Mutational and Structural Analysis of HCV Non Structural Protein 2 (NS2) Revealed Genotype Specific Motif in TMD3 of the Protein

Sadia Anjum, Faryal Mahwish, Tahir Ahmad, and Muhammad Sohail Afzal

***Abstract*—Hepatitis C virus (HCV) nonstructural protein 2 (NS2) is a transmembranal protein with a hydrophobic amino-terminal subdomain containing up to three putative transmembrane segments and a carboxyterminal cytoplasmic domain. It is believed that NS2 plays a crucial role in major processes during the propagation of virus such as viral replication, assembly, regulation of cellular gene expression and in the induction of apoptosis. The present study describes the sequence and mutational analysis of NS2 from Pakistani isolates of 3a genotype (3a GT). A total of 18 amino acid changes were observed out of which fourteen were frequently reported in other genotypes. Our data however revealed 4 rare mutations in NS2. The effects of these mutations were then examined in the secondary and tertiary structures. At secondary structure level, a significant difference in the transmembrane segment 3 helix1 was noticed which is critically involve in viral assembly on the other hand the protease domain was fully conserved. This study revealed that though NS2 is relatively conserved its N terminal transmembranal part exhibit genotype specific variations and need further investigations across all genotypes.**

***Index Terms*—HCV, mutations, structure prediction, secondary structure.**

I. Introduction

Approximately 200 million individuals around the globe are suffering from HCV infection constituting roughly 3.3% of the total world population today [1]. In Pakistan alone, about 10 million people are suffering from this horrendous infection, just about 6% of the total inhabitants of the country [2]. Being a single stranded positive sense RNA virus, HCV contains a single open reading frame encoding approximately 3,000-amino acid polyprotein precursor, which is co- and post translationally cleaved into individual proteins by host and viral proteases. The amino-terminus harbors the structural protein core, envelope proteins E1 and E2 followed by a p7, a small transmembranal protein that forms the hexameric ion channel while the carboxyterminus is occupied by nonstructural (NS) protein, NS2, NS3, NS4A, NS4B, NS5A and NS5B in their respective order [3].

NS2 is 217 amino acids (aa), 23 kDa, nonstructural protein with a hydrophobic amino-terminal subdomain containing up to three putative transmembranal segments (TMS) and a carboxyterminal cytoplasmic domain [4]. The C-terminal domain (residues 94-217) of NS2, together with residues 1-181 of NS3, forms the NS2-3 protease [5]. Cleavage at the NS2/NS3 junction is required for RNA replication of full length HCV replicons [6], in the infectious tissue culture system [7] and in chimpanzees [8]. It is believed that NS2 plays a crucial role in major processes during the propagation of virus such as regulation of cellular gene expression [9], in the induction of apoptosis [10], in the viral assembly process [5] and in viral replication process [9]. The exact role NS2 in virus assembly is not yet completely deciphered and lot of information is still missing. Study of this particular gene of the virus and the role of its protein product needs to be investigated as it might help in the generation of important information for the development of strategies to inhibit viral propagation. The aim of present study was to clone the NS2 gene of hepatitis C virus from Pakistani isolates to identify its mutations and to see their impact on structure and functions of the protein.

Manuscript received December 30, 2012; revised February 28, 2013.

The author are with Atta-ur-Raman School of applied Biosciences; National University of Science and Technology, Islamabad, Pakistan (email: sadiaezza@gmail.com).

II. Material and Method

A. Extraction of Nucleic Acid

HCV patients infected with 3a genotype were included in the study; Patient's blood samples (1300 µl) were taken in EDTA vacutainer tubes. The samples were centrifuged at 12,000 g for 2 min to isolate serum. Viral RNA was extracted from serum by Qiagen RNA extraction kit according to the manufacturer protocols (Qiagen, Germany, Hamburg). RNA was stored at -20 ℃.

B. Specific Primers

For PCR amplification of the HCV NS2 gene, primers were designed by retrieving sequences of the specific viral gene of genotype 3a from NCBI Nucleotide database. The sequence of the sense primer containing restriction enzyme site BamHI was 5' AAAGGATCCGTGGTCGGGTGAAGACAGCG 3' and that of the antisense primer containing restriction enzyme site XhoI was 5' AAGCTCGAGCAACAGACGCCAGC CCATCT 3'.

C. Complementary DNA (cDNA) Synthesis, PCR Amplification and Cloning

Extracted RNA was used as the template for the cDNA synthesis. cDNA synthesis, PCR amplification and cloning of the gene was carried out as described earlier [11].Only the specific primers were changed.

D. Sequencing, Mutational and Structural Analysis

Positive clones were subjected to sequencing by using Beckman coulter CEQ 8000 (USA) as described earlier [11]. Three sequences from each clone were obtained and aligned in multalin (multalin.toulouse.inra.fr/ multalin) to compile a consensus sequences and the consensus sequences were submitted to NCBI (HQ108092-99). Multalin was also used for comparison of different sequences.

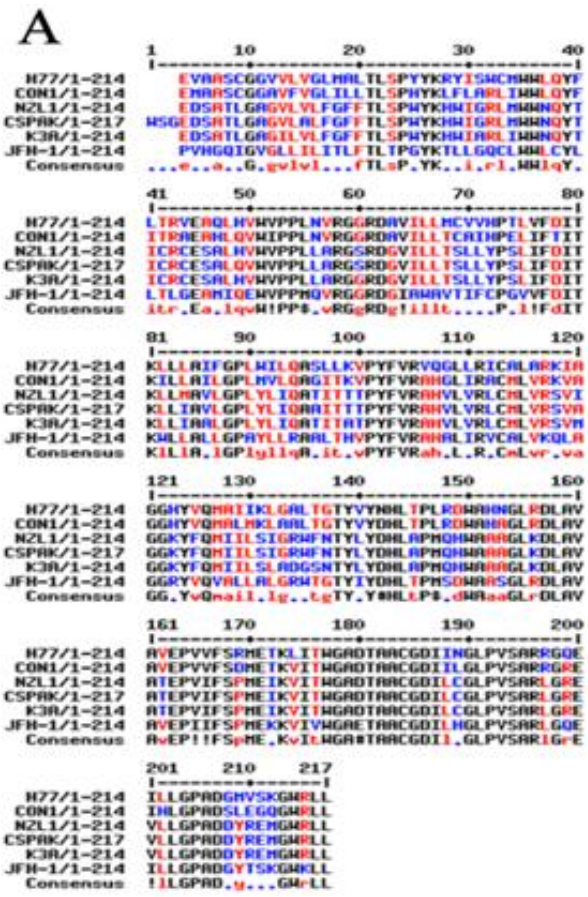
Protein products were predicted and analyzed using PSIPRED (McGuffin *et al.*, 2000) for secondary structure prediction and tertiary structure was predicted and analyzed using web based tool, I-TASSER [12]. Mutations observed were categorized as significant and non significant and their role in the secondary structure was observed in comparison with other reported sequences.

III. Results

Current study deals with impact of certain mutations observed in genotype 3a on the proposed structure of the NS2 protein. Fig. 1 and 2 and Table I shows the observed mutations.

TABLE I: Rare Mutations Observed in HCV NS2 GT3A Isolates from Pakistan

Position	mutation	). of mutations
	W-C	4
99	R-Y	6
132	L-F	1
182	A-T	10



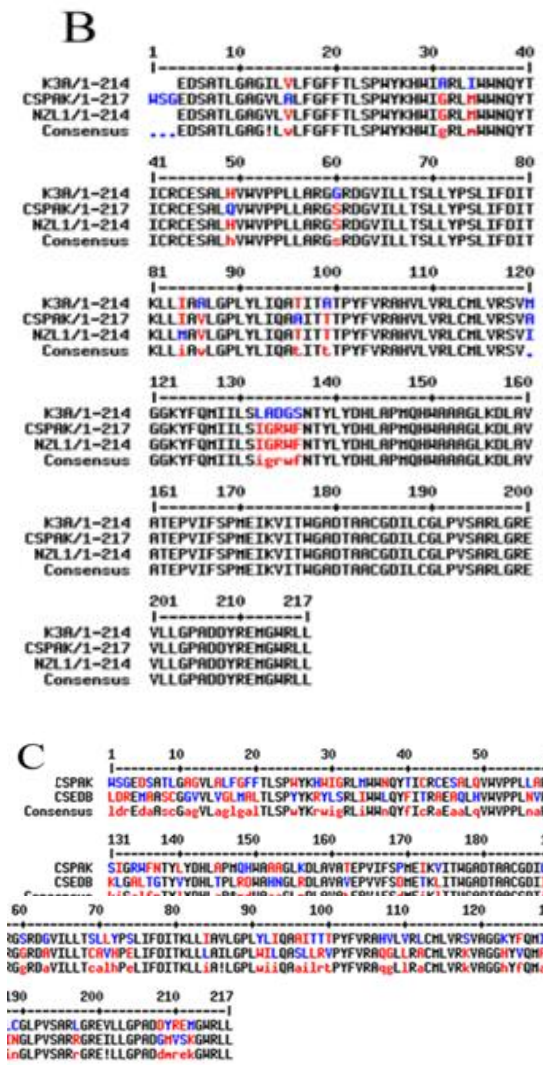


Fig. 1. Multiple sequence alignment of HCV NS2 Pakistani consuses sequence; A showing alignment with reference strains of genotype 1a, 1b, 3a and 2a respectively, B showing alignment with 3a reference strains, C showing alignment with European HCV Database consensus sequence. Black letters refers to high consensus; red low consensus and blue indicate neutral change.

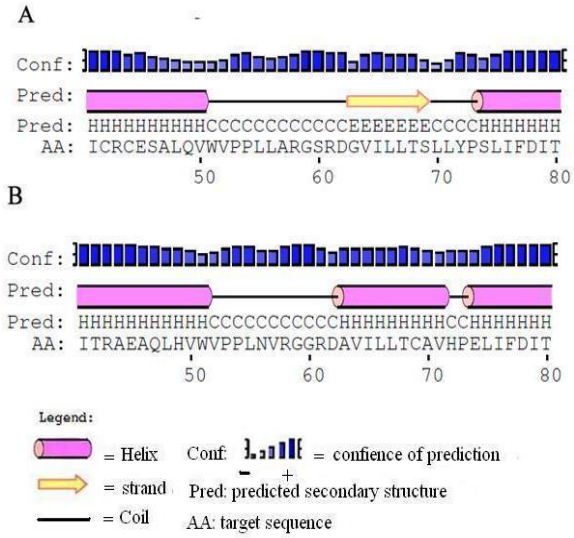


Fig. 2. Secondary structure prediction using PSIRPDB soft ware. Only aa 41-80 are shown. There was no significant difference for the rest of the protein. A represents secondary structure of consuses sequence of Pakistani GT3a isolates (consNS2PK), B shows secondary structure of European HCV data base consensus sequence for NS2 protein (consNS2Eu). At position 61-69 a strand is present in consNS2PK while 63-71 is occupied by a short helix.

IV. Discussion

The exact role of one of the Non-structural protein 2 (NS2) still remains largely unknown once it is dissociated from NS3 protease [13]. Particularly its role in viral assembly is very vague. The present study describes the cloning of HCV NS2 from 3aGT Pakistani isolates, its sequencing and *in silico* analysis. Multiple sequence alignment of the 217 aa protein was performed with consensus sequence of Pakistani isolates (consNS2PK) and consensus sequence of European data base (consNS2EU) (96645 entries till 7-7-2012). In addition aa alignment with various reference strains of genotype 1a, 1b, 2a and 3a was also conducted. The aa alignment results (Figure 1) showed 116 residues with high consensus and 46 residues with neutral mutations when compared with the reference strains such as JFH1 (2a), H77 (1a), Con1 (1b) k3a and NZL1 (3a). When the conNS2PK was considered with reference consNS2EU for the protein, 137 aa were found highly conserved showing 63% high consensus; on the other hand when cons NS2 PK was compared with other 3a isolates 91% residues were found highly conserved and only 18 mutations were seen (Fig. 1). Most of these mutations were neutral mutations with no implications on the structure and properties of the protein or were frequently reported in other genotypes as well. Protein conformations are insensitive to such variations [14] and for that reason these mutations were not investigated further.

Four of these mutations however, were rare mutations (Table I) that were scarcely reported (1-10 times) in other isolates. The significance of these mutations was further analyzed through the prediction of secondary and tertiary structures of the protein. For this purpose secondary structures were predicted by using PSIPRED online soft ware and the data was compared with NMR studies by Jirasko *et al* [4]. A very significant difference was observed in the TMS3 where both the NMR structure and the PSIPRED showed the presence of an α helix at position 63-71 both in JFH1 (2a) and Con1 (1b) isolates and in European consensus sequence, interestingly in 3aPK isolates a slightly smaller 63-69 strand was found at this position. Interestingly the sequence conservation of this particular motif was low among JFH1 and Con1 (37%) than the rest of the domain which showed 58 to 81% conservation and the swap mutant analysis of this particular motif showed that the swapping of these residues results in 1000 fold decrease both in virus production and infectivity [4]. Our data further support the idea that this particular motif is not only genotype specific it even varies in structure in GT3a as compared to GT2a and 1b.This particular motif need father investigations across all genotypes.

Since three of the rarely reported mutations were found in the cystein protease domain of the protein at position R99Y, L132F and A182T; to check whether any structural change is induced into this domain at tertiary level, the structure of consNS2PK3a and that of the consNS2Euro were predicted and compared. The structures were predicted using Protein Databank (PDB) entry 2hd0 as template. PDB 2hd0 is the reported crystalline structure of NS2-3 protease excluding the first 100 amino acids of NS2. The tertiary structures predicted were visualized, compared and analyzed using Swiss PDB viewer. The structure of this protein well tolerated all mutations (data not shown). The observations are quite logical, being a protease domain any significant mutation in this domain will render the virus non infectious and only non disrupting mutations will be tolerated in this domain.

References

[1] J. R. Wands, "Prevention of hepatocellular carcinoma," *N Engl J Med*, vol. 351, no. 15, pp. 1567-70, Oct. 7, 2004.

[2] N. S. Raja and K. A. Janjua, "Epidemiology of hepatitis C virus infection in Pakistan," *J Microbiol Immunol Infect*, vol. 41, no, 1, pp. 4-8, Feb. 2008.

[3] B. D. Lindenbach and C. M. Rice, "Flaviviridae: The Viruses and Their Replication," *Fields Virology*, 5th ed., D. M. Knipe, P. M. Howley, Eds., Philadelphia: Lippincott Williams & Wilkins, PA, USA, vol. 1, 2007, pp. 1101-1152.

[4] V. Jirasko, R. Montserret, J. Y. Lee, J. Gouttenoire, D. Moradpour *et al*., "Structural and Functional Studies of Nonstructural Protein 2 of the Hepatitis C Virus Reveal Its Key Role as Organizer of Virion Assembly," *PLoS Pathog*, vol. 6, no. 12, pp. e1001233, 2010.

[5] M. Pallaoro, A. Lahm, G. Biasiol, M. Brunetti, C. Nardella *et al*., "Characterization of the hepatitis C virus NS2/3 processing reaction by using a purified precursor protein," *J. Virol.* vol. 75, pp. 9939-9946, 2001.

[6] S. Welbourn, R. Green, I. Gamache, S. Dandache, V. Lohmann *et al*., "Hepatitis C virus NS2/3 processing is required for NS3 stability and viral RNA replication," *J. Biol. Chem*., vol. 280, pp. 29604-29611, 2005.

[7] C. T. Jones, C. L. Murray, D. K. Eastman, J. Tassello, and C. M. Rice, "Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus," *J. Virol.*, vol. 81, pp. 8374-8383, 2007.

[8] A. A. Kolykhalov, K. Mihalik, S. M. Feinstone, and C. M. Rice, "Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication *in vivo*," *J. Virol*., vol. 74, pp. 2046-2051, 2000.

[9] F. L. Dumoulin, A. V. D. Bussche, J. Li, L. Khamzina *et al*., "Hepatitis C virus NS2 protein inhibits gene expression from different cellular and viral promoters in hepatic and nonhepatic cell lines," *Virology.*, vol. 305, no. 2, pp. 260-6, Jan. 20, 2003.

[10] M. Yi, Y. Ma, J. Yates, and S. M. Lemon, "Trans-complementation of an NS2 defect in a late step in hepatitis C virus (HCV) particle assembly and maturation," *PLoS Pathog*., vol. 5, pp. e1000403, 2009.

[11] Y. Waheed, S. Tahir, T. Ahmad, and I. Qadri, "Sequence comparison and phylogenetic analysis of core gene of hepatitis C virus from Pakistani population," *African Journal of Biotechnology*, vol. 9, no. 29, pp. 4561-4567, July 19, 2010.

[12] A. Roy, A. Kucukural, and Y. Zhang, "I-TASSER: A unified platform for automated protein structure and function prediction," *Nat Protoc.*, vol. 5, no. 4, pp. 725-38, Apr. 2010.

[13] S. M. Lemon, J. A. McKeating, T. Pietschmann, D. N. Frick *et al*., "Development of novel therapies for hepatitis C," *Antiviral Res.*, vol. 86, no. 1, pp. 79-92, Apr. 2010.

[14] Z. Doszt ányi, A. Fiser, and I. Simon, "Stabilization centers in proteins: identification, characterization and predictions," *J Mol Biol*., vol. 272, no. 4, pp. 597-612, Oct. 3, 1997.

Sadia Anjum was born on 10 June, 1971 in Sialkot, Pakistan. She got her M.Phil in 2003, GCU, in Lahore, Pakistan. She got her PhD 2013 (notification awaited) in National University of Sciences and Technology, Islamabad, Pakistan. Her fields of specialization are Molecular Biology and Virology. She is regular PhD scholar at Atta-ur-Rahman School of Biosciences (ASAB), National University of Sciences and Technology, Islamabad, Pakistan. She is part of the faculty (DEMONSTRATOR) since 2008.