 mutant protein, which lacks five residues in the tip of the Δ1 loop. Electron micrographs confirm that the majority of the particles in the m26–30

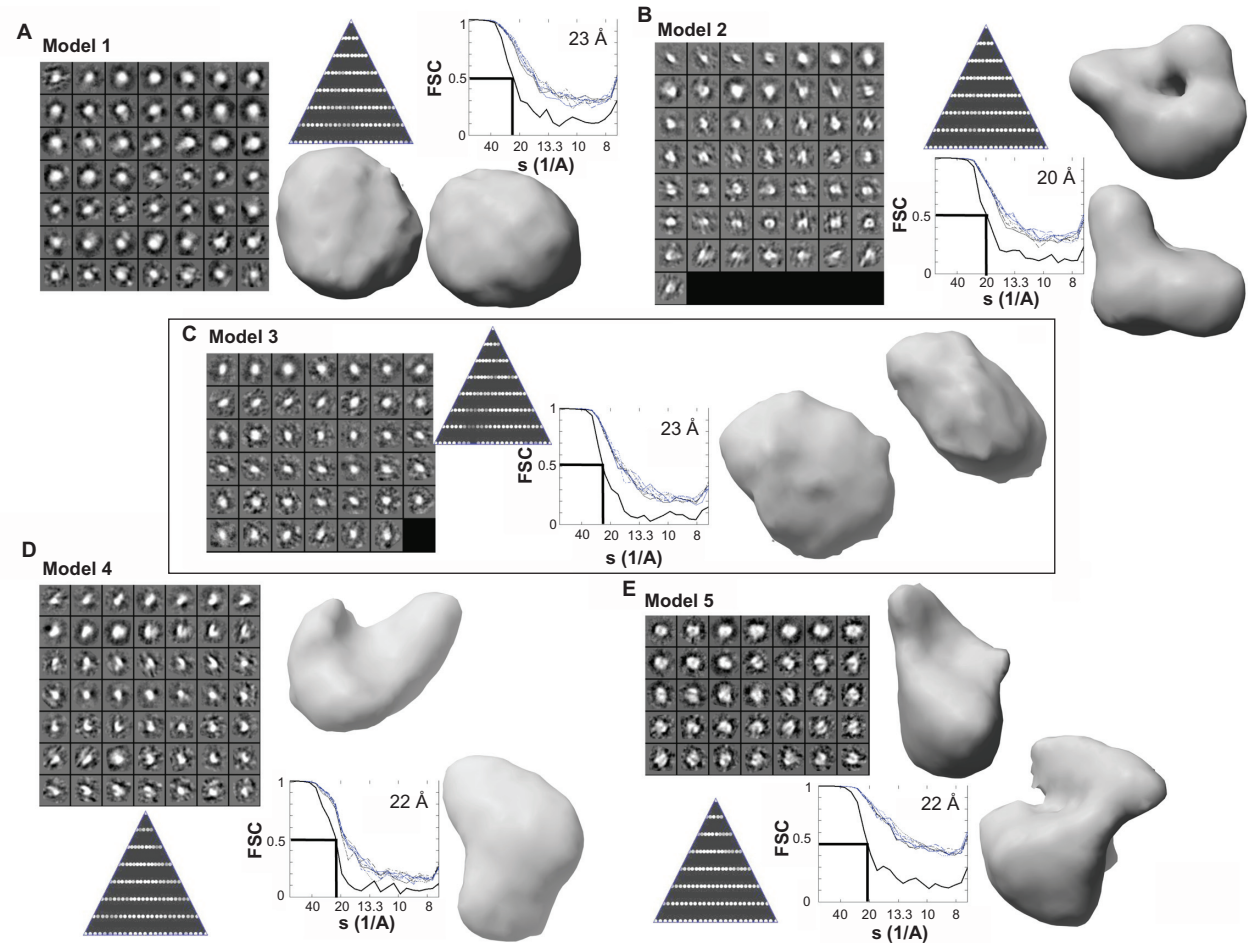



**Figure 4** Data used to construct the five models depicting different conformations of the Δ21 monomer. Panels A–E all contain the class averages generated for each of the five monomer models, the asymmetric triangles that depicts the coverage of all of the views in each model, a Fourier shell correlation plots showing the convergence of the eight rounds of model refinement. The final calculation of the structures, and two views of the model are shown as electron density shells in grey.

preparation are also more homogenous than the particles of Δ21 (Figure 5A).

We selected ~4000 particles from over 20 micrographs and used the EMAN program to generate a 3D model. The class averages suggest that the vast majority of the particles are monomers and belonged to one model (Figure 5B). The reconstruction of m26–30 was achieved in two independent trials using different data sets and yielded the same models at between 18 and 20 Å resolution (Figure 5C and data not shown). The final reconstructed model is a good match to the different views in the class average (Figure 5D), giving confidence in the validity of the model (compare Figure 5B and Figure 5D).

Protein m26–30 was predicted to exist in a predominantly open conformation based on preferential binding to dsRNA.[13] Distinct thumb, finger, and palm subdomains are obvious in the model (Figure 5E), and the exposed template channel has a width of 36 Å at the most constricted juncture, a feature compatible with the ability to bind dsRNA. To facilitate a comparison, the shape of the RdRp determined by X-ray crystallography (pdb: 1QUV) is shown in Figure 5F.

To explore further the arrangement of the subdomains, we attempted to fit the electron density of m26–30 to the existing atomic resolution structures and found that none matched with the extra protrusion observed in the model. However the typical closed form of RdRp can generally fit into the rest of the electron density for m26–30. In our previous studies, we generated a simulated annealing model of RdRp, which depicted a slightly open conformation with an unfolded Δ1 loop and flexibility of the nearby helices.[13] To improve the fit, additional models were generated by varying the annealing temperature and effective radius for the structures of Δ21

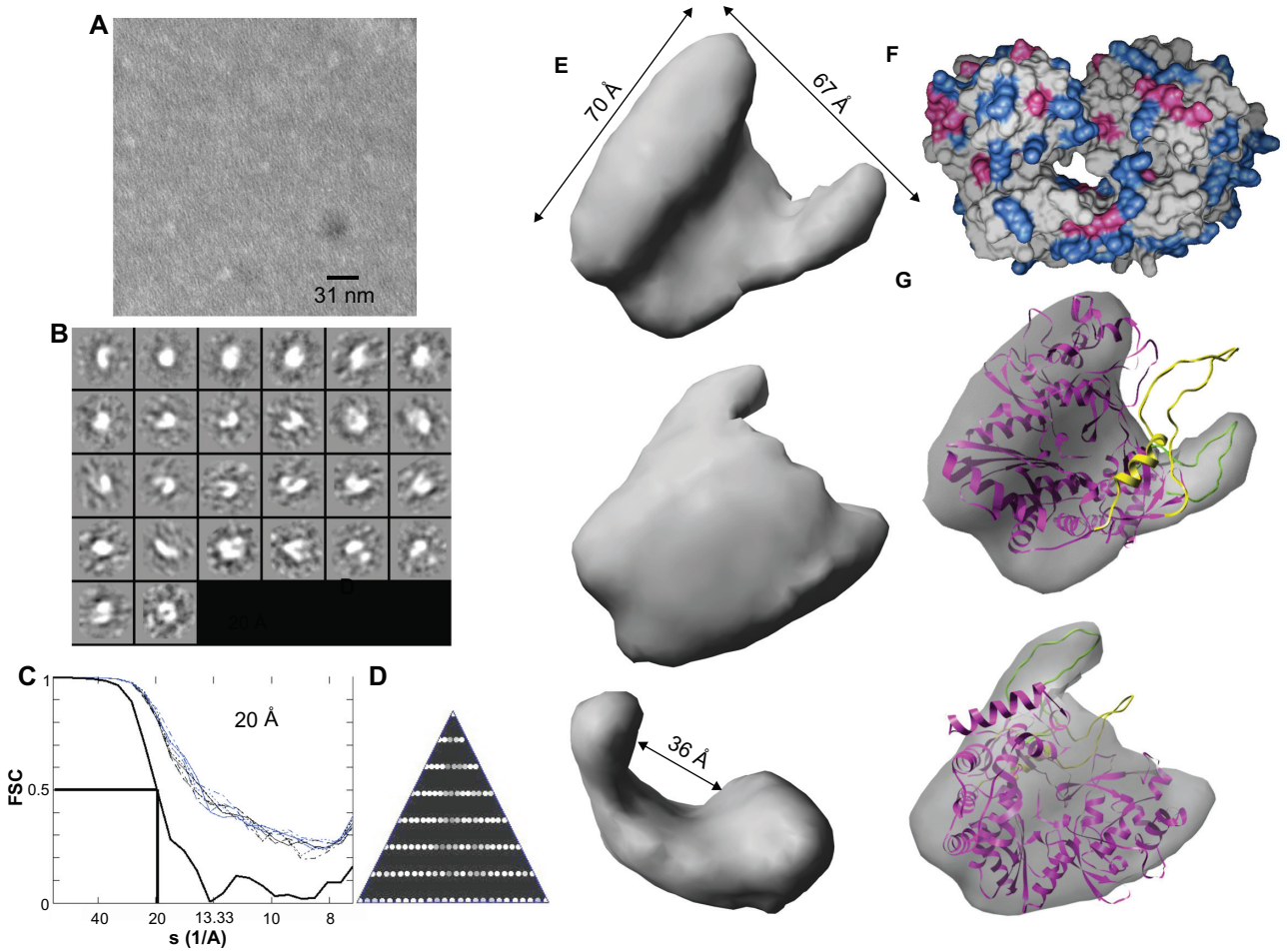



**Figure 5** 3D reconstruction of the m26-30 monomer. **A**) A sample micrograph of the m26–30 protein used for single particle reconstruction. **B**) Class averages of the ca. 4000 particles picked for this analysis. **C**) Fourier Shell Correlation of the refinement process and EOTEST showing that the resolution of the converged m26–30 structure was at 20 Å. **D**) An asymmetric triangle showing the distribution of the particles. **E**) Three views of the model of the m26–30 molecule. The EM envelopes of the reconstructions are shown. **F**) Model of a closed form of HCV RdRp intended to illustrate the difference between the open conformation of m26–30 and the closed conformation of the HCV RdRp. The surface residues that are acidic are shown in red and the basic ones in blue. **G**) Docking of the molecular model generated from the crystal structure of the HCV RdRp into the electron density of m26–30. Two views that differ by 180° are shown. The ribbon structure of the HCV RdRp was generated by replacing residues 26–30 of the HCV crystal structure (pdb 1quv) with glycines and then calculating the lower energy state of the resulting molecule. The altered Δ1 loop is shown in yellow.

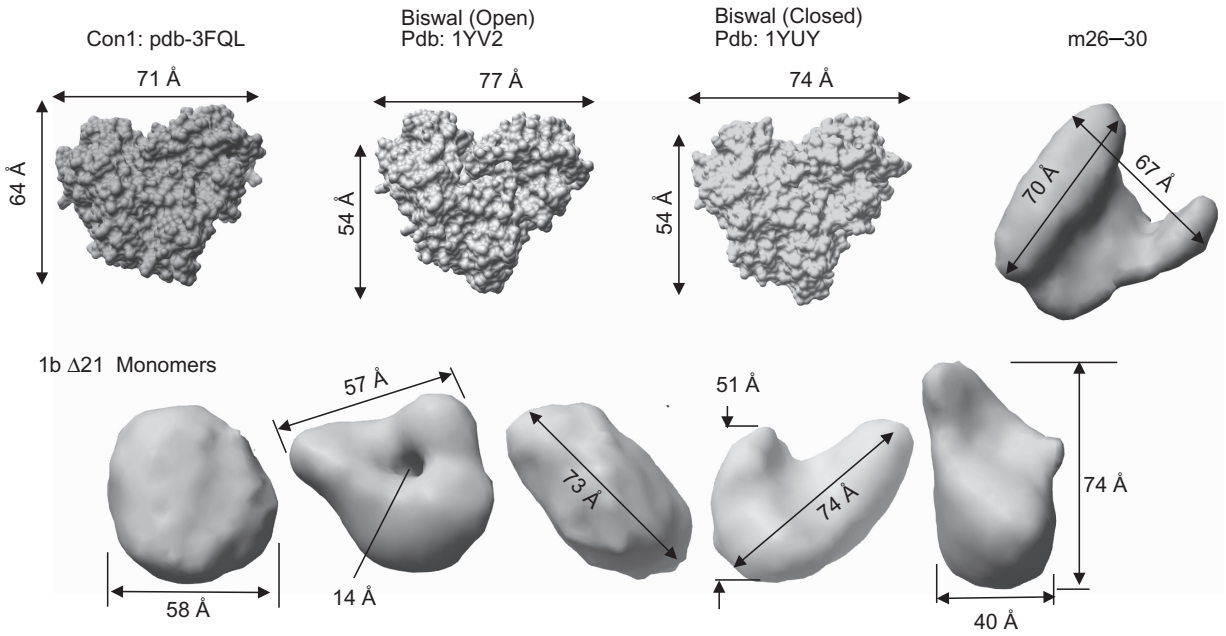



**Figure 6** A compilation of the conformations of the monomers reconstructed or reported in the crystal structure for the HCV RdRp. The crystal structures are shown with higher resolution electron density structures along with their associated pdb numbers. The dimensions for each of the particles were measured using Chimera.

using initial manual fitting followed by further refinement using the Mapfit option of Chimera (UCSF). Two views of the best fit between the calculated structure and the m26–30 3D model are shown in Figure 5G. The correlation of the final fit is ~70% and showed that the palm domain is not dramatically changed. However, the fingertip region has completely moved away from the thumb helices while the simulated model represents an intermediate position in this movement of the loop region. There may even be small changes in the secondary structure(s) as observed by Biswal et al[14] that may account for some portions of the two molecules not fitting better. These results show that single particle reconstruction can be used to identify conformational states of the HCV RdRp that are difficult to capture by X-ray crystallography. This is partly confirmed in Figure 6, where we compare the conformations of the five states for the Δ21 monomers and the open conformation of m26–30 along with the dimensions of the X-ray structures of a 1b RdRp.

## Residues involved in regulation of open and closed RdRp conformations

The reconstruction of the m26–30 structure is consistent with previous claims showing that the Δ1 loop is involved in the regulation of the open and closed conformations. Recently, the crystal structure of the 2a RdRp was solved and also shown to be more capable of *de novo* initiation than the 1b RdRp.[8] The increase in *de novo* initiation is likely due to formation of a more closed conformation. We have confirmed that the 2a RdRp is approximately 2–3 times more effective in *de novo* initiation than the 1b RdRp (data not shown). Our analysis of the RdRp conformations above suggests that changes in the interaction between the Δ1 loop and the regions that it contacts could be partially responsible for the differences in *de novo* initiation. To better understand these interactions, 230 sequences of the Δ1 loop of 1b and 12 sequences of the 2a HCV strains were analyzed to determine which residues in the Δ1 loop are conserved for interactions. Interestingly, a highly conserved portion included residues 26–35 which has a stretch that could form a short α-helix in the 1b RdRp. In addition, the sequences close to where the Δ1 loop originated in the finger domain (residues 10–18 and 40–45) were also more conserved (residues 12–18 and 40–45). We also note that the acidic or basic residues, E17, E18, R32, H34, and R43, tended to be highly conserved. With regard to a comparison between the 1b and the 2a Δ1 loops, the 2a Δ1 loop contained a glutamic acid at position 19 (where a serine or threonine existed for the 1b RdRp) and a lysine at position 36 where methionine or leucine normally existed in other HCV strains. The 2a Δ1 loop thus has two additional charged residues when compared with all of the HCV isolates. This analysis prompted us to examine more closely how the residues in the Δ1 loop could interact with the residues in the thumb domain.

We first used the PIC Webserver[36] program to predict the locations of potential bond forming partners within the Δ1 loops of the 1b and the 2a RdRps (Figure 7b). As would be expected, a number of interactions are predicted in both the Δ1 loops of the 1b and the 2a RdRps. Both contained a number of interactions within the short α-helical region, as

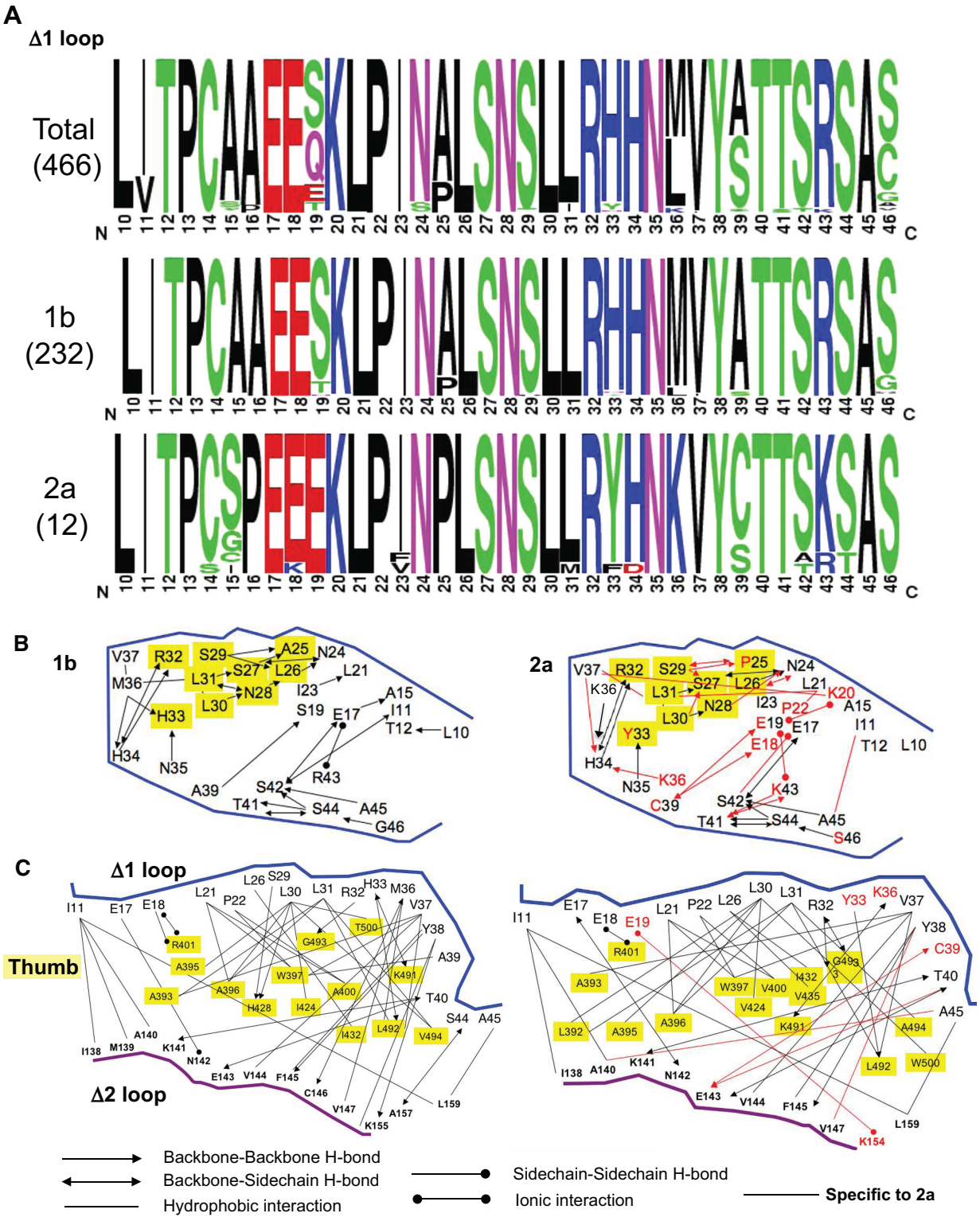



**Figure 7** Analysis of the HCV RdRp sequences involved in the formation of the open or closed structures. **A**) Alignments of the D1 loop sequences from a combination of HCV RdRp sequences from different genotypes. The degree of conservation is shown by the size and number of single amino acid code at each position. The number of sequences used for each analysis is shown in parenthesis under the HCV genotype. **B**) Predicted interactions within the D1 loops for the 1b and 2a HCV RdRps. The key to the interactions are shown at the bottom of the figure. The residues highlighted in yellow are the ones that are at the tip of the Δ1 loop and contain a short a-helix. **C**) The interactions between the D1 loops of the 1b and 2a RdRps with the thumb subdomain (residues highlighted in yellow) and the D2 loop. Both the D1 and the D2 loops have been splayed to allow the interactions to be shown more clearly.

well as between the two termini of the loop as they emerged from the finger domain. In addition, a number of additional interactions (highlighted in red, Figure 7B, right panel) are seen with the 2a Δ1 loop. A detailed summary of the interactions in the Δ1 loops of the 1b and the 2a RdRps are in supplemental Tables 1 and 2, respectively.

Next, we looked for possible interactions between the Δ1 loop and the rest of the RdRp. The Δ1 loop had interactions with the thumb subdomain and with the Δ2 loop. In addition, we observed that there were several new interactions within the Δ1 loop of the 2a RdRp with the Δ2 loop. Of special interest is the interaction between glutamic acid 19 and lysine 154, which adds a sidechain-sidechain interaction that should be of interest for future analysis of the interactions that promote *de novo* initiation. The other new charged residue may H-bond with K491 in the thumb domain. A detailed summary of the interactions in the Δ1 loops of the 1b and the 2a RdRps with the thumb subdomain and the Δ2 loop are in supplemental Tables 3 and 4, respectively.

There are several key residues in the interactions between the Δ1 loop and the thumb domain. In the thumb domain, a key sidechain-sidechain interaction between E18 and R401 is conserved between the two genotypes of HCV RdRps. Several thumb residues also participate in multiple conserved interactions, including W397 (interacting with L21, P22, and V37, and A39 in the 1b RdRp), and A396 and L492 that participate in multiple, primarily hydrophobic interactions with the residues in the Δ1 loop. In the Δ1 loop, I11, L21, P22, L30, L31, Y38, and V37 anchor a set of hydrophobic interactions. We noted that the 1b RdRp also had M36 involved in hydrophobic interactions while this residue was replaced with a K36 in the 2a RdRp. Altogether, while most of the interactions between the Δ1 loop and the thumb are conserved in the two RdRps, there are sufficient differences in the 1b and the 2a HCV RdRps to account for the altered conformational changes that could result in an increase in *de novo* initiation by the 2a RdRp.

## Discussion

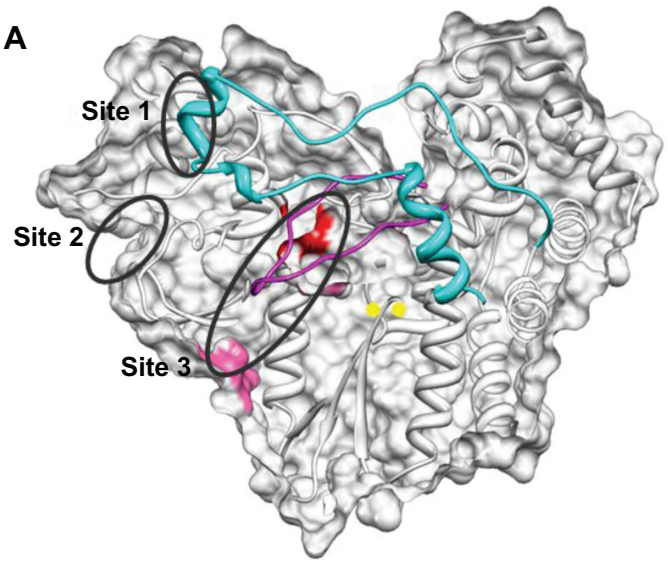
Viral RNA replication and transcription products are recognized by the host innate immunity receptors to signal a viral infection.[37] Therefore, there should be strong selection pressure to regulate viral RNA synthesis both in terms of timing, types, and amounts of replication and transcription products. A likely regulatory step is the conformations of the polymerase that is required for initiation and elongation. In this work, we provide evidence to show that Δ21 exists in multiple conformations and that the interaction between the Δ1 loop and the residues that it contacts can regulate the resulting conformations. This evidence includes low-resolution structures of the monomeric HCV RdRps and the open conformation formed by a mutant RdRp with a short deletion in the Δ1 loop. These conformations are likely to be important factors for RNA-dependent RNA synthesis and for interaction with cellular proteins, such as Rb.

### RdRp conformation and RNA synthesis

The conformation of the HCV RdRp has been reported to distinguish the modes of RNA synthesis.[13] We note that the open and closed structures observed in the monomers could participate in extension from a primed template (which requires lower concentrations of RdRp) and in the formation of oligomeric structures that can direct *de novo* initiation. This more complex scenario is necessitated by our results demonstrating that higher concentrations of the HCV RdRp are more competent for *de novo* initiation, suggesting that monomers may not be sufficient for the initiation of RNA synthesis. In addition, the requirements of an oligomeric HCV RdRp complex to initiate RNA synthesis have been previously proposed for RNA synthesis by the viral RdRps.[24–30] For the poliovirus RdRp, contact between the oligomerized polymerase subunits has been reported to increase the cooperativity in RNA synthesis in poliovirus RdRp.[24,25] Similar results have been observed for the norovirus RdRp.[27] The contacts between poliovirus 3D polymerase monomers and the poliovirus 3C protease have also been proposed to RNA synthesis by regulating protein-priming.[23] We propose that the opening of the monomers could be a key state to allow the interaction between HCV RdRp subunits, or between the RdRp and cellular proteins, such as Rb which interacts with the exposed palm domain of the HCV RdRp.[15]

### RdRp conformation and NNI interaction

The open and closed conformations of the HCV RdRp will have important implications for the design and efficacy of HCV-specific NS5B inhibitors. There are three distinct sites for the binding of non-nucleoside analog inhibitors (NNI, Figure 8A). Site 1 involves the Δ1 loop and thumb domain interface. Site 2 is a pocket in the thumb domain that is present near Site 1. Site 3 lies within the template channel and involves the "primer-grip site" (Figure 8A). The binding of benzimidazole/indoles to Site 1 will affect the interaction between the Δ1 loop and the rest of the polymerase.[38,39] Indeed, two of the residues in the thumb domain, W397 and H428, that mediate multiple interactions with the Δ1 loop overlap with the region that bind benzimidazoles/indole




B

| Class | Compounds | Inhibition stage |
|---|---|---|
| Site 1 | benzimidazoles/ Indoles | Initiation |
| Site 2 | thiophenes pyranindoles | NA Elongation |
| Site 3 | benzothiadiazines/ benzofurans | Initiation |

**Figure 8** The locations for the binding of three classes of nonnucleoside inhibitors (NNIs) of the HCV RdRp. **A**) The partial surface and ribbon representation of the 1a RdRp are rendered from the structure in pdb ID: 1QUV. Sites 1 and 2 are on the surface of the thumb domain and site 3 (colored in red) is located in the template channel. The residues that bind to the primer are shown in pink. The divalent metal coordinating residues are in yellow. **B**) A table summarizing the mechanism of action of the NNIs that bind to the three sites.

compounds (Figure 8). Analysis of the effects of these compounds was consistent with the mutational analyses of the HCV RdRp,[13,40] and the inhibitors did not affect the ability of the HCV RdRp in elongation, but prevented initiation, consistent with a closed structure (or one that requires Δ1 loop-thumb interactions between two or more RdRp subunits) needed for initiation complex assembly.

Site 2 binds compounds based on the thiophene scaffold and acts at the base of the thumb some distance from the Δ1 loop. The binding of the compounds to the 1b RdRp (to residues L419, M423, L474, and W528) did not reveal a dramatic change in RdRp conformation.[14,41,42] However, binding to the 2a RdRp did cause a reconfiguration of the tip of the Δ1 loop structure,[14] confirming that the 1b and 2a RdRps have distinct interactions that could influence RNA synthesis and binding to inhibitors.

The compounds that bind to Site 3, including those from the benzothiadiazine class[43,44] showed differences in response by the different HCV genotypes.[45] For example, 1,5-benzodiazepines were effective against the 1a and 1b HCV replicons but not the 2a replicon. We speculate that the propensity to form an open RdRp conformation and the strength of the interactions between the Δ1 loop, the thumb domain, and the Δ2 loop could facilitate the entry of these inhibitors.

The results from this work show that the dynamic conformations of the HCV RdRp could represent an important regulatory step in RNA synthesis and, more importantly, be exploited by designing NNIs that target the HCV polymerase. In addition, a number of predicted interactions, such as ones between the Δ1 loop and the Δ2 loop, should be examined in detail for effects on RNA synthesis.

## Conclusion

This study documents some of the conformations assumed by the HCV RdRp using electron microscopy and single-molecule reconstruction. The study also helps to demonstrate that an oligomerization state could be a major regulatory factor in the RdRp being able to perform *de novo* initiation or extend from a primed template. The differences in the interactions in this area could help explain the modes of RNA synthesis by the 1b and 2a RdRps and aid in the design of HCV genotype-specific inhibitors.

## Acknowledgements

We thank Laura Kao for editing the manuscript, the Kao laboratory for numerous helpful discussions, and Pengyun Li of the Bioprocessing Facility for purification of several of the recombinant proteins. The Cultivation and Bioprocessing Facility is supported by the Indiana META-Cyt Program funded in part by an endowment from the Lilly Foundation. This work was funded by the National Institute of Allergy and Infectious Diseases for grant 1RO1AI073335 to CK.

## References

1. Wasley A, Alter MJ. Epidemiology of hepatitis C: Geographic differences and temporal trends. *Semin Liver Dis*. 2000;20:1–16.

2. Gomez L, Martell M, Quer J, et al. Hepatitis C viral quasispecies. *J Viral Hepat*. 1999;6:3–16.
3. Feld JJ, Hoofnagle JH. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature*. 2005;436:967–972.
4. Burton JR Jr, Everson GT. HCV NS5B polymerase inhibitors. *Clin Liver Dis*. 2009;13:453–465.
5. Kao CC, Singh P, Ecker DJ. De novo initiation of viral RNA-dependent RNA synthesis. *Virology*. 2001;287:251–260.
6. Van Dijk AA, Makeyev EV, Bamford DH. Initiation of viral RNA-dependent RNA polymerization. *J Gen Virol*. 2004;85:1077–1093.
7. Kao C, Sun JH. Initiation of minus-strand RNA synthesis by the brome mosaic virus RNA-dependent RNA polymerase: Use of oligoribonucleotide primers. *J Virol*. 1996;70:6826–6830.
8. Simister P, Schmitt M, Geitmann M, et al. Structural and functional analysis of hepatitis C virus strain JFH1 polymerase. *J Virol*. 2009;83:11926–11939.
9. Bressanelli S, Tomei L, Roussel A, et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc Natl Acad Sci U S A*.1999;96:13034–13039.
10. Lesburg CA, Cable MB, Ferrari E, et al. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol*. 1999;6:937–943.
11. De Haseth PL, Zupancic M, Record MT. RNA polymerase-promoter interactions: The comings and goings of RNA polymerase. *J Bacteriol*. 1998;180:3019–3025.
12. Butcher SJ, Grimes JM, Makeyev EV, et al. A mechanism for initiating RNA-dependent RNA polymerization. *Nature*. 2001;410:235–240.
13. Chinnaswamy S, Yarbrough I, Palaninathan S, et al. A locking mechanism regulates RNA synthesis and host protein interaction by the hepatitis C virus polymerase. *J Biol Chem.* 2008;283:20535–20546.
14. Biswal BK, Cherney MM, Wang M, et al. Crystal structures of the RNA-dependent RNA polymerase genotype 2a of hepatitis C virus reveal two conformations and suggest mechanisms of inhibition by non-nucleoside inhibitors. *J Biol Chem*. 2005;280:18202–18210.
15. Munakata T, Nakamura M, Liang Y, et al. Down-regulation of the retinoblastoma tumor suppressor by the hepatitis C virus NS5B RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A*. 2005;102:18159–18164.
16. Liu Z, Robida JM, Chinnaswamy S, et al. Mutations in the hepatitis C virus polymerase that increase RNA binding can confer resistance to cyclosporine A. *Hepatology*. 2009;50:25–33.
17. Ranjith-Kumar CT, Kim YC, Gutshall L, et al. Mechanism of *de novo* initiation by the hepatitis C virus RNA-dependent RNA polymerase: Role of divalent metals. *J Virol*. 2002;76:12513–12525.
18. Ludtke SJ, Jakana J, Song JL, et al. A 11.5 A single particle reconstruction of GroEL using EMAN. *J Mol Biol*. 2001;314:253–262.
19. Sun J, Duffy KE, Ranjith-Kumar CT, et al. Structural and functional analyses of the human Toll-like receptor 3. Role of glycosylation. *J Biol Chem*. 2006;281:11144–11151.
20. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25:1605–1612.
21. Los Alamos National Laboratory. Hepatitis C virus (HCV) Database Project. 2009. Available from: http://hcv.lanl.gov/. Accessed on February 20, 2010.
22. Cramer J, Jaeger J, Restle T. Biochemical and pre-steady-state kinetic characterization of the hepatitis C virus RNA polymerase (NS5B delta21, HC-J4). *Biochemistry*. 2006;45:3610–3619.
23. Pathak HB, Ghosh SK, Roberts AW, et al. Structure-function relationships of the RNA-dependent RNA polymerase from poliovirus (3Dpol). A surface of the primary oligomerization domain functions in capsid precursor processing and VPg uridylylation. *J Biol Chem*. 2002;277:31551–31562.
24. Lyle JM, Bullitt E, Bienz K, et al. Visualization and functional analysis of RNA-dependent RNA polymerase lattices. *Science*. 2002;296:2218–222.
25. Pata JD, Schultz SC, Kirkegaard K. Functional oligomerization of poliovirus RNA-dependent RNA polymerase. *RNA*. 1995;1:466–477.
26. Hobson SD, Rosenblum ES, Richards O, et al. Oligomeric structures of poliovirus polymerase are important for function. *EMBO J*. 2001;20:1153–1163.
27. Högbom M, Jäger K, Robel I, et al. The active form of the norovirus RNA-dependent RNA polymerase is a homodimer with cooperative activity. *J Gen Virol*. 2009;90:281–291.
28. Qin W, Luo H, Nomura T, et al. Oligomeric interaction of hepatitis C virus NS5B is critical for catalytic activity of RNA-dependent RNA polymerase. *J Biol Chem*. 2002;277:2132–2137.
29. Wang QM, Hockman MA, Staschke K, et al. Oligomerization and cooperative RNA synthesis activity of hepatitis C virus RNA-dependent RNA polymerase. *J Virol*. 2002;76:3865–3872.
30. Gu B, Gutshall LL, Maley D, et al. Mapping cooperative activity of the hepatitis C virus RNA-dependent RNA polymerase using genotype 1a–1b chimeras. *Biochem Biophys Res Commun*. 2004;313:343–350.
31. Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc*. 2007;2:2212–2221.
32. Yi G, Vaughan R, Yarbrough L, et al. RNA binding by the brome mosaic virus capsid protein and the regulation of viral RNA accumulation. *J Mol Biol.* 2009;391:314–326.
33. Kim MJ, Kao CC. Factors regulating template switch in vitro by viral RNA-dependent RNA polymerases: Implications for RNA-RNA recombination. *Proc Natl Acad Sci U S A*. 2001;98:4972–4977.
34. Murali A, Li X, Ranjith-Kumar CT, et al. Structure and function of LGP2, a DEX(D/H) helicase that regulates the innate immunity response. *J Biol Chem*. 2008;283:15825–15833.
35. Ranjith-Kumar CT, Ayaluru M, Dong W, et al. Agonist and antagonist recognition by RIG-I, a cytoplasmic innate immunity receptor. *J Biol Chem*. 2009;284:1155–1165.
36. Tina KG, Bhadra R, Srinivasan N. PIC: Protein Interactions Calculator. *Nucleic Acids Res*. 2007;35:W473–W476.
37. Takeuchi O, Akira S. Innate immunity to virus infection. *Immunol Rev*. 2009;227:75–86.
38. Kukolj G, McGibbon GA, McKercher G, et al. Binding site characterization and resistance to a class of non-nucleoside inhibitors of the hepatitis C virus NS5B polymerase. *J Biol Chem*. 2005;280:39260–39267.
39. Di Marco S, Volpari C, Tomei L, et al. Interdomain communication in hepatitis C virus polymerase abolished by small molecule inhibitors bound to a novel allosteric site. *J Biol Chem*. 2005;280:29765–29770.
40. Tomei L, Altamura S, Bartholomew L, et al. Characterization of the inhibition of hepatitis C virus RNA replication by non-nucleosides. *J Virol*. 2004;78:938–946.
41. Love RA, Parge HE, Yu X, et al. Crystallographic identification of a noncompetitive inhibitor binding site on the hepatitis C virus NS5B RNA polymerase enzyme. *J Virol*. 2003;77:7575–7581.
42. Biswal BK, Wang M, Cherney MM, et al. Non-nucleoside inhibitors binding to hepatitis C virus NS5B polymerase reveal a novel mechanism of inhibition. *J Mol Biol*. 2006;361:33–45.
43. Dhanak D, Duffy K, Johnston VK, et al. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. *J Biol Chem.* 2002;277:38322–38327.
44. Gu B, Johnston VK, Gutshall L, et al. Arresting initiation of hepatitis C virus RNA synthesis using heterocyclic derivatives. *J Biol Chem*. 2003;278:16602–16607.
45. Nyanguile O, Pauwels F, Van den Broeck W, et al. 1,5-Benzodiazepines, a novel class of hepatitis C virus polymerase nonnucleoside inhibitors. *Antimicrob Agents Chemother*. 2008;52:4420–4431.

# Supplementry materials

**Table S1** Summary of HCV Genotype 1b NS5B Δ1 loop interaction within Δ1 loop

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| L10 | L 98.7%, M 1.3% | T12 | T 100% | B-B |
| I11 | I 99.1%,V 0.9% | S42 | S 99.1%,T 0.9% | B-S |
| | | A45 | A 100.0% | HI |
| T12 | T100% | L10 | L 98.7%, M 1.3% | B-B |
| P13 | P 100% | S42 | S 99.1%,T 0.9% | S-S |
| C14 | C100% | NONE | | |
| A15 | A 97.8%, G 0.4%, S 0.4%,T 0.4%, V 0.9% | E17 | E 100% | B-B |
| A16 | A 99.6%, S 0.4% | NONE | | |
| E17 | E 100% | A15 | A 97.8%, G 0.4%, S 0.4%, T 0.4%, V 0.9% | B-B |
| | | S42 | S 99.1%,T 0.9% | B-S |
| | | R43 | R 100% | II |
| E18 | E 99.6%, G 0.4% | NONE | | |
| S19 | A 0.4%, N 1.7%, S 88.4%,T 9.5% | A39 | A 95.7%, S 3.9%, V 0.4% | B-B |
| K20 | A 0.4%, Q 0.4%, E 0.4% K 98.7% | NONE | | |
| L21 | L100% | I23 | I 99.6%, V 0.4% | B-B |
| | | V37 | I 2.6%, V 97.4% | HI |
| P22 | P 100% | NONE | | |
| I23 | I 99.6%,V 0.4% | L21 | L 100% | B-B |
| N24 | N 99.6%, I 0.4% | L26 | L 100% | B-B |
| | | S27 | N 0.4%, S 99.6% | B-B |
| | | | | B-S |
| | | N28 | N 99.6%, Y 0.4% | B-B |
| A25 | A 85.3%, P 14.2%, S 0.4% | S27 | N 0.4%, S 99.6% | B-B |
| | | N28 | N 99.6%, Y 0.4% | B-B |
| | | S29 | P 1.7%, S 97.8%,T 0.4% | B-B |
| L26 | L 100% | N24 | N 99.6%, I 0.4% | B-B |
| | | N28 | N 99.6%,Y 0.4% | B-B |
| | | S29 | P 1.7%, S 97.8%,T 0.4% | B-B |
| S27 | N 0.4%, S 99.6% | N24 | N 99.6%, I 0.4% | B-S |
| | | A25 | A 85.3%, P 14.2%, S 0.4% | B-B |
| | | L31 | L 100% | B-B |
| N28 | N 99.6%,Y 0.4% | N24 | N 99.6%, I 0.4% | B-B |
| | | A25 | A 85.3%, P 14.2%, S 0.4% | B-B |
| | | L26 | L 100% | B-B |
| | | L30 | Q 0.4%, L 99.6% | B-B |
| | | L31 | L 100% | B-S |
| S29 | P 1.7%, S 97.8%,T 0.4% | A25 | A 85.3%, P 14.2%, S 0.4% | B-B |
| | | L26 | L 100% | B-B |
| L30 | Q 0.4%, L 99.6% | N28 | N 99.6%, Y 0.4% | B-B |
| L31 | L 100% | S27 | N 0.4%, S 99.6% | B-B |
| | | N28 | N 99.6%, Y 0.4% | B-S |
| R32 | R 100% | H34 | R 0.9%, H 99.1% | B-B |
| | | | | B-S |

(*Continued*)

**Table S1** (*Continued*)

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| H33 | R 0.4%, N 1.7%, H 96.1% | N35 | N 99.1%, S 0.9% | B-B |
| | | M36 | L 6.5%, K 0.4%, M 93.1% | B-B |
| H34 | R 0.9%, H 99.1% | R32 | R 100% | B-B |
| | | | | B-S |
| | | M36 | L 6.5%, K 0.4%, M 93.1% | B-B |
| | | V37 | I 2.6%, V 97.4% | B-B |
| N35 | N 99.1%, S 0.9% | H33 | R 0.4%, N 1.7%, H 96.1% | B-B |
| M36 | L 6.5%, K 0.4%, M 93.1% | H33 | R 0.4%, N 1.7%, H 96.1% | B-B |
| | | H34 | R 0.9%, H 99.1% | B-B |
| | | V37 | I 2.6%, V 97.4% | HI |
| V37 | I 2.6%, V 97.4% | H34 | R 0.9%, H 99.1% | B-B |
| Y38 | Y 100% | NONE | | |
| A39 | A 95.7%, S 3.9%, V 0.4% | S19 | A 0.4%, N 1.7%, S 88.4%, T 9.5% | B-B |
| T40 | A 0.4%, T 99.6% | NONE | | |
| T41 | T 100% | S44 | I 0.4%, S 99.6% | B-S |
| S42 | S 99.1%, T 0.9% | 111 | I 99.1%, V 0.9% | B-S |
| | | P13 | P 100% | B-S |
| | | E17 | E 100% | B-S |
| | | S44 | I 0.4%, S 99.6% | B-B |
| | | A45 | A 100% | B-B |
| R43 | R 100% | NONE | | |
| S44 | I 0.4%, S 99.6% | T41 | T 100% | B-B |
| | | | | B-S |
| | | S42 | S 99.1%, T 0.9% | B-B |
| | | G46 | N 1.7%, C 0.9%, G 11.2%, S 84.5%, T 1.3%, V 0.4% | B-B |
| | | L47 | R 0.4%, Q 76.3%, L 22.8%, K 0.4% | B-B |
| | | R48 | R 100% | B-B |
| A45 | A 100% | S42 | S 99.1%, T 0.9% | B-B |
| G46 | N 1.7%, C 0.9%, G 11.2%, S 84.5%, T 1.3%, V 0.4% | S44 | I 0.4%, S 99.6% | B-B |

**Notes:** Residues in red have one amino acid present at more than 99%.
Cyan denotes positions where all substitutions are all amin acids had polar or all non-polar side chains.
**Abbreviations:** B, Backbone; R, R-group; (H), H bond; (ϕ), hydrophobic; I, Ionic; NONE, none observed in structure.

**Table S2** Summary of the interactions within Δ1 loop of the HCV 2a NS5B

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| L10 | L 100% | NONE | | |
| 111 | I 100% | A45 | A 100% | ϕ |
| | | A49 | A 100% | ϕ |
| T12 | T 100% | NONE | | |
| P13 | P 100% | NONE | | |
| C14 | C 91.7%, S 8.3% | NONE | | |
| S15 | C 8.3%, G 8.3%, I 8.3%, S 75.0% | NONE | | |
| P16 | P 100% | NONE | | |
| E17 | E 100% | NONE | | |
| E18 | E 91.7%, K 8.3% | NONE | | |

(*Continued*)

**Table S2** (*Continued*)

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| E19 | E 100% | K20 | K 100% | I |
| | | K43 | R 8.3%, K 91.7% | I |
| K20 | K 100% | NONE | | |
| L21 | L 100% | P22 | P 100% | ϕ |
| | | L31 | L 100% | ϕ |
| | | V37 | V 100% | ϕ |
| P22 | P 100% | NONE | | |
| I23 | I 83.3%, F 8.3%, V 8.3% | NONE | | |
| N24 | N 100% | NONE | | |
| P25 | P 100% | NONE | | |
| L26 | L 100% | N24 | N 100% | B-B (H) |
| | | | | B-R (H) |
| S27 | S 100% | N24 | N 100% | B-B (H) |
| | | | | B-R (H) |
| N28 | N 100% | N24 | N 100% | B-B (H) |
| S29 | S 100% | P25 | P 100% | B-B (H) |
| | | | | B-R (H) |
| | | S27 | S 100% | B-B (H) |
| L30 | L 100% | S27 | S 100% | B-B (H) |
| | | N28 | N 100% | B-B (H) |
| | | L31 | L 100% | ϕ |
| L31 | L 100% | S27 | S 100% | B-B (H) |
| | | V37 | V 100% | ϕ |
| R32 | R 100% | NONE | | |
| Y33 | Y 100% | NONE | | |
| H34 | D 8.3%, H 91.7% | R32 | R 100% | B-B (H) |
| | | | | B-R (H) |
| N35 | N 100% | Y33 | Y 100% | B-B (H) |
| K36 | K 100% | Y33 | Y 100% | B-B (H) |
| | | H34 | D 8.3%, H 91.7% | B-B (H) |
| V37 | V 100% | H34 | D 8.3%, H 91.7% | B-B (H) |
| Y38 | Y 100% | NONE | | |
| C39 | C 83.3%, S 16.7% | E18 | E 91.7%, K 8.3% | B-R (H) |
| | | E19 | E 100% | B-B (H) |
| T40 | T 100% | NONE | | |
| T41 | T 100% | NONE | | |
| S42 | A 8.3%, S 83.3%, T 8.3% | E17 | E 100% | B-R (H) |
| | | | | R-R (H) |
| K43 | R 8.3%, K 91.7% | T41 | T100% | B-B (H) |
| | | | | B-R (H) |
| S44 | S 83.3%, T 16.7% | T41 | T100% | B-B |
| | | | | B-R (H) |
| | | S42 | A 8.3%, S 83.3%, T 8.3% | B-B (H) |
| A45 | A 100% | S42 | A 8.3%, S 83.3%, T 8.3% | B-B (H) |
| | | A49 | A 100% | B-B (H) |

(*Continued*)

**Table S2** (*Continued*)

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| S46 | S 100% | S44 | S 83.3%, T 16.7% | B-B (H) |
| | | R48 | R 100% | B-B (H) |
| | | A49 | A 100% | B-B (H) |
| | | K50 | K 100% | B-B (H) |

**Notes:** Residues in red have one amino acid present at more than 91%.
Cyan denotes positions where all substitutions are all amin acids had polar or all non-polar side chains.
**Abbreviations:** B, Backbone; R, R-group; (H), H bond; ϕ, hydrophobic; I, Ionic; NONE, none observed in structure.

**Table S3** Summary of the Δ1 loop interaction with the thumb domain and the Δ2 loop in the HCV 1b NS5B

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| L10 | L 98.7%, M 1.3% | NONE | | |
| I11 | I 99.1%,V 0.9% | I138 | I 92.2%, V 7.8% | ϕ |
| | | M139 | M 100% | ϕ |
| | | A140 | A 100% | ϕ |
| | | L159 | L 95.3%, F 4.7% | ϕ |
| T12 | T 100% | NONE | | |
| P13 | P 100% | NONE | | |
| C14 | C 100% | NONE | | |
| A15 | A 97.8%, G 0.4%,S 0.4%, T 0.4%, V 0.9% | NONE | | |
| A16 | A 99.6% S 0.4% | NONE | | |
| E17 | E 100% | N142 | N 89.7%, S 9.9%, T 0.4% | R-R (H) |
| E18 | E 99.6%, G 0.4% | R401 | R 97.4%, K 2.6% | I<br>R-R (H) |
| S19 | A 0.4%, N 1.7%, S 88.4%,T 9.5% | NONE | | |
| K20 | A 0.4%, Q 0.4%, E 0.4%, K 98.7% | NONE | | |
| L21 | L 100% | A396 | A 100.0% | ϕ |
| | | W397 | W 100.0% | ϕ |
| | | A400 | A 98.3%, S 0.4%, V 1.3% | ϕ |
| P22 | P 100% | W397 | W 100% | ϕ |
| | | A400 | A 98.3%, S 0.4%, V 1.3% | ϕ |
| I23 | I 99.6%,V 0.4% | NONE | | |
| N24 | N 99.6%, I 0.4% | NONE | | |
| A25 | A 85.3%, P 14.2%, S 0.4% | NONE | | |
| L26 | L 100% | I432 | N 0.4%, I 99.1%, T 0.4% | ϕ |
| S27 | N 0.4%, S 99.6% | NONE | | |
| N28 | N 99.6%,Y 0.4% | NONE | | |
| S29 | P 1.7%, S 97.8%,T 0.4% | H428 | H 100% | B-B (H) |
| L30 | Q 0.4%, L 99.6% | A395 | A 100% | ϕ |
| | | A396 | A 100% | ϕ |
| | | I424 | I 91.4%, V 8.6% | ϕ |
| | | V494 | V 100% | ϕ |
| | | T500 | A 8.2%, I 0.4%, T 2.6%, V 88.8% | ϕ |
| | | H428 | H 100% | B-B (H) |

(*Continued*)

**Table S3** (*Continued*)

| Residues in Δ1 loop | | Interaction residues | | Bond type |
|---|---|---|---|---|
| L31 | L 100% | A393 | A 98.3%, G 0.9%, S 0.9% | ϕ |
| | | A396 | A 100% | ϕ |
| | | L492 | L 100% | ϕ |
| | | V494 | V 100% | ϕ |
| R32 | R 100% | G493 | G 100% | B-B (H) |
| H33 | R 0.4%, N 1.7%, H 96.1% | K491 | Q 0.4%, K 99.6% | B-R (H) |
| | | L492 | L 100% | B-B (H) |
| H34 | R 0.9%, H 99.1% | NONE | | |
| N35 | N 99.1%, S 0.9% | NONE | | |
| M36 | L 6.5%, K 0.4%, M 93.1% | C146 | D 0.4%, C 98.7%, V 0.9% | B-R (H) |
| | | V147 | I 3.9%, V 96.1% | ϕ |
| | | L492 | L 100% | ϕ |
| V37 | I 2.6%,V 97.4% | A393 | A 98.3%, G 0.9%, S 0.9% | ϕ |
| | | W397 | W 100% | ϕ |
| | | L492 | L 100% | ϕ |
| | | V144 | I 1.7%, V 98.3% | ϕ |
| Y38 | Y 100% | F145 | I 0.4%, F 99.6% | ϕ |
| | | | | B-B (H) |
| | | K155 | K 100% | ϕ |
| | | | | ϕ |
| A39 | A 95.7%, S 3.9%, V 0.4% | W397 | W 100% | ϕ |
| T40 | A 0.4%,T 99.6% | K141 | K 100% | B-R (H) |
| | | E143 | E 100% | B-B (H) |
| T41 | T 100% | NONE | | |
| S42 | S 99.1%,T 0.9% | NONE | | |
| R43 | R 100% | NONE | | |
| S44 | I 0.4%, S 99.6% | A157 | A 100% | B-R (H) |
| A45 | A 100% | L159 | L 95.3%, F 4.7% | ϕ |
| G46 | N 1.7%, C 0.9%, G 11.2%, S 84.5%, T 1.3%,V 0.4% | NONE | | |

**Notes:** Residues in red have one amino acid present at more than 99%.
Cyan denotes positions where all substitutions are all amino acids had polar or all non-polar side chains.

**Abbreviations:** B, Backbone; R, R-group; (H), H bond; ϕ, hydrophobic; I, Ionic; NONE, none observed in structure.

**Table S4** HCV Genotype 2a interaction between the Δ1 loop and the thumb domain and the Δ2 loop

| Residues in Δ1 loop | | Interacting residue(s) | | Bond type |
|---|---|---|---|---|
| L10 | L 100% | NONE | | |
| I 11 | I 100% | I138 | I 100% | ϕ |
| | | A140 | A 100% | ϕ |
| | | L159 | L 100% | ϕ |
| T12 | T 100% | NONE | | |
| P13 | P 100% | NONE | | |
| C14 | C 91.7% , S 8.3% | NONE | | |
| S15 | C 8.3%, G 8.3%, I 8.3%, S 75.0% | NONE | | |
| P16 | P 100% | NONE | | |
| E17 | E 100% | N142 | N 100% | R-R (H) |
| E18 | E 91.7%, K 8.3% | R401 | R 100% | I |
| E19 | E 100% | K154 | K 100% | I |
| K20 | K 100% | NONE | | |
| L21 | L 100% | A396 | A 100% | ϕ |
| | | W397 | R 8.3%, W 91.7% | ϕ |
| | | V400 | V 100% | ϕ |
| P22 | P 100% | W397 | R 8.3%, W 91.7% | ϕ |
| | | V400 | V 100% | ϕ |
| I23 | I 83.3%, F 8.3%,V 8.3% | NONE | | |
| N24 | N 100% | NONE | | |
| P25 | P 100% | NONE | | |
| L26 | L 100% | I432 | I 100% | ϕ |
| | | V435 | A 100% | ϕ |
| S27 | S 100% | NONE | | |
| N28 | N 100% | NONE | | |
| S29 | S 100% | NONE | | |
| L30 | L 100% | L392 | I 83.3%, L 8.3%, X 8.3% | ϕ |
| | | A395 | A 100% | ϕ |
| | | A396 | A 100% | ϕ |
| | | V424 | V 100% | ϕ |
| | | W500 | W 100% | ϕ |
| L31 | L 100% | L392 | I 83.3%, L 8.3%, X 8.3% | ϕ |
| | | A396 | A 100% | ϕ |
| | | L492 | L 100% | ϕ |
| | | A494 | A 100% | ϕ |
| R32 | R 100% | G493 | G 100% | B-B (H) |
| | | | | B-R (H) |
| Y33 | Y 100% | L492 | L 100% | B-B (H) |
| H34 | D 8.3%, H 91.7% | NONE | | |
| N35 | N 100% | NONE | | |
| K36 | K 100% | K491 | K 100% | B-R (H) |
| V37 | V 100% | V144 | V 100% | ϕ |
| | | A393 | A 91.7%, S 8.3% | ϕ |
| | | W397 | R 8.3%, W 91.7% | ϕ |
| | | L492 | L 100% | ϕ |
| Y38 | Y 100% | F145 | F 100% | ϕ |
| | | | | B-B (H) |
| | | V147 | V 100% | ϕ |

(*Continued*)

**Table S4** (*Continued*)

| Residues in Δ1 loop | | Interacting residue(s) | | Bond type |
|---|---|---|---|---|
| C39 | C 83.3%, S 16.7% | E143 | E 100% | B-R (H) |
| T40 | T 100% | K141 | K 100% | B-R (H) |
| | | E143 | E 100% | B-B (H) |
| | | | | B-R (H) |
| T41 | T 100% | NONE | | |
| S42 | A 8.3%, S 83.3%, T 8.3% | NONE | | |
| K43 | R 8.3%, K 91.7% | NONE | | |
| S44 | S 83.3%,T 16.7% | NONE | | |
| A45 | A 100% | A140 | A 100% | ϕ |
| | | L159 | L 100% | ϕ |
| S46 | S 100% | | | |

**Notes:** Residues in red have one amino acid present at more than 91%.
Cyan denotes positions where all substitutions are all amin acids had polar or all non-polar side chains.

**Abbreviations:** B, Backbone; R, R-group; (H), H bond; ϕ, hydrophobic; I, Ionic; NONE:, none observed in structure; X, sequencing issue precludes interpretation.