J. Biochem. **133**, 351–359 (2003)
DOI: 10.1093/jb/mvg047

Structure/Function Analysis of an RNA Aptamer for Hepatitis C Virus NS3 Protease

Satoru Sekiya, Fumiko Nishikawa, Kotaro Fukuda and Satoshi Nishikawa*

Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8566

Received December 6, 2002; accepted January 6, 2003

RNA aptamers that bind specifically to hepatitis C virus (HCV) NS3 protease domain (ΔNS3) were identified in previous studies. These aptamers, G9-I, -II, and –III, were isolated using an *in vitro* selection method and they share a common loop with the sequence 5′-GA(A/U)UGGGAC-3′. The aptamers are potent inhibitors of the NS3 protease *in vitro* and may have potential as anti-HCV compounds. G9-I has a 3-way stem-loop structure and was selected for further characterization using site-directed mutagenesis. Mutations or deletions in stem-loop II do not interfere with binding or inhibition of ΔNS3, but mutations or deletions in stem I and stem-loop III destroy the G9-I active conformation and interfere with inhibition of NS3 protease. A 51 nt fragment of 74 nt G9-I was identified (ΔNEO-III) as is the minimal fragment of G9-I that is an effective inhibitor of the NS3 protease. Tertiary interactions involving functionally important nucleotides were identified in the active structure of G9-I using nucleotide analog interference mapping (NAIM). Strong interferences were focused in the conserved loop involving stem-loop III and stem I. For example, analog-interference caused at A(+8) and C(+24)-G(–36) base pair implied an A-minor motif involving the intramolecular base triple A(+8)·C(+24)-G(–36), which is further supported by mutagenesis. These results suggested the interaction of stem I and stem-loop III is essential for the function of G9-I aptamer.

Key words: aptamer, HCV, NS3 protease, nucleotide analog interference mapping, RNA structure.

Abbreviations: HCV, hepatitis C virus; NS3, non-structural protein 3; NAIM, nucleotide analog interference mapping; NTPαS, phosphorothioate nucleotide; 2APαS, 2-aminopurine riboside; PRαS, purine riboside; IαS, inosine riboside; m^5UαS, 5-methyluridine. Enzymes: Taq DNA polymerase [EC 2.7.7.7] T7 RNA polymerase [EC 2.7.7.6]; T4 polynucleotide kinase [EC 2.7.1.78].

Hepatitis C virus (HCV) is the major etiological agent of post-transfusion non-A, non-B hepatitis. HCV infection is a serious disease worldwide, which can develop into chronic hepatitis, liver cirrhosis or hepatocellular carcinoma. The number of HCV carriers has increased to approximately 300 million worldwide; nevertheless, effective antiviral drugs have not yet been developed, and the principal anti-HCV drugs are based on interferon. HCV has a single positive-stranded RNA genome approximately 9.6 kb in length. The viral genome encodes a large polyprotein approximately 3,010 amino acids in length. The precursor polyprotein is processed into structural (core protein C; envelope glycoproteins E1 and E2) and non-structural (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) proteins by host signal peptidases and two viral proteases, NS2-3 and NS3. NS3 has a trypsin-like serine protease activity in its N-terminal domain and a helicase activity in its C-terminal domain. The protease activity is essential for processing viral non-structural proteins NS3, NS4A, NS4B, NS5A, NS5B and is necessary for viral replication *in vivo*. Thus, it is likely that a specific inhibitor of the NS3 protease activity would have potential as an effective anti-HCV drug (reviewed in Ref. *1*).

To find potential anti-HCV drugs, we applied *in vitro* selection (*2–4*) from several kinds of random RNA pools (*5*, *6*). RNA aptamers were identified that were specific for the protease domain of NS3 (ΔNS3; amino acids 1027–1218) (*7*, *8*). These aptamers bind to ΔNS3 with high specificity (K_d = 10 nM) and inhibit approximately 90% of the protease activity of ΔNS3 and full-length NS3 (NS3 amino acids 985–1647 fused with maltose binding protein, ref. *9*). The ΔNS3-specific RNA aptamers belong to three classes, designated G9-I, II, and III (Fig. 1A), which have a common sequence 5′-GA(A/U)UGGGAC-3′. This sequence forms a loop structure when modeled with the Mulfold program (*10*), and this secondary structure has been confirmed by nuclease partial digestion in G9-I (*7*). Previous mutagenesis studies of G9-I demonstrated that stem I and stem-loop III are required for NS3 protease inhibition but stem II is not, that stem I and stem-loop III may interact. Recently we found that the G9-I aptamer binds to a patch of positive amino acid residues on the surface of NS3 (*11*).

This study investigates the functional structure of G9-I using site-specific mutagenesis and nucleotide analog interference mapping (NAIM). Mutations and deletions

*To whom correspondence should be addressed. Tel: +81-298-61-6085, Fax: +81-298-61-6159, E-mail: satoshi-nishikawa@aist.go.jp

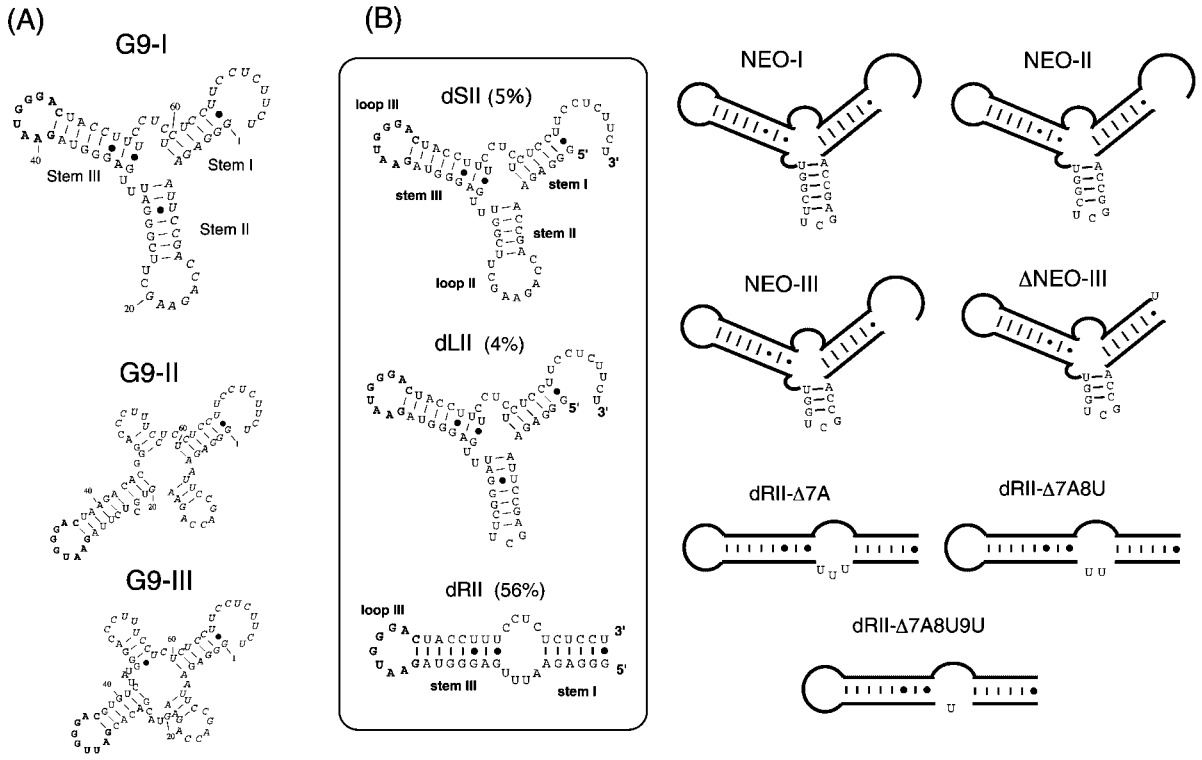

Fig. 1. **Secondary structures of G9 aptamers and deletion variants of G9-I.** (A) The secondary structure models of G9 aptamers (G9-I, II and III) were predicted using the Mulfold program. The portion of the aptamer that was randomized and then subjected to *in vitro* selection is shown in bold or normal letters. Bold letters indicate the consensus sequence (5′-GANUGGGAC-3′, N=A or U) common to all G9 aptamers. Italic letters indicate the non-randomized sequence regions that flank the randomized N30 core. (B) In the previous studies, dSII (U-A and U-G deleted in stem II), dLII (most of loop II deleted) and dRII were constructed and characterized. Figures in parentheses indicate residual proteolytic activity (*7*). The mutations in dSII and dLII were combined into the mutant NEO-I. Stem II in NEO-I was further deleted one bp at a time to form 2 additional mutants NEO-II and NEO-III. Eight nucleotides at the 3′ terminus of NEO-III were deleted to construct ΔNEO-III. In dRII, all of stem-loop II was removed except the two nucleotides of the terminal base pair, and stem-loop III is connected to stem I with junction single strands (*7*). Three dRII mutants, dRII-Δ7A, dRII-Δ7A8U and dRII-Δ7A8U9U, were constructed by deleting A, AU, or AUU residues in the 5′ side of the spacer region.

in the stem-loop II region do not interfere with the catalytic function of the aptamer; however, stem loop II plays a role in stabilizing the tertiary structure of G9-I. A 51-nucleotide (nt) fragment (ΔNEO-III) lacking stem-loop II was the minimal fragment of 74 nt G9-I that is an efficient inhibitor of the NS3 protease. Tertiary interactions and functional groups of G9-I were identified using NAIM with thiophosphate nucleotide analogs (*12*). NAIM is a powerful tool that has been used to identify structure-function relationships in catalytic RNAs such as *Tetrahymena* group I intron (*12–14*), RNase P (*15*, *16*), and hairpin ribozyme (*17*). In this report, NAIM is used to predict important intramolecular base triple interactions between the conserved loop and stem I in the active G9-I structure. The results suggest that A(+8) docks into the minor-groove of C(+24)-G(–36) and forms a tertiary interaction with ΔNS3 protease.

EXPERIMENTAL PROCEDURES

Preparation of Deletion Variants of G9-I RNA Aptamer—G9-I deletion mutants, 3′ tail–deleted mutant ΔG9-I and site-specific mutants in C(+24)-G(–36) were constructed by PCR mutagenesis using substituted (+) and/or (–) primers and G9-I as a template. The mutated PCR products were subcloned into pGEM-T Easy vector© (Promega) and confirmed by sequencing using a *BigDye* Terminator Cycle Sequencing Kit© (Applied Biosystems) on a 377 automatic sequencer (Applied Biosystems).

The double-stranded DNA for transcription was generated by PCR (EX Taq, Takara) with a (+) primer at the T7 promoter sequence and a (–) primer in the plasmid vector. The PCR product was used as a template for *in vitro* transcription by T7 RNA polymerase (T7 Ampliscribe Kit, Epicentre Technologies, USA). Reactions were allowed to proceed for 3 h at 37°C, treated with DNase I for 15 min at 37°C, and terminated with an equal volume of stop solution (50 mM EDTA, 9 M urea, 0.1% xylene cyanol, 0.1% bromophenol blue). The mixture was heated at 90°C for 2 min, snap-cooled on ice, and separated by electrophoresis on an 8% polyacrylamide gel (PAGE) containing 7 M urea. After visualization by UV irradiation, the appropriate band was excised from the gel and eluted with 0.3 M sodium acetate. The RNA was recovered by ethanol precipitation and quantified.

Protease Inhibition Assay Using a Synthetic Substrate—ΔNS3 (amino acids 1027–1218) was expressed in *Escherichia coli* HB101 and purified as described previously (*8*). Protease inhibition assay was essentially the same as described previously (*7*). Usually protease ΔNS3

and RNA aptamer were mixed at 25°C in a reaction mixture containing 50 mM Tris-HCl (pH 7.8), 30 mM NaCl, 5 mM $CaCl_2$, and 10 mM DTT. A dansyl-labeled peptide substrate corresponding to the NS5A-5B junction (S1; 86 μM) was added to this solution, which included ΔNS3 protease (0.8 μM) and G9-I aptamer or its variants in the presence of NS4A peptide (P41; 22.5 μM). The reaction mixture (25 μl) was incubated for 40 min at 25°C, then quenched with NaOH (0.25 N). The reactions were analyzed by reverse phase HPLC (TSKgel ODS-120T; 4 ml) with a linear gradient of acetonitrile in 50 mM ammonium acetate (pH 6.5). The ratio of cleaved product to uncleaved substrate was calculated from the fluorescence signal (excitation, 340 nm; emission, 510 nm) using a fluorescence monitor (*9*). For comparison between G9-I, ΔG9-I derivatives, and RIII, ΔNS3 and RNA were mixed at molar ratios of 1:2, 1:5, or 1:10 in the presence of P41 (7.9 μM).

Nucleotide Analog Interference Mapping (NAIM)—NAIM experiments were carried out using the procedure of Ryder and Strobel (*18*).

Preparation of Nucleotide Analog-Incorporated RNA Aptamers—Twelve phosphorothioate nucleotide analogs were obtained from Glen Research as follows: NTPαS (AαS, GαS, CαS, UαS and their base analogs), 2-aminopurine riboside (2APαS), purine riboside (PRαS), inosine (IαS), 5-methyluridine (m^5UαS), and 2′-deoxy derivatives (dAαS, dGαS, dCαS, dUαS). Each nucleotide analog was incorporated at a level of approximately 5% (except dCαS at approximately 10%) into RNA aptamers by T7 RNA polymerase during *in vitro* transcription. Under these conditions, each RNA molecule contains on average one analog base. In the case of dNTPαS incorporation, 1 mM Mn^{2+} was added to improve efficiency of incorporation (*19*). Transcription and RNA purification were carried out as described above. RNA transcripts were dephosphoryrated with alkaline phosphatase (New England Biolabs), and 5′-radiolabeled with T4 polynucleotide kinase (Takara) using 30 μCi of [γ^{32}P]ATP (Amersham). The 5′-labeled RNA was purified as described above and incorporated radioactivity was quantified.

Preparation of Bound RNA to ΔNS3—The 5′-end labeled RNA (300 kcpm) and 20 nM ΔNS3 were incubated for 30 min at 25°C under the standard binding conditions described previously. ΔNS3-bound RNA was separated from unbound RNA by filtration through a 0.45 μm nitrocellulose filter (Millipore). The RNA-protein complex on the filter was recovered by soaking the filter in elution buffer [3.5 M Urea, 0.3 M sodium acetate (pH 5.2), 2.5 mM EDTA (pH 8.0)] for 2 min at 90°C, then snap-cooling on ice. The eluted material was washed with phenol-chloroform, precipitated with ethanol, and radioactivity was quantified.

Iodide Cleavage of Nucleotide Analog Incorporated RNA Aptamers—The bound (+) fraction of 5′ labeled thiophosphate-modified RNA was cleaved with 10 mM iodine/ethanol in a reaction mixture containing 5 μg yeast tRNA (Boehringer Mannheim), 3.5 M urea, 0.025% xylene cyanol, and bromophenol blue. Mixtures were incubated for 2 min at 94°C and snap-cooled on ice. A control cleavage reaction was carried out using untreated 5′-labeled thiophosphate-modified RNA. The cleavage products were analyzed by electrophoresis on 8% and 20% PAGE containing 7 M urea, and the band intensity for each position was quantified using a BAS 2000 image analyzer (Fuji Film).

Quantitation of Nucleotide Analog Interference—Nucleotide analog interference was quantified at each aptamer position using the ratio: [intensity of RNA bound with ΔNS3 (+)]/[intensity of RNA untreated with ΔNS3 (–)]. Interference values were normalized by dividing the value at each position by the average value at all positions. Nucleotide analog interferences at each position were the average of three independent experiments; the standard deviation was in all cases within 10%. If the interference value of AαS, GαS, CαS, or UαS was less than 0.6, the thiophosphate at that position was considered to cause binding interference. The extent of base-modification or 2′-deoxy interference was calculated as follows: the averaged normalized value of nucleotide analog (δαS) at each position (M) was divided by the averaged normalized value of the parental nucleotide at the corresponding position (S) to obtain (M)/(S). When [(M)/(S)] was less than 0.5, the analog at that position was considered to cause interference. The results are summarized in Table 1.

RESULTS

Mutational Analysis at Stem-Loop II in G9-I Aptamer—Stem I and loop III [5′-GA(A/U)UGGGAC-3′] are conserved in all three G9 aptamers and are required for inhibition of NS3 protease activity (*7*). Three deletion mutants of stem-loop II were characterized to determine its role in aptamer function. dSII (a 2 bp deletion in stem II) and dLII (deletion of most of loop II) inhibit NS3 protease to the same extent as G9-I (Fig. 1B). This result suggested that stem-loop II is dispensable for the inhibitory function of G9-I. However, dRII (Fig. 1B), which lacks the entire stem-loop II except its terminal base pair, was a less effective inhibitor of NS3 than G9-I (~45% inhibition relative to G9-I) (*7*).

The structure/function relationships of G9-I were tested further to define the minimum structural requirements for a fully active aptamer. Three deletion mutants of dRII, designated dRII-Δ7A, dRII-Δ7A8U, and dRII-Δ7A8U9U, were constructed and analyzed (Fig. 1B). The results show that dRII-Δ7A, dRII-Δ7A8U, and dRII-Δ7A8U9U partially inhibited ΔNS3 protease activity (15–25% inhibition relative to G9-I) and were less effective than dRII (Fig. 2). These mutations change the length of the 5′ spacer region of dRII but do not restore inhibitory function to the level of G9-I. Thus, it is likely that stem II clamps stem I and stem III to support proper folding of the aptamer.

Additional aptamers were constructed with deletions in the stem-loop II region (Fig. 1B). The mutations in dSII and dLII were combined into the mutant NEO-I. Stem II in NEO-I was further deleted one bp at a time to form two additional mutants NEO-II and NEO-III. NEO-I, II, and III were efficient inhibitors of ΔNS3 protease (≈95% inhibition relative to G9-I) (Fig. 3). These results suggest that the duplex structure of stem II is functionally required for maximal inhibition of ΔNS3 protease; however, the sequence of stem-loop II is tolerant to muta-

Table 1. **Quantification of NAIM interference in G9-I.**

Base	Position	A	dA/A	2AP/A	PR/A	G	dG/G	I/G	C	dC/C	U	dU/U	m^5U/U
G	*–36*					1.00	0.61	*0.48*					
A	–35												
G	–34												
A	–33					0.98	0.79	0.62					
A	–32												
U	–31												
U	–30												
C	–29												
C	–28												
G	–27					1.06	1.09	0.88					
A	–26												
C	–25								1.74	1.09			
C	–24								1.77	1.10			
A	–23	1.48	1.08	0.89	0.79								
G	–22					1.48	1.39	1.13					
A	–21	1.05	1.11	1.22	2.53								
A	–20	1.02	1.17	0.95	0.91								
G	–19					1.15	0.86	1.24					
C	–18								1.08	0.89			
U	–17										1.05	1.09	0.99
U	–16										1.32	0.90	0.88
C	–15								0.98	1.11			
G	–14					0.96	1.16	1.10					
G	–13					1.12	0.83	0.94					
G	–12					0.84	0.98	1.38					
A	–11	0.74	1.69	1.38	1.20								
U	–10										0.93	1.02	1.08
U	*–9*										1.05	*0.37*	0.89
U	–8										1.06	1.21	0.95
G	–7					1.28	0.71	0.77					
A	*–6*	0.77	*0.31*	0.90	0.68								
G	–5					0.66	0.85	1.32					
G	*–4*					*0.52*	*0.33*	0.75					
G	–3					0.95	0.82	0.95					
U	–2										1.00	0.54	0.95
A	–1	0.73	1.27	1.44	0.75								
G	+1					1.23	0.81	0.67					
A	+2	0.89	0.48	0.90	1.09								
A	*+3*	0.83	*0.10*	1.16	0.82								
U	*+4*										0.91	*0.11*	1.14
G	+5					0.93	1.04	0.78					
G	+6					0.95	0.95	0.63					
G	*+7*					*0.44*	0.84	0.73					
A	*+8*	1.08	*0.16*	*0.06*	*0.52*								
C	*+9*								0.88	*0.15*			
U	*+10*										1.12	*0.45*	0.84
A	+11	1.04	0.83	0.92	0.68								
C	+12								0.88	1.31			
C	+13								0.86	1.31			
U	+14										1.05	0.95	0.99
U	+15										0.95	1.05	1.15
U	+16										1.00	1.00	0.98
C	+17								0.87	0.52			
C	*+18*								0.86	*0.44*			
U	*+19*										*0.42*	*0.50*	1.10
C	+20								0.73	1.01			
U	+21										0.90	1.02	1.10
C	+22								0.93	1.27			
U	+23										1.07	0.91	0.90
C	*+24*								0.79	*0.24*			
C	+25								0.86	0.86			

Measurements were based on three independent experiments, and the standard deviation was 10% or less. The abbreviations 2AP, PR, I, and m^5U denote 2-aminopurine riboside, purine riboside, inosine, and 5-methyluridine, respectively. The interference values from G(–36) to C(+25) are indicated. A value of 1 indicates no interference and the value decreases with increasing interference. Strong interference effects are indicated by italics.

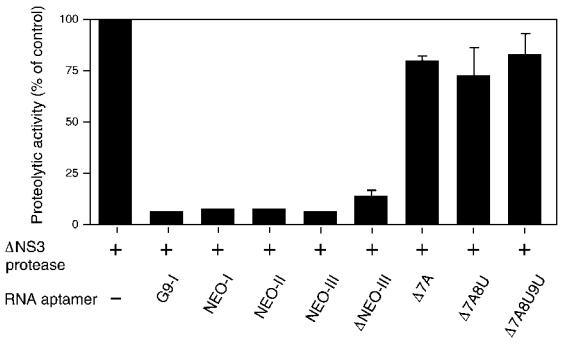

Fig. 2. **Protease inhibition activity of deletion variants of G9-I.** Protease assay was carried out as described in Materials and Methods. Proteolytic activity is normalized relative to a reaction containing only ΔNS3. Proteolytic activity varies slightly with the preparation of ΔNS3 and other factors.

tion and stem-loop II does not directly influence binding of the aptamer to ΔNS3.

In a previous study, the 3′-end boundary of G9-I was defined, and it was concluded that the 3′ tail of the molecule is not required for ΔNS3 binding (*7*). Most of 3′ tail region was deleted from NEO-III to create the 51 nt mutant ΔNEO-III (Fig. 2). ΔNEO III inhibits the ΔNS3 protease nearly as efficiently as G9-I (Fig. 2) and binds to full-length NS3 protease (data not shown). No more deletion of stem II was possible without altering the secondary structure of stem I and stem-loop III. Thus, ΔNEO-III is the smallest fragment of G9-I that is fully active as a specific inhibitor of ΔNS3.

Nucleotide Analog Interference Mapping (NAIM)—Tertiary interactions between G9-I and ΔNS3 protease were characterized by NAIM (*12*) in order to identify important functional groups in the RNA aptamer. Twelve nucleotide phosphorothioate analogs including parental phosphorothioates (AαS, GαS, CαS, UαS) and their base analogs, 2-aminopurine riboside (2APαS), purine riboside (PRαS), inosine (IαS), 5-methyluridine (m^5UαS), and 2′-deoxy derivatives (dAαS, dGαS, dCαS, dUαS) were partially incorporated in aptamers synthesized by *in vitro* transcription. For the purpose of this analysis, G9-I nucleotides were numbered with coordinate +1 at the beginning of the conserved loop (Fig. 3A). Negative coordinates are therefore upstream of loop III.

An analog interference effect was detected from position G(–36) to C(+25) by comparing band intensity on denaturing PAGE (Fig. 3B). The thiophosphate interference and/or modification interference was quantified as described in Materials and Methods and the results are shown in Table 1. Analog interferences were identified at 13 positions of G9-I (Fig. 3 and Table 1, italic letters/numbers). The positions that cause distinct interference were strongly concentrated in the conserved loop region, confirming the importance of this region (*7*). This moiety is clearly essential for binding of G9-I to ΔNS3. Nucleotide analog interference also occurs in stem I and the two junction regions between stems. Analog interference was not detected in stem II. This result is consistent with the results of mutagenesis on stem-loop II, which indicated that this region is not required for G9-I binding to ΔNS3.

Thiophosphate Interference—Thiophosphate interference was observed at G(+7) in the conserved loop, G(–4) and U(+19). This result suggests that non-bridging R_p oxygens of 5′-phosphates at these positions contribute to the tertiary structure of the G9-I RNA-ΔNS3 protein complex, but the exact role played by these phosphate oxygens is not revealed by this data. The phosphate oxygens could directly contact the ΔNS3 protein, could

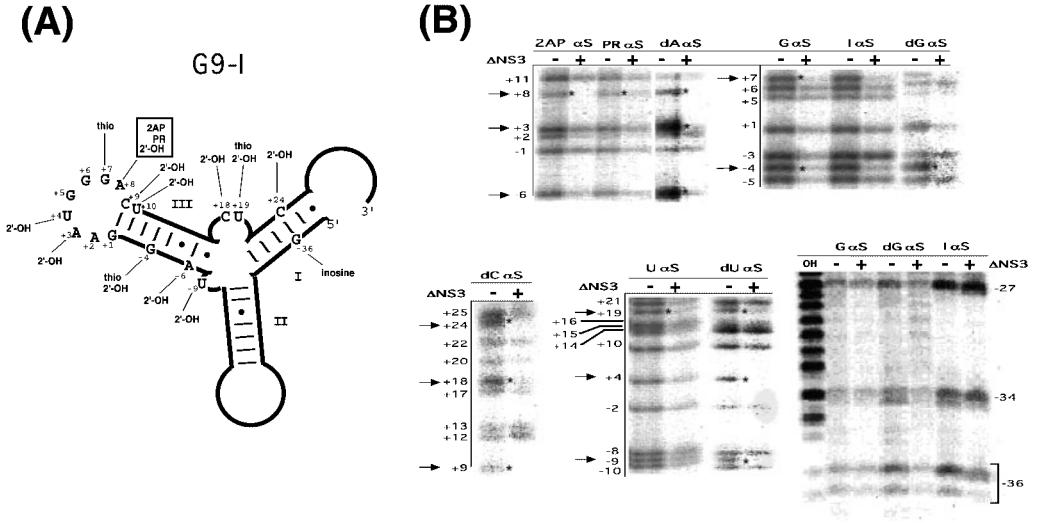

Fig. 3. **Nucleotide analog interference mapping (NAIM) of G9-I aptamer.** (A) Interference positions and residues. The N30 randomized region (*7*) is shown in blue and red. Red letters indicate the conserved loop sequence of G9 aptamers. The canonical base pairs are shown as dashed lines and dot indicates a G·U wobble base pair. The nucleotides affected by analog incorporation are indicated with letters. "2′-OH" denotes an interaction with 2′ hydroxyl group. "thio" denotes the thio interference by substitution of Rp oxygen with Rp-S. "2AP", "PR" and "inosine" denote interference by substitution with 2-aminopurine riboside, purine riboside and inosine riboside, respectively. The strong interference effects at A(+8) are boxed. Nucleotide numbers start with +1 at the 5′ end of the conserved sequence. (B) Gel images. An asterisk indicates retardation of a band and an arrow indicates a base discussed in the text. Analog abbreviations are 2-amino purine riboside (2APαS), purine riboside (PRαS) and inosine (IαS).

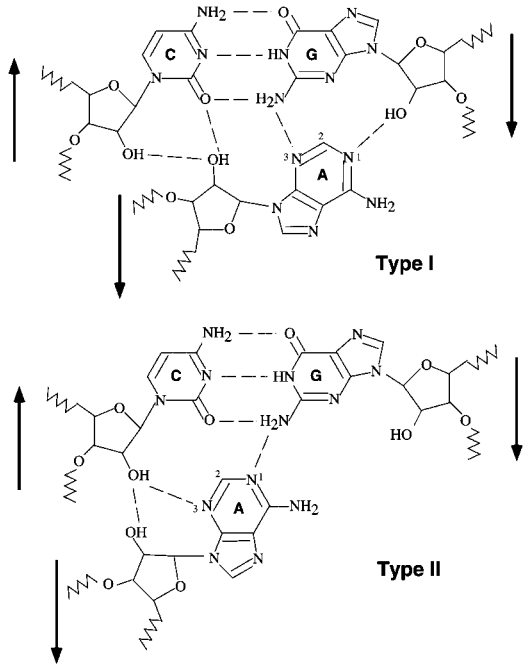

Fig. 4. **The examples of adenosine docking into minor-groove C-G base pair (Type I and Type II).** Arrow indicates the strand polarity. A dotted line indicates hydrogen bonding. The 2-NH_2 group interferes by steric hindrance with 2-aminopurine docking into the minor-groove C-G base pair.

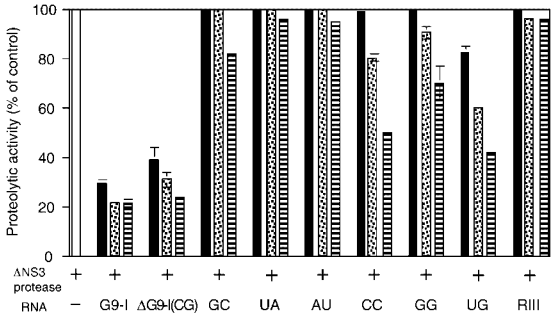

Fig. 5. **Protease inhibition activity of variants at C(+24)-G(-36) base pair.** Proteolytic activity was compared at a ΔNS3:RNA molar ratio of 1:0 (white bar), 1:2 (black bar), 1:5 (dotted bar) and 1:10 (striped bar) in stem-loop III (RIII), ΔG9-I and its derivatives.

coordinate a divalent cation with or without a water molecule, or could be involved in a hydrogen bonding interaction.

G9 aptamers I, II and III, which share the conserved loop sequence, inhibit ΔNS3 protease to a similar extent. Secondary structures of the G9 aptamers involve 3-way and 4-way junctions as shown in Fig. 1A; these structures suggest that tertiary structure and folding may distinguish G9-I and G9-II, III. In particular, a divalent metal ion (Mg^{2+}, Ca^{2+}) is essential for the active G9 aptamer conformation, and the structure of G9-I, but not of G9-II and G9-III, favors Ca^{2+} binding (data not shown). It is possible that a Ca^{2+} binding pocket is present in G9-I. For example, Ca^{2+} might efficiently neutralize electrostatic repulsion between the helices of stem I and stem-loop III and thus stabilize the tertiary structure of the aptamer.

2′-Deoxy Analog Interference—2′-Deoxy analog substitution caused strong interference (normalized interference <0.2) at nucleotide positions A(+3), U(+4), A(+8), and C(+9) in the conserved loop (Table 1). These residues are thought to be a part of a hydrogen-bonding network that organizes higher-order structure in the conserved loop. This structure might involve a ribose zipper (*20*) similar to structures observed in many catalytic RNAs. A large interference effect was also detected at C(+24) in stem I. Other nucleotide positions where 2′-deoxy analog substitution caused interference were G(–4) and A(–6) in stem III, U(–9) at the junction of stem II and stem III, C(+18) and U(+19) at the junction of stem I and stem III, respectively. These 2′ hydroxyl residues may stabilize the interaction of stem I and stem III through water-mediated or nonwater-mediated hydrogen bonds, forming triple helices as observed in X-ray crystallography of the vitamin B12 aptamer (*21*). The 2′ deoxy modifications could potentially cause local structural perturbation by altering sugar pucker (*22*). It has not yet been possible to assign the counterparts interacting with these 2′ hydroxyl groups.

Base-Modification Interference and Docking Adenosine A(+8)—Substitution of 2-aminopurine (2AP) at position A(+8) in the conserved loop caused very strong interference (0.06), and substitution of purine riboside (PR) caused a slight decrease in binding activity. Thus, at position A(+8), the 2-amino moiety of 2AP alters the tertiary interactions between G9-I and ΔNS3. The specific interference at A(+8) by dA, 2AP, and PR analogs implies some tertiary interactions such as a typical adenosine docking (Fig. 4, Refs. *23* and *24*), which forms a base triple in the tertiary arrangement of the aptamer. It is possible that the corresponding harbor site is the C(+24)-G(–36) base pair, because 2′ deoxy interference at C(+24) and inosine interference at G(–36) are detected. These results suggest that there is some interaction between stem I and stem-loop III of G9-I in its active structure.

Deletion and mutation analyses indicate that the G-triplet G(+5)G(+6)G(+7) in the conserved loop of G9-I is necessary for ΔNS3 binding (*7*); however, 2′ hydroxyl and 2 amino group analogs did not cause interference at those positions. In addition, 5-methyluridine ($m^5U\alpha S$) incorporation at U(+4) did not cause interference, suggesting that methyl substitution at this position does not affect ΔNS3 binding by G9-I.

A Minor-Groove Base Triple A(+8)·C(+24)-G(–36)—A model for A(+8) docking into C(+24)-G(–36) shown in Fig. 4 was tested using a series of base pair mutants of G9-I (Fig. 5). The nonessential 3′ tail was deleted and the terminal bp of stem I was changed from the 5′G·U wobble base pair to 5′G-C to maintain the same secondary structure as G9-I. This derivative was called ΔG9-I (CG) and the following base pair mutants were introduced into the C(+24)-G(–36) derivative: Watson-Crick (W-C) base pair; G(+24)-C(–36), U-A, A-U (GC, UA, AU); and non W-C base pair, C·C, G·G, U·G (CC, GG and UG). The ΔG9-I(CG) binds to ΔNS3 and inhibits protease activity nearly as efficiently as G9-I. In contrast, the W-C base pair mutants GC, UA and AU have almost no inhibitory activity under similar reaction conditions. The amount of residual inhibitory activity is comparable to the variant

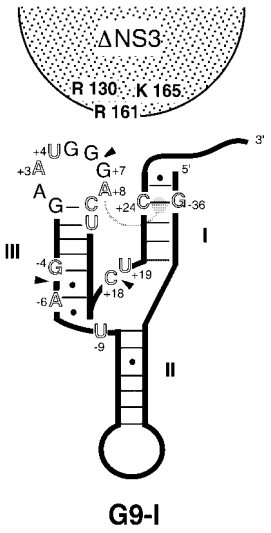

Fig. 6. **Schematic model of the interaction between G9-I and ΔNS3.** Italic letters indicate the nucleotides with 2′ hydroxyl group interactions. An arrow indicates the adenosine that docks into the minor-groove the C(+24)-G(-36) base pair. Arrowheads indicate the positions where thiophosphate interference occurred. R130, R161 and K165 indicate the amino acids of ΔNS3 that bind to G9-I (*11*).

RIII, which has only a portion of stem-loop III; however, the GC mutant has a weak inhibitory activity at high concentration ([ΔNS3]: [RNA] = 1:10). Therefore, the C(+24)-G(–36) base pair plays an essential role in the active conformation of G9-I. Non W-C base pair mutants with one point mutation possess weak inhibitory activity, but their activity is higher than the W-C base pair mutants. The wobble base pair mutant UG, which is closer in spatial orientation to CG, has approximately 50% of the activity of the CG aptamer. Thus, C(+24)-G(–36) tolerates a non W-C base pair substitution to a greater extent than the W-C base pair substitutions G-C, U-A, or A-U. This base pair specificity is consistent with A(+8) docking into the minor-groove of C(+24)-G(–36) to form a minor-groove base triple (*24*). This structure would mediate an interaction between stem I and stem-loop III.

Two major types of base triples have been identified in which adenosine in a single-stranded region docks into the helix minor-groove (*23–25*), and these are designated type I and type II (Fig. 4). These base triples were identified in the crystal structures of the hammerhead ribozyme (*26*) and the P4-P6 domain of the *Tetrahymena* group I intron (*20*). In the type I base triple, the N1, C2, N3, and 2′ hydroxyl groups of the adenosine moiety lie along the minor groove face of the helix and contact the entire minor groove surface of the contact base pair including both ribose sugars. Alternatively, in the type II base triple, the N1, C2, N3, and 2′ hydroxyl groups of the adenosine contact approximately one-half of the minor groove of a base pair. It is not yet concluded whether type I or type II is accepted in A(+8)·C(+24)-G(–36).

DISCUSSION

The protease domain of HCV NS3 encodes a trypsin-like serine protease activity. To date, several aptamers for trypsin-like proteases, such as thrombin etc. have been identified (*27–29*). X-ray crystallographic analysis of the thrombin-DNA aptamer complex indicated that the aptamer binds at the anion binding exosite, which involves an Arg cluster that interacts with fibrinogen (*30*). Activated protein C (APC) is also a trypsin-like enzyme that regulates the blood clotting cascade (*31*). X-ray crystallography of APC indicated that an anion binding exosite involving a Lys cluster lies on the surface of the protease. Studies with APC mutants suggested that this region binds the APC aptamer (Nishikawa *et al*., unpublished data). A recent study also identified the region of ΔNS3 that binds G9-I using alanine scanning mutagenesis targeting positively charged residues on the surface of ΔNS3 (*11*). The results indicate that Arg161, Arg130 and Lys165 bind the aptamer and are clustered in a highly solvent-exposed surface near the catalytic site of ΔNS3. Arg161 may play a key role in binding G9-I, because it is a candidate for a specific interaction with the P6 position of the protease substrate (*32–34*).

Previous studies led to three conclusions concerning the structure and function of aptamer G9-I: (i) the conserved loop 5′-GA(A/U)UGGGAC-3′ is important for aptamer function; (ii) both stem I and stem-loop III are essential for aptamer function, and probably have some contacts with each other; (iii) stem-loop II is dispensable for the function. This study shows that stem-loop II does not contact ΔNS3, but it supports the tertiary structure of the aptamer by clamping stem I and stem III. A 51 nt deletion variant (ΔNEO-III) of the 74 nt G9-I was created, in which the 3′ tail and stem-loop II were deleted, and this mutant showed efficient ΔNS3 protease inhibition activity (Fig. 2). The interaction between ΔNS3 and G9-I was also analyzed by NAIM, which identified functional groups and nucleotides of the RNA aptamer (Fig. 3A). NAIM provided preliminary information on functionally important groups of G9-I, but a detailed interpretation of the NAIM data is not yet possible and will require additional experiments.

The A-minor motif is a specific interaction between an adenosine moiety and the minor groove of an RNA helix. This motif is a universal base triple forming a compact active structure that is stabilized by interbase, ribose-ribose and ribose-base hydrogen bonds. The structure forms between a remote single-strand RNA region and a duplex RNA region, supporting essential tertiary interactions (*23*). The A-minor motif has been reported widely in the structure of RNA molecules and RNA-protein complexes.

NAIM and further mutagenesis data in this study propose that G9-I has an essential base triple involving positions A(+8)·C(+24)-G(–36), which forms a triple helix between stem I and loop III. A(+8) is flipped out to dock into the C(+24)-G(–36) base pair, which allows a loop structure that can contact ΔNS3. The RIII lacks stem I and is incapable of A(+8) docking; this may cause the conserved loop to be flexible in structure and strongly reduce the inhibitory activity of this RNA. A one base pair deletion in stem III also strongly decreased the functional activity of the aptamer (*7*), which may indicate that A(+8) is unable to dock in this mutant because of altered distance between A(+8) and its receptor site. The conserved loop sequence is required for interaction between G9-I

and ΔNS3 protease, and extension of the terminus of stem I also interrupted binding to ΔNS3 (unpublished data). Therefore, ΔNS3 is located in proximity to the conserved loop and the terminus of stem I, but the ΔNS3 binding site for G9-I is likely to be limited to the conserved loop portion of the RNA aptamer.

The data presented here suggest that there are four distinct 2′ hydroxyl-mediated interactions in the conserved loop of G9-I involving A(+3), U(+4), A(+8) and C(+9). In addition, the Rp oxygen at G(+7) may also have a significant interaction. The G9-I binding site on ΔNS3 includes a potential anion-binding region involving the positively charged amino acids R130, R161, K165 (*11*). These results suggest that the R_p oxygen at G(+7) contacts ΔNS3. However, to date, the corresponding thiophosphate interferences have not been assigned. It is supposed that A(+8) participates in adenosine minor groove docking into the C(+24)-G(–36) base pair. Other functions in loop III of the active G9-I complex have not yet been defined. NAIM data also indicate that 2′ hydroxyl interactions and thio effects exist in stem III and junctions. Interactions were observed at G(–4), A(–6), and U(+19), which may stabilize the formation of triple helices between stem III and single strand region C(+17)–C(+20) and mediate an interaction between stem I and stem-loop III. The 2′ hydroxyl at U(–9) may facilitate a turn in the RNA single strand, which could change the direction of the strand in the active structure. These data are summarized in Fig. 6, which shows a schematic model of putative interactions between G9-I and ΔNS3. For the data obtained from mutagenesis and NAIM in the active structure, our proposed model consists of mutual interaction between stem I and stem-loop III, but the idea that stem I and stem-loop III interact with NS3 protein individually is not excluded. Additional experiments to confirm this model are required; such experiments will compare the structure/function relationships of G9-I, II and III and examine G9-I structure/function at higher resolution.

An important future goal will be to isolate a competitive inhibitor of NS3 protease. One novel approach will be to conjugate G9-I with a small molecule general serine protease inhibitor. This would be a new type of protease inhibitor that could have novel and useful properties. NS3 protein includes both a protease and helicase domain. This study shows that the 3′ tail of G9-I is not required for interaction with ΔNS3. Therefore, it may be possible to design a dual functional aptamer that inhibits both NS3 protease and helicase using the 3′ terminal stem I of G9-I to connect a linker. These research projects will hopefully facilitate discovery of protease inhibitors that have potential as antiviral compounds against HCV.

We thank Dr. Joonsung Hwang for preparation of ΔNS3. This research was supported by the grants from the Ministry of Economy, Trade and Industry (METI) of Japan, especially by ″R & D for Evolutionary Molecular Engineering.″ F.N. and K.F. are grateful for the New Energy and Industrial Technology Development Organization (NEDO) fellowship.

REFERENCES

1. De Francesco, R., Pessi, A., and Steinkuhler, C. (1998) The hepatitis C virus NS3 proteinase: structure and function of a zinc-containing serine proteinase. *Antivir. Ther.* **3**, 99–109
2. Ellington, A.D. and Szostak J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822
3. Joyce, G.F. (1989) Amplification, mutation and selection of catalytic RNA. *Gene* **82**, 83–87
4. Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510
5. Urvil, P.T., Kakiuchi, N., Zhou, D-.M., Shimotohno, D-.M., Kumar, P.K.R., and Nishikawa, S. (1997) Selection of RNA aptamers that bind specifically to the NS3 protease of hepatitis C virus. *Eur. J. Biochem.* **248**, 130–138
6. Kumar, P.K.R. Machida, K., Urvil, P.T., Kakiuchi, N., Vishnuvardhan, D., Shimotohno, K., Taira, K., and Nishikawa, S. (1997) Isolation of specific RNA aptamers to NS3 protein of Hepatitis C Virus from a pool of completely random RNA. *Virology* **237**, 270–282
7. Fukuda, K., Vishnuvardhan, D., Sekiya, S., Hwang, J., Kakiuchi, N., Taira, K., Shimotohno, K., Kumar, P.K.R., and Nishikawa, S. (2000) Isolation and characterization of RNA aptamers specific for the hepatitis C virus NS3 protease. *Eur. J. Biochem.* **267**, 3685–3694
8. Vishnuvardhan, D., Kakiuchi, N., Urvil, P.T., Shimotohno, K., Kumar, P.K.R., and Nishikawa, S. (1997) Expression of highly active recombinant NS3 protease domain of hepatitis C virus in *E. coli*. *FEBS Lett.* **402**, 209–212
9. Kakiuchi, N., Hijikata, M., Komoda, Y., Tanji, Y., Hirowatari, Y., and Shimotohno, K. (1995) Bacterial expression and analysis of cleavage activity of HCV serine proteinase using recombinant and synthetic substrate. *Biochem. Biophys. Res. Commun.* **210**, 1059–1065
10. Zuker, M. (1989) Computer prediction of RNA structure. *Methods Enzymol.* **180**, 202–287
11. Hwang, J.S., Fauzi, H., Fukuda, K., Sekiya, S., Kakiuchi, N., Shimotohno, K., Taira, K., and Nishikawa, S. (2000) The RNA aptamer-binding site of hepatitis C virus NS3 protease. *Biochem. Biophys. Res. Commun.* **279**, 557–562
12. Strobel, S.A. and Shetty, K. (1997) Defining the chemical groups essential for *Tetranymena* group I intron function by nucleotide analog interference mapping. *Proc. Natl Acad. Sci. USA* **94**, 2903–2908
13. Ortoleva-Donnelly, L., Szewczak, A.A., Gutell, R.R., and Strobel, S.A. (1998) The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA* **4**, 498–519
14. Strauss-Soukup, J.K. and Strobel, S.A. (2000) A chemical phylogeny of Group I Introns based upon Interference Mapping of a Bacterial Ribozyme. *J. Mol. Biol.* **302**, 339–358
15. Hardt, W.D., Erdmann, V.A., and Hartmann, R.K. (1996) Rp-deoxy-phosphorothioate modification interference experiments identify 2′-OH groups in RNase P RNA that are crucial to tRNA binding. *RNA* **2**, 1189–1198
16. Heide, C., Pfeiffer, T., Nolan, J.M., and Hartmann, R.K. (1999) Guanosine 2-NH2 groups of *Escherichia coli* RNase P RNA involved in intramolecular tertiary contacts and direct interactions with tRNA. *RNA* **5**, 102–116
17. Ryder, S.P. and Strobel, S.A. (1999) Nucleotide Analog Interference Mapping of the Hairpin ribozyme: implications for secondary and tertiary structure formation, *J. Mol. Biol.* **291**, 295–311
18. Ryder, S.P. and Strobel, S.A. (1999) Nucleotide analog interference mapping. *Methods* **18**, 38–50
19. Conrad, F., Hanne, A., Gaur, R.K., and Krupp, G. (1995) Enzymatic synthesis of 2′-modified nucleic acids: Identification of important phosphate and ribose moieties in RNase P substrates. *Nucleic Acids Res.* **23**, 1845–1853
20. Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. (1996) Crystal

structure of a group I ribozyme domain: Principles of RNA packing. *Science* **273**, 1678–1685
21. Sussman, D., Nix, J.C., and Wilson, C. (2000) The structural basis for molecular recognition by the vitamin B12 RNA aptamer. *Nat. Struct. Biol.* **7**, 53–57
22. Saenger, W. (1984) *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.
23. Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B., and Steitz, T.A. (2001) RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc. Natl Acad. Sci. USA* **98**, 4899–4903
24. Doherty, E.A., Batey, R.T., Masquida, B., and Doudna, J.A. (2001) A universal mode of helix packing in RNA. *Nat. Struct. Biol.* **8**, 339–343
25. Leontis, N.B. and Westhof, E.Q. (1998) Conserved geometrical base-pairing patterns in RNA. *Quart. Rev. Biophys.* **31**, 399–455
26. Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994) Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature* **372**, 111–113
27. Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H., and Toole, J.J. (1992) Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **355**, 564–566
28. Lin, Y., Padmapriya, A., Morden, K.M., and Jayasena, S.D. (1995) Peptide conjugation to an *in vitro*-selected DNA ligand improves enzyme inhibition. *Proc. Natl Acad. Sci. USA* **92**, 11044–11048
29. Gal, S.W., Amontov, S., Urvil, P.T., Vishnuvardhan, D., Nishikawa, F., Kumar, P.K., and Nishikawa, S. (1998) Selection of a RNA aptamer that binds to human activated protein C and inhibits its protease function. *Eur. J. Biochem.* **252**, 553–562
30. Padmanabhan, K., Padmanabhan, K.P., Ferrara, J.D., Sadler, J.E., and Tulinsky, A. (1993) The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer, *J. Biol. Chem.* **268**, 17651–17654
31. Stubbs, M.T. and Bode, W. (1995) The clot thickens: Clues provided by thrombin structure. *Trends Biochem. Sci.* **20**, 23–28
32. Yao, N., Reichert, P., Taremi, S.S., Prosise, W., and Weber, P.C. (1999) Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* **7**, 1353–1363
33. Cicero, D.O., Barbato, G., Koch, U., Ingallinella, P., Bianchi, E., Nardi, M.C., Steinkuhler, C., Cortese, R., Matassa, V., DeFrancesco, R., Pessi, A., and Bazzo, R. (1999) Structural characterization of the interactions of optimized product inhibitors with the N-terminal proteinase domain of the Hepatitis C virus (HCV) NS3 protein by NMR and modelling studies. *J. Miol. Biol.* **289**, 385–396
34. Laplante, S.R., Cameron, D.R., Aubry, N., Lefebvre, S., Kukolj, G., Maurice, R., Thibeault, D., Lamarre, D., and Llinas-Brunet, M. (1999) Solution structure of substrate-based ligands when bound to Hepatitis C virus NS3 protease domain. *J. Biol. Chem.* **274**, 18618–18624