Understanding the RNA-Specificity of HCV RdRp: Implications for Anti-HCV Drug Discovery

Jinyoung Kim[†] and Youhoon Chong[†,‡,*]

[†]*Department of Biomedical Science and Technology,* [‡]*Department of Molecular Biotechnology, Konkuk University, Seoul 143-701, Korea.* [*]*E-mail: chongy@konkuk.ac.kr*

Received September 30, 2005

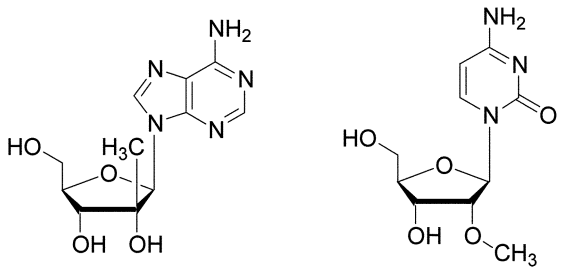
Unlike other viral polymerases, HCV RNA-dependent RNA polymerase (RdRp) has not been successfully inhibited by nucleoside analogues presumably due to its strong substrate specificity for RNA. Thus, in order to understand the RNA-specificity of HCV RdRp, the structural characteristics of the active site was investigated. The hereto unknown 2-OH binding pocket at the active site of RdRp provides invaluable implication for the development of novel anti-HCV nucleoside analogues.

Key Words : HCV RNA-dependent RNA polymerase (RdRp), Nucleoside inhibitors, Molecular modeling, Drug design

Introduction

Hepatitis C virus (HCV)[1] is the main agent responsible for contagious hepatitis not due to the "classical" hepatitis viruses A and B.[2] There is presently no vaccine or widely effective treatment for the resulting liver damaging infection that can eventually lead to death.[3] Because current HCV therapies are inefficient, costly and prone to producing debilitating side effects, HCV infection has been viewed as a growing threat to human health worldwide and therefore, new antiviral therapies are clearly needed.

Among the viral proteins encoded by HCV, the NS5B RNA-dependent RNA polymerase (RdRp)[4-7] is an essential enzyme for viral replication and therefore represents a valid target for therapeutic intervention through the design of specific inhibitors. As the knowledge regarding HCV RdRp structure and function accumulates, development of specific inhibitors has been greatly promoted. However, in general, HCV RdRp is not sensitive to most nucleoside/tide analogs that inhibit other viral polymerases: it was not inhibited by most of the known inhibitors for DNA-dependent DNA polymerases (DdDp) or reverse transcriptases (RT).[8-10] In this context, it is worth while to note that HCV RdRp is a RNA-specific enzyme whereas DdDp and RT take deoxyribonucleotides (DNA) as substrates. More interestingly, the most potent inhibitors of HCV RdRp discovered to date are 2'-substituted nucleoside analogues (2'-*C*-methyladenosine and 2'-*O*-methylcytidine) which have the '*ribo*'-sugar moieties in common (Figure 1). Unfortunately, however, there has been little efforts to correlate the HCV RdRp's insensitivity to DNA polymerase inhibitors with its RNA specificity. Thus, it is of our great interest to find the characteristic structural features of the active site of HCV RdRp which might be constructed in such a way as to selectively accommodate the RNA substrate. Eventually, understanding the RNA-specificity of HCV RdRp would serve as the key for the future development of novel anti-HCV nucleoside analogues. Herein, we report our recent efforts to develop potent and selective inhibitors of HCV RdRp by elucidating the structural characteristics of the active site of HCV RdRp.



2'-*C*-Methyladenosine 2'-*O*-Methylcytidine

Figure 1. 2'-Substituted 'ribo'-nucleoside analogues.

Experimental Section

All molecular modeling studies of the enzyme-substrate complexes were carried out using SYBYL 7.0 (Tripos Inc. St. Louis, Mo.) on linux enterprise operating system.

Polymerase Structure. Three dimensional crystal structures of HIV-1 reverse transcriptase and HCV RNA-dependent RNA polymerase were obtained from the protein data bank (PDB code 1RTD and 1QUV, respectively). The two structures were superimposed by fitting seven amino acid residues around the active site with RMSD (Root-Mean-Square-Deviation) of 0.89 Å: Arg72 (RT)-Arg158 (RdRp), Asp110-Asp220, Asp113-Cys223, Ala114-Phe224, Tyr115-Asp225, Asp185-Asp318, Asp186-Asp319. DNA template-primer duplex, two Mg atoms and TTP (thymidine triphosphate) in RT were then merged into the HCV RdRp active site. DNA template-primer duplex except three pairs from the active site (U-G-G: A-C-C, primer:template) and template overhang were removed for the sake of simplicity, and the DNA pairs were modified into the corresponding RNAs. Gästeiger-Hückel charge[11,12] was given to the ligand

with formal charges (+2) to two Mg atoms in the active site. Then, Kollman-All-Atom charges[13,14] were loaded to the enzyme site from the biopolymer option in SYBYL 7.0. In order to eliminate local strains resulting from merging inhibitors, residues inside 8 Å from the merged inhibitors were annealed[15] until the energy change from one iteration to the next was less than 0.05 kcal/mol (hot region: 8 Å, interesting region: 12 Å). The annealed enzyme-inhibitor complexes were minimized by using Tripos force field for 5,000 iterations.

Modeling Binding Modes of 2'-*O*-Methylcytidine and 2'-*C*-Methyladenosine at the Active Site of RdRp. Crystallographic structure of D-adenosine triphosphate (ATP, 3'-endo conformation) was used as a template for construction of 3'-endo conformation of 2'-*O*-methylcytidine triphosphate (2'-*O*-Me-CTP) and 2'-*C*-methyladenosine triphosphate (2'-*C*-Me-ATP). In contrast, the D-adenosine triphosphate was modified into its 2'-endo conformation first to provide the three dimensional models of 2'-*O*-methylcytidine triphosphate (2'-*O*-Me-CTP) and 2'-*C*-methyladenosine triphosphate (2'-*C*-Me-ATP) in 2'-endo conformation. The modeled nucleoside monophosphate analogues were superimposed into the substrate (UTP, uridine triphosphate) bound at the active site of RdRp, merged and replaced with UTP. The resulting RdRp: 2'-*O*-Me-CTP and RdRp: 2'-*C*-Me-ATP complexes were energy-minimized by using the same protocol as described above.

2'-OH Binding Site. Crystallographic water molecules, RNA duplex and two Mg atoms were removed from the RdRp:ligand complexes and the Connolly surface of the enzyme was generated by using the 'Fast Connolly Channel' option in the 'Molcad Surface Generation' menu of Sybyl 7.0. The surface, thus generated, was overlaid with the enzyme-RNA duplex complex containing water molecules for investigation of the 2'-OH binding pocket.

Results and Discussion

Crystal Structure of HCV RdRp. HCV belongs to the *Flaviviridae* family, which contains a positive-sense single-stranded RNA. The HCV genomic RNA contains a single open reading frame which encodes a polyprotein precursor of ~3010-3033 amino acids. The polyprotein precursor is further processed by cellular signal peptidase and virally encoded proteases to generate three structural (core, E1, E2) and seven non-structural (P7, NS2, NS3, NS4A, NS4B, NS5A, NS5B) proteins. HCV NS5B is the viral RNA-dependent RNA polymerase (RdRp) that is responsible for viral genome replication. HCV RdRp contains a total of 591 amino acids and the mutational analysis has suggested that the catalytic core of HCV RdRp is contained in the *N*-terminal ~540 amino acids.[16] Three groups have reported NS5B structures in the absence of substrates, at 2.8 Å,[17] 1.9 Å,[18] and 2.5 Å[19] resolution. In all structures reported, the HCV RdRp protein displays the fingers, palm, and thumb subdomains characteristic of all known RNA and DNA polymerases as shown in Figure 2. However, unlike other

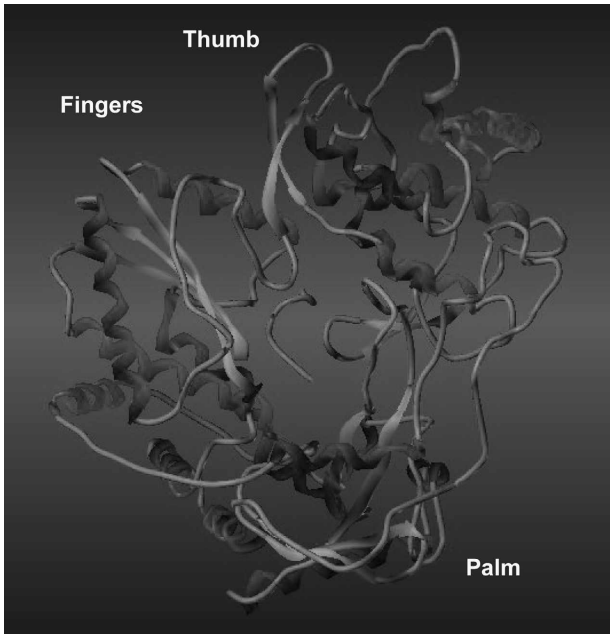


Figure 2. HCV RdRp has three subdomains like other polymerases, but the fingers and thumb regions completely encircles the active site.

polymerases that resemble a right hand and bind DNA or RNA in a cleft between the fingers and thumb, HCV NS5B has an enlarged thumb and expanded fingers region that completely encircles the active site (Figure 2).

HCV RdRp has a "closed" Conformation. It is well known that viral DNA and RNA polymerases undergo conformational changes upon binding of primed template and substrate at the active site, which accompanies closing the gap between the fingers and the thumb subdomains to make the enzyme catalytically competent.[20] In the absence of proper substrates, all other polymerases have been observed in an "open" conformation, with the fingers rotated away from the thumb, and they only adopt the "closed" conformation in the presence of a primed template and the proper NTP (nucleoside triphosphate) that is to be next inserted. The crystal structures of HCV RdRp[17-19] do not have primed template and substrate, but they represent the catalytically competent conformations of the enzyme because HCV RdRp includes an encircled active site with an overall globular shape instead of the typical U shape found in other polymerases (Figure 2). Such an encircled active site is due to the extensive interactions between the fingers and the thumb subdomains, which close the gap between the two subdomains.[17-19]

Construction of HCV RdRp-template/Primer-rNTP Ternary Complex. Although none of the crystal structures of HCV RdRp was solved with RNA or metal ions bound to the active site, the residues involved in binding of RNA and metals have been highlighted by superimposition with the ternary complex of HIV reverse transcriptase.[20] For example, two aspartates, Asp220 and Asp318, were found to be the key residues for nucleotide metal binding, consistent with the biochemical data obtained from site-directed

mutagenesis of HCV RdRp.[21] Thus, RNA duplex and metal ions, which are missing from the crystal structure of HCV RdRp but required for ligand binding studies, could be modeled into the HCV RdRp by adopting those of HIV-1 RT after superimposition of the two enzymes. Similarly, both Bressanelli[17] and Lesburg[19] reported that the DNA duplex of HIV-1 RT (PDB code 1RTD) could be used as a template to model RNA duplex into HCV RdRp structures without major conformational changes in the fingers and thumb domains. In our study, DNA template-primer was modeled into HCV RdRp by overlaying HCV RdRp (PDB code 1QUV) with HIV-1 RT (PDB code 1RTD) around seven matching amino acid pairs (Figure 3, Arg158-Arg72; Asp318-Asp185; Asp319-Asp186; Asp220-Asp110; Cys223-Asp113; Phe224-Ala114; Asp225-Tyr115) between the two enzymes (RMSD 0.89 Å). The modeled DNA duplex was truncated from the active site leaving three base pairs, which were modified into the corresponding RNA's. The ternary complex of RT showed that, although the DNA appears to be predominantly in the B-form, the base pairs close to the polymerase active site have an A-like conformation with a widened minor groove, which is the standard conformation of RNA duplex. Therefore, A-form DNA duplex around the active site of RT could be successfully used as a template for RNA duplex of HCV RdRp without significant conformational change upon modification. Like other polymerases, HCV RdRp also requires divalent cations for catalysis. The most preferred metals have been identified as Mg^{2+} at concentrations of ~5–10 mM or Mn^{2+} at lower concentrations (< 1 mM). The involvement of divalent cations in the reactions catalyzed by HCV RdRp can be explained by the two metal ion mechanism established for other DNA and RNA polymerases.[22] As mentioned above, these metals, together with several conserved aspartates (Asp220, Asp319, Asp318), are responsible for coordination of nucleotide binding to the enzyme, and thus, modeled into the active site of the enzyme accordingly (Figure 3).

RNA Specificity of HCV RdRp. HCV NS5B is indeed an RNA-dependent RNA polymerase: it does not contain DNA-dependent RNA polymerase activity or RNA-dependent DNA polymerase (reverse transcriptase) activity because it cannot utilize DNA as template or deoxynucleotide as substrate.[10,23]

Active Site Analysis of HCV NS5B Polymerase: 2-OH Binding Pocket. Ever since the landmark discovery of acyclovir, viral polymerases have been very popular targets for antiviral drugs, which resulted in the wide use of nucleoside analogues for treatment of hepatitis B virus and HIV infections.[24] However, it should be noted that the antiviral nucleoside research has been focused on the inhibition of HIV reverse transcriptase or HBV polymerase both of which take DNA as substrates. The three dimensional structure of HIV-1 RT[20] clearly explains how the enzyme achieves its DNA-specificity: one of the key active site residues, Tyr115 does not allow RNA to bind the active site by acting as a "steric gate" against the 2'-OH group of the RNA (Figure 4). Naturally, nucleoside analogues with 2'-OH group have not drawn specific attention whereas 2'-deoxy or 2',3'-dideoxy nucleoside analogues played major roles in antiviral nucleoside research. However, in the context of the RNA-specificity of HCV NS5B RdRp, there must be specific interaction between the 2'-OH group of the RNA substrate and the active site residues. Thus, it was our intention to understand the structural background which

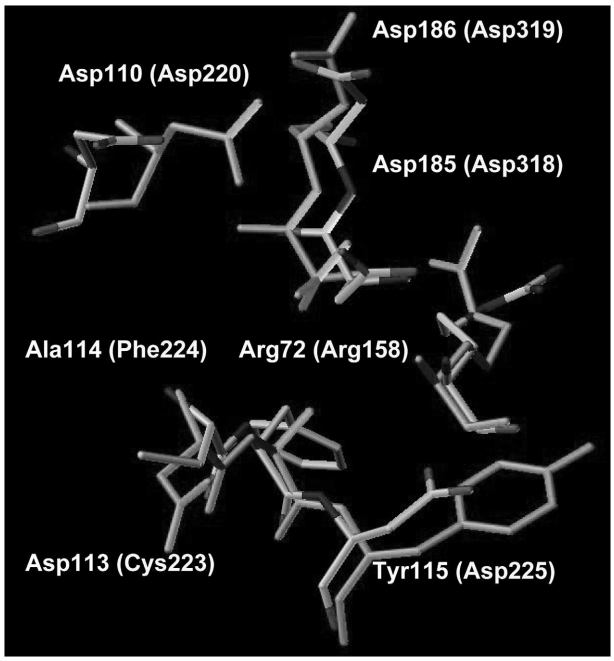


Figure 3. Seven matching amino acid pairs between HIV-1 RT (green color, white legend) and HCV RdRp (atom type color, yellow legend).

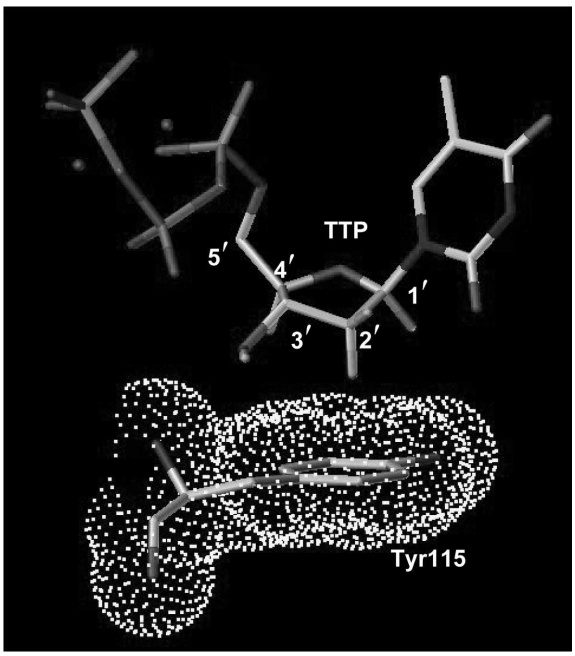


Figure 4. The 'steric gate' of HIV-1 RT: The bulky aromatic side chain of Tyr115 lies right below the 2'-H to prevent ribonucleotides from binding the active site.

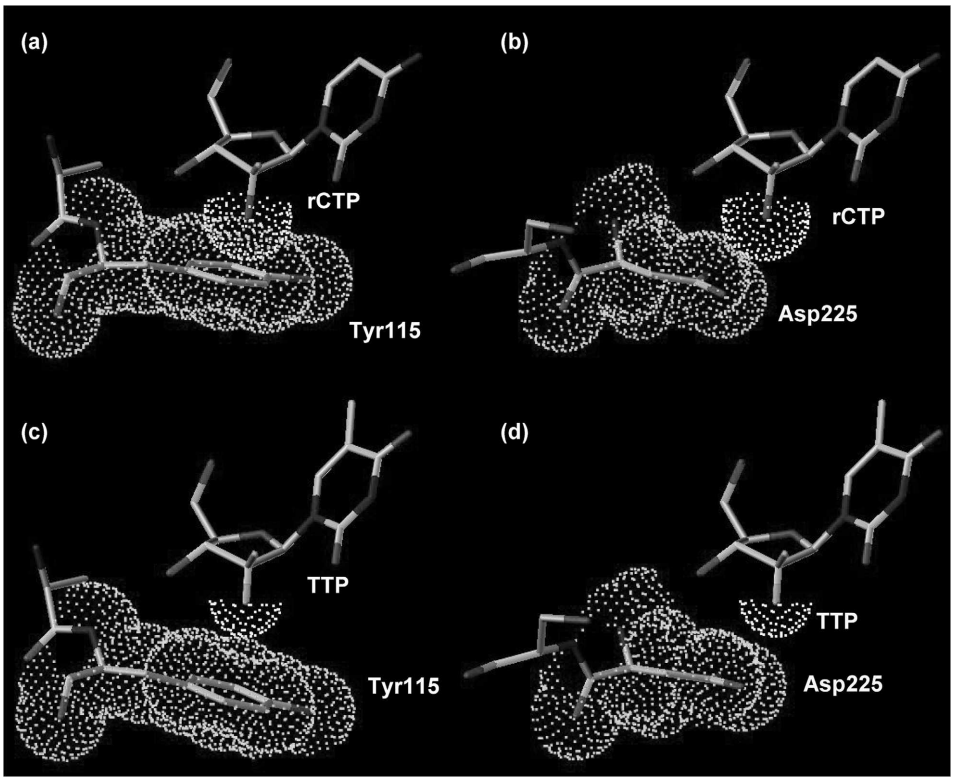


Figure 5. Asp225 of HCV RdRp (b, d) corresponds to the steric gate (Tyr115) of HIV-1 RT (a, c), but the small side chain of Asp225 provides enough space to the active site for binding of ribonucleotides. Ribonucleoside rCTP experiences steric hindrance with the aromatic side chain of Tyr115 of RT (a), but deoxyribonucleoside TTP is well separated from Tyr115 (c).

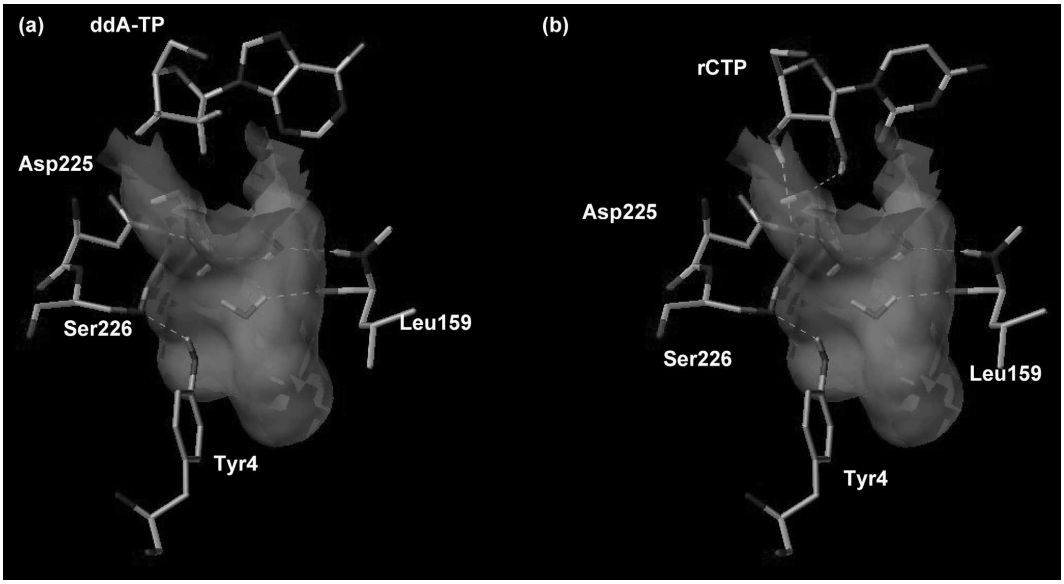
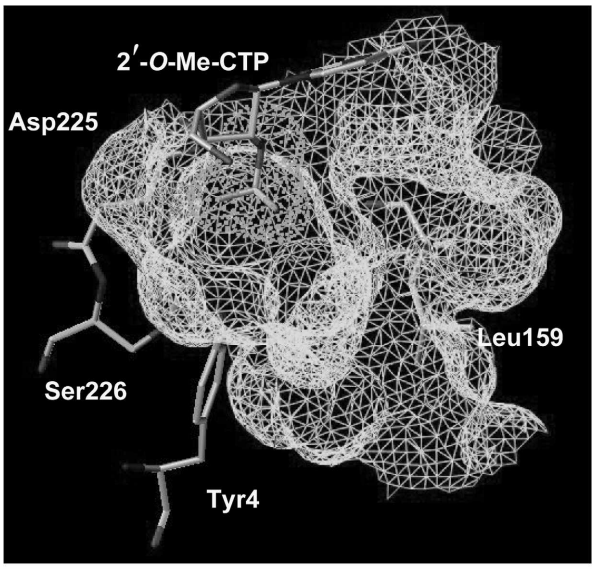


Figure 6. (a) Anti-HIV drug ddA-triphosphate is located at the active site of HCV RdRp. Nucleoside triphosphate does not have any specific interaction with the pocket residues or water molecules inside the pocket (yellow surface). (b) '2'-OH pocket' (yellow surface) composed of Asp225, Ser226, Tyr4 and Leu159 contains water molecules which interact with both the 2'-OH of rCTP and pocket residues through hydrogen bonding networks.

supports the RNA-specificity of HCV RdRp for the purpose of designing more potent and specific HCV RdRp inhibitors.

Sequence alignment showed that the "steric gate" of HIV-1 RT (Tyr115) corresponds to Asp225 in HCV RdRp (Figure 3), but the small side chain of Asp225 provides an open space at the active site enough to accommodate the ribose sugar moiety of RNA substrate (Figure 5). The side chain carboxylate of Asp225 is in close proximity to the 3-OH of rNTP (ribonucleoside triphosphate), which is reminiscent of the hydrogen bonding interaction between the amide



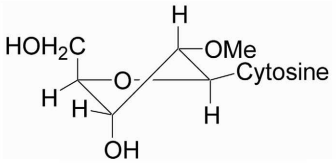
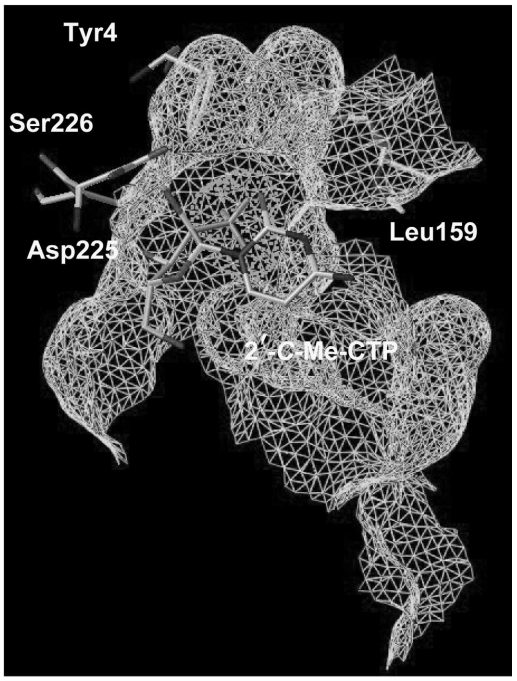


Figure 7. 2'-*O*-methyl group of 2'-*O*-Me-CTP can be located inside the '2'-OH pocket' through van der Waals interaction with the pocket residues. 2'-substituents with higher steric demands than methyl looks to be accommodated by the pocket. South (2'-endo) conformation of 2'-*O*-Me-CTP (right) was used for binding study.

backbone of Tyr115 and 3'-OH of NTP in HIV-1 RT structure (Figure 5). However, no active site residue could be found close to the 2'-OH group of rNTP for specific interaction, which did not support our assumption that the 2'-OH group must be involved in the RNA-specificity of HCV RdRp. Thus, the only way by which the 2'-OH group could achieve a specific interaction with the enzyme residue seemed to be by way of bridging water molecules inside the active site. To our surprise, crystallographic water molecules could be found around the 2'-OH group of the RNA substrate and more interestingly, those water molecules were tightly held at the middle of an 11 Å long channel located next to Asp225 (Figure 6). The hereto unknown active site channel is composed of 11 residues, which are in excellent three dimensional arrangements to hold water molecules. The water molecules act either as a hydrogen bond donor or an acceptor to the 2'-OH group of the RNA substrate (Figure 6). In addition to the critical role in the binding of the RNA substrate, the pocket provides more important insights into the design of potent and specific inhibitors of HCV RdRp. First, we could understand the characteristic binding modes of 2'-*O*-Me-CTP and 2'-*C*-Me-ATP, which are the most potent two nucleoside analogues developed so far. Recently, two separate studies[25,26] suggested that 2'-*O*-methylcytidine and 2'-*C*-methyladenosine would have south (2'-endo) and north (3'-endo) conformations, respectively (Figures 7 and



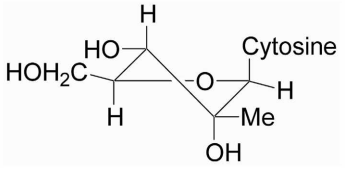


Figure 8. Top view of the binding mode of 2'-*C*-Me-CTP at the active site of HCV RdRp. 2'-*C*-methyl substituent is well-oriented just inside the 2'-OH pocket for stable binding interaction (van der Waals) with the pocket residues. North (3'-endo) conformation of 2'-*C*-Me-CTP (right) was used for binding study.

8). It was intriguing that only the south (2'-endo) conformation of 2'-*O*-methylcytidine could be successfully located at the active site of HCV RdRp (Figure 7) while the north (3'-endo) conformation experienced severe steric hindrance between its sugar moiety and the active site residues. In addition, 2'-*O*-methyladenosine (2'-endo) showed its 2'-*O*-methyl group located in the middle of the 2'-OH pocket of the enzyme, which provides an interesting insight into designing more potent HCV RdRp inhibitors (Figure 7). As the 11 Å long pocket has residues with polar side chains in the middle for coordination of water molecules, longer 2'-*O* substituents capable of electrostatic interaction with the pocket residues would have greater chance to selectively bind to the active site of HCV RdRp. In contrast, 2'-*C*-methyl group of 2'-*C*-methyladenosine looks like to be the substituent of choice because the 2'-*C*-methyl snuggly fits into a small groove at the top of the 2'-OH pocket when the sugar moiety adopts north (3'-endo) conformation (Figure 8), which is in good accordance with the poor antiviral activity of 2'-*C*-ethyladenosine.[25] The south (2'-endo) conformation of 2'-C-methyladenosine results in an abortive binding mode due to the lack of any specific interaction with the active site residues.

Conclusion

Unlike other viral polymerases, HCV RNA-dependent RNA polymerase (RdRp) has not been successfully inhibited by nucleoside analogues presumably due to its strong substrate specificity for RNA. In this study, the RNA-specificity of HCV RdRp was investigated by molecular modeling studies to give information about the existence of the hereto unknown 2-OH binding pocket at the active site of RdRp, which provides invaluable implication for the development of novel anti-HCV nucleoside analogues.

Acknowledgments. This work was supported by the faculty research fund of Konkuk University in 2005.

References

1. Alter, M. J.; Kruszon-Moran, D.; Nainan, O. V.; McQuillan, G. M.; Gao, F.; Moyer, L. A.; Kaslow, R. A.; Margolis, H. S. *N. Engl. J. Med.* **1999**, *341*, 556-562.
2. Choo, Q.-L.; Kuo, G.; Weiner, A. J.; Overby, R. L.; Bradley, D. W.; Houghton, M. *Science* **1989**, *244*, 359-362.
3. Saito, I.; Miyamura, T.; Ohbayashi, A.; Katayama, T.; Kikuchi, S.; Watanabe, Y.; Koi, S. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6547-6549.
4. Hwang, S. B.; Park, K.-J.; Kim, Y.-S.; Sung, Y. C.; Lai, M. M. C. *Virology* **1997**, *227*, 439-446.
5. Behrens, S.-E.; Tomei, L.; De Francesco, R. *EMBO J.* **1996**, *15*, 12-22.
6. Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; Andoh, T.; Yoshida, I.; Okayama, H. *J. Virol.* **1991**, *65*, 1105-1113.
7. Kim, J.; Lee, M.; Kim, Y.-Z. *Bull. Korean. Chem. Soc.* **2005**, *26*, 285.
8. Ishii, K.; Tanaka, Y.; Yap, C. C.; Aizaki, H.; Matsuura, Y.; Miyamura, T. *Hepatology* **1999**, *29*, 1227-1235.
9. Ferrari, E.; Wright-Minogue, J.; Fang, J. W.; Baroudy, B. M.; Lau, J. Y.; Hong, Z. *J. Virol.* **1999**, *73*, 1649-1654.
10. Johnson, R. B.; Sun, X. L.; Hockman, M. A.; Villarreal, E. C.; Wakulchik, M.; Wang, Q. M. *Arch. Biochem. Biophys.* **2000**, *377*, 129-134.
11. Gästeiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219-322.
12. Purcell, W. P.; Singer, J. A. *J. Chem. Eng. Data* **1967**, *12*, 235-246.
13. Blaney, J. M.; Weiner, P. K.; Dearing, A.; Kollman, P. A.; Jorgensen, E. C.; Oatley, S. J.; Burridge, J. M.; Blake, C. C. F. *J. Am. Chem. Soc.* **1982**, *104*, 6424-6434.
14. Wipff, G.; Dearing, A.; Weiner, P. K.; Blaney, J. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1983**, *105*, 997-1005.
15. Chong, Y.; Borroto-Esoda, K.; Furman, P. A.; Schinazi, R. F.; Chu, C. K. *Antivir. Chem. Chemother.* **2002**, *13*, 115-128.
16. Tomei, L.; Vitale, R. L.; Incitti, I.; Serafini, S.; Altamura, S.; Vitelli, A. *J. Gen. Virol.* **2000**, *81*, 759-767.
17. Bressanelli, S.; Tomei, L.; Roussel, A.; Incitti, I.; Vitale, R. L.; Mathieu, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13034-13039.
18. Ago, H.; Adachi, T.; Yoshida, A.; Yamamoto, M.; Habuka, N.; Yatsunami, K. *Structure Fold Des.* **1999**, *7*, 1417-1426.
19. Lesburg, C. A.; Cable, M. B.; Ferrari, E.; Hong, Z.; Mannarino, A. F.; Weber, P. C. *Nat. Struct. Biol.* **1999**, *6*, 937-943.
20. Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. *Science* **1998**, *282*, 1669-1675.
21. Lohmann, V.; Korner, F.; Herian, U.; Bartenschlager, R. *J. Virol.* **1997**, *71*, 8416-8428.
22. Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395-17398.
23. Yamashita, T.; Kaneko, S.; Shirota, Y.; Qin, W.; Nomura, T.; Kobayashi, K. *J. Biol. Chem.* **1998**, *273*, 15479-15486.
24. Chu, C. K. *Antiviral Nucleosides: Chiral Synthesis and Chemotherapy*; Elsevier: Amsterdam, Neth., 2003.
25. Tomassini, J. E.; Getty, K.; Stahlhut, M. W.; Shim, S.; Bhat, B.; Eldrup, A. B.; Prakash, T. P.; Carroll, S. S.; Flores, O.; MacCoss, M.; McMasters, D. R.; Migliaccio, G.; Olsen, D. B. *Antimicrob. Agents Chemother.* **2005**, *49*, 2050-2058.
26. Eldrup, A. B.; Prhave, M.; Brooks, J.; Bhat, B.; Prakash, T. P.; Song, Q.; Bera, S.; Bhat, N.; Dande, P.; Dan Cook, P.; Bennett, C. F.; Carroll, S. S.; Ball, R. G.; Bosserman, M.; Burlein, C.; Colwell, L. F.; Fay, J. F.; Flores, O. A.; Getty, K.; LaFemina, R. L.; Leone, J.; MacCoss, M.; McMasters, D. R.; Tomassini, J. E.; Von Langen, D.; Wolanski, B.; Olsen, D. B. *J. Med. Chem.* **2004**, *47*, 5284-5297.