Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 374–377
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Short communication

Expression of alternate reading frame protein (F1) of hepatitis C virus in *Escherichia coli* and detection of antibodies for F1 in Indian patients

Munpally Shesheer Kumar [a,*], Khareedu Venkateswara Rao [a], Chittor Mohammed Habeebullah [b], Vudem Dashavantha Reddy [a]

[a] *Centre for Plant Molecular Biology, Osmania University, Hyderabad 500 007, AP, India*
[b] *Centre for Liver Diseases, Owisi College of Medical Sciences, Kanchan Bagh, Hydearbad, AP, India*

Received 12 July 2007; received in revised form 15 December 2007; accepted 26 December 2007
Available online 5 January 2008

Abstract

Apart from the core (21 kD), a novel hepatitis C virus (HCV) frame shift protein (F1) is synthesized from the initiation codon of the polyprotein sequence followed by ribosomal frame shift into the −2/+1 reading frame. To date, no information is available on F1 protein of Indian isolates, and hence detection of antibodies for F1 protein in Indian patients assumes great relevance. Specific primers have been designed to amplify sequence coding for 120aa of truncated F1 (tF1). The amplified tF1 has been cloned in bacterial expression vector, pET21b for expression in *Escherichia coli*. Partially purified expressed protein has been subjected to western blot analysis using patients' sera. Three HCV positive sera employed in western analysis showed positive signals to tF1, while sera from uninfected individuals failed to give any signals. Further, results of western blots, carried out with patients sera titrated with purified core protein, confirmed the presence of antibodies specific to F1. The positive signal observed for F1 in western analysis with HCV infected sera suggests that F1 protein is synthesized in the natural course of HCV infection in Indian patients as well. Presence of antibodies against F1 protein of subtype 1c has been demonstrated, for the first time, in Indian patients.
© 2008 Elsevier B.V. All rights reserved.

Keywords: HCV; F1; ARFP; Frame shift; Indian isolate; Antibody

1. Introduction

Hepatitis C virus (HCV) is the major causative agent of post-transfusion and parenterally transmitted, non-A, non-B hepatitis throughout the world (Alter and Seeff, 2000). HCV is an enveloped RNA virus that is classified in the family Flaviviridae (Robertson et al., 1998). HCV has high genomic variability and at least six different genotypes and an increasing number of subtypes have been reported (Simmonds, 1999). The genome of HCV comprises a single stranded positive-sense RNA of ~9.6 kb in length and contains a single open reading frame (ORF) that encodes for a non-functional polyprotein of about 3000 aminoacids (Grakoui et al., 1993). This polyprotein is cleaved co- and post-translationally by cellular and viral proteases to yield 10 different functional proteins. Structural proteins are the major components of the mature virion, which are coded by the 5′ quarter of the ORF and arranged as C–E1–E2 and p7, while the non-structural proteins are coded by the 3′ three-quarters of the ORF in the order NS2, NS3, NS4A, NS4B, NS5A and NS5B (Barbara and Contreras, 1991); these proteins are involved in polyprotein processing and replicative functions of the virus (Suzuki et al., 1999; Suzuki et al., 2007). Translation of the HCV polyprotein sequence was reported to be regulated by a cap-independent mechanism that requires most of the 5′-non-coding region and the first nine codons of the polyprotein-coding sequence to serve as the internal ribosomal entry sequence (IRES) (Rijnbrand and Lemon, 2000). Initial expression studies indicated that, besides the core protein (21 kD), another protein (17 kD) also expressed from the same core protein coding sequence both *in vitro* and in mammalian cells and it was thought to be a truncated core protein (Lo et al., 1994, 1995; Basu et al., 2004). Recent observation, in members

* Corresponding author. Tel.: +91 40 27098087; fax: +91 40 27096170.
E-mail address: shesheer@gmail.com (M. Shesheer Kumar).

1567-1348/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.12.008

```
     M   A   R   I   L   N   L   K   E   K   P   N   V   T   P   T   D   A   H   R   T
CAT ATG GCA CGA ATC CTA AAC CTC AAA GAA AAA CCA AAC GTA ACA CCA ACC GAC GCC CAC AGA ACG
 L   S   S   R   V   A   A   R   S   L   A   E   F   A   C   C   R   A   G   V   P   E
TTA AGT TCC CGG GTG GCG GCC AGA TCG TTG GCG GAG TTT GCT TGT TGC CGC GCA GGG GTC CCA GAG
 W   V   C   A   R   R   G   R   L   P   S   G   H   N   L   A   E   G   V   S   L   F
TGG GTG TGC GCG CGA CGA GGA AGA CTT CCG AGC GGT CAC AAC CTC GCG GAA GGC GTC AGC CTA TTC
 P   R   P   A   D   P   R   A   G   P   G   R   S   P   G   T   L   G   P   S   M   A
CCA AGG CCC GCC GAC CCG AGG GCA GGT CCT GGG CGC AGC CCG GGT ACC CTT GGC CCC TCT ATG GCA
 T   R   A   V   G   G   Q   D   G   S   C   P   P   A   A   P   G   L   V   G   A   P
ACG AGG GCT GTG GGT GGG CAG GAT GGC TCT TGT CCC CCC GCG GCT CCC GGC CTA GTC GGG GCC CCT
 L   T   P   G   A   G   H   A   I   W   V   S   L   R   P   H   S   S   T   T   T   T
CTG ACC CCC GGC GCA GGT CAC GCA ATT TGG GTA AGC TTG CGG CCG CAC TCG AGC ACC ACC ACC ACC
 T   T   E   I   R   L   L   T   K   P   E   R   K   L   S   W   L   L   P   P   L   S
ACC ACT GAG ATC CGG CTG CTA ACA AAG CCC GAA AGG AAG CTG AGT TGG CTG CTG CCA CCG CTG AGC
 N   N   *
AAT AAC TAG
```

Fig. 1. Nucleotide sequence of clone pET21b-tF1 with deduced aminoacids. The deduced aminoacid sequence of F1 is represented in red and the sequence in black indicate the aminoacids derived from pET21B plasmid. The sequence in bold represents the sequence of cloning sites.

of family Flaviviridae, revealed that a novel translation mechanism of a ribosomal frame shift exists within the capsid-encoding region, which results in a frame shift protein (Varaklioti et al., 2002; Walewski et al., 2001; Xu et al., 2001). The frame shift protein was named as F1 or ARFP (alternative ribosomal frame shift protein) based on translation initiated at a non-AUG codons in a −2/+1 reading frame relative to the polyprotein of HCV (Baril and Brakier-Gingras, 2005). Similar synthesis of capsid protein via frame shift was also observed in various other viruses such as SARS-CoV (Baranov et al., 2005).

It was first demonstrated that the 17 kD protein was synthesized by ribosomal frame shift and was mostly derived from the coding sequence that overlaps the HCV core protein reading frame (Xu et al., 2001; Choi et al., 2003). The expressed F1 protein was localized in the cytoplasm of HepG2 cells, with a notable perinuclear localization (Roussel et al., 2003), and was found to be associated with the endoplasmic reticulum (Xu et al., 2003). This subcellular localization of HCV F1 protein is similar to that of the HCV core and NS5A proteins, raising the hypothesis that the F1 protein may participate in HCV morphogenesis or replication (Xu et al., 2003). In addition, sera from patients who were positive for HCV genotype 1a or 1b were shown to react differently to synthetic peptides of F1 (Boulant et al., 2003). The present study mainly deals with the cloning and expression of F1 coding sequence of the HCV Indian isolate belonging to genotype 1c. Further, antibodies against F1 protein have been detected in Indian patients. The recombinant plasmid containing HCV core coding sequence, belonging to genotype 1c (GenBank Acc. No. AY051292), was used as template for amplification of truncated F1 (tF1) coding sequence employing F1F 5′ATT CAT ATG GCA CGA ATC CTA AAC C 3′ and F1R 5′ ATT AAG CTT ACC CAA ATT GCG TGA CCT GC 3′ as forward and reverse primers, respectively. The PCR amplification was performed using conditions of 94 °C/30 s, 52 °C/45 s, 72 °C/1 min for 35 cycles and a final extension of 72 °C/5 min. PCR product was gel purified and subjected to restriction digestion using *Nde*I and *Hin*dIII and subsequently cloned at same sites of pET21b. The pET21b-tF1 was subjected to automated DNA sequencing. *E. coli* BL21(DE3) competent cells were transformed with pET21b-tF1 to carryout the expression studies. Western blot analysis was carried out using three positive (anti-HCV and HCV RNA positive) and three negative sera (anti-HCV and HCV RNA negative). Westerns were also performed employing patients' sera titrated with purified core protein. The deduced aminoacid sequences in +1 reading frame of the standard reference set representing various genotypes were utilized in multiple alignment of F1 sequences. The phylogenetic tree was generated based on alignment using clustalW program (http://swift.embl-heildelberg.de).

Analyses of F1 coding −2/+1 reading frame indicated a protein product of 142aa in the Indian isolate AY051292 belonging to genotype 1c. Different genotypes of HCV were reported to code for varied lengths of F1—the genotype 1a encoded 162aa, 1b coded for 144aa and 2a coded for 126aa

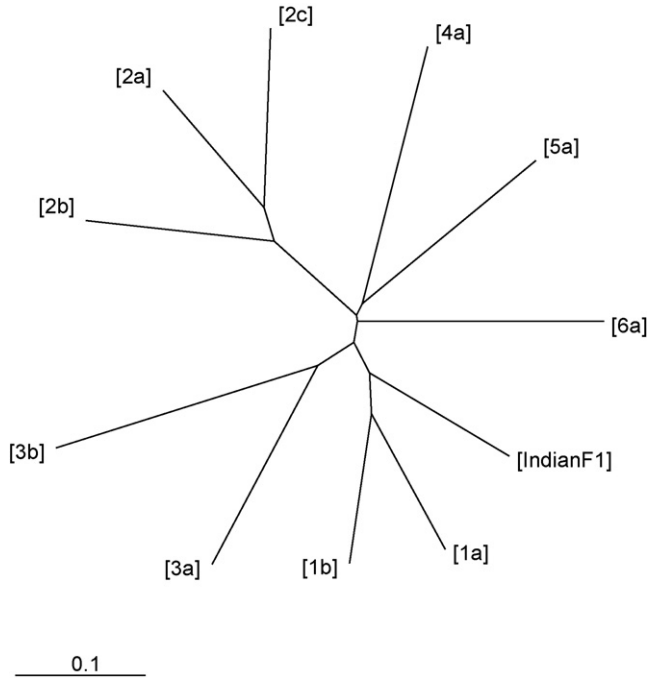

Fig. 2. Phylogenetic analysis based on deduced aminoacid sequences of F1 of various genotypes of HCV. The phylogenetic analysis was done using ClustalW and the tree was constructed using TreeView program. Scale bar shows number of nucleotide substitutions *per* site.

(Kolykhalov et al., 1997; Lohmann et al., 1999; Yanagi et al., 1999; Xu et al., 2001). PCR amplification product of ~370 bp region coding for truncated F1 (tF1) was cloned at *Nde*I and *Hin*dIII sites of pET21b. The clone pET21b-tF1 released a fragment of ~370 bp upon double digestion with *Nde*I and *Hin*dIII.

The clone having the tF1 insert when subjected to sequencing revealed the presence of F1 coding sequence (Fig. 1). The deduced aa of tF1 sequence subjected to BLAST search exhibited domain based identity of 74–77% with POCO45, a core frame shift product of HCV isolate H belonging to type 1a. Sequence alignment of deduced aa of F1 belonging to different genotypes displayed substantial diversity in F1 sequences. Despite these variations, presence of various conserved aa clusters in F1 indicates the conserved nature of its secondary structure among isolates. Phylogenetic analysis of F1 showed close clustering of sequences belonging to various subtypes of specific genotype (Fig. 2) implicating that F1 sequences are genotype specific. Motif search analysis of F1 revealed the presence of Caesin Kinase 2-phosphorylation site; Protein Kinase C-phosphorylation site and LDL class B (LDLRB) receptor binding site. The function of F1 protein in the life cycle of HCV remains unknown (Baril and Brakier-Gingras, 2005). Presence of a binding site for LDLRB indicates the possibility of interaction of F1 with lipids in the natural course of infection. The protein expressed upon IPTG induction yielded ~17 kD band which was absent in un-induced samples (Fig. 3A). The expressed protein tF1 was in the insoluble fraction as inclusion bodies. Inclusion bodies were purified by washing the pellet after lysis using 0.1% Triton-X100 and subsequently the pellet was dissolved in phosphate buffer (pH 8.0) containing 0.5% sodium laural sarcosine (SLS). The partially purified tF1 was employed in western blot analyses. Three HCV positive sera employed in western analysis showed the presence of antibodies to tF1, while sera from uninfected individuals failed to give any signals (Fig. 3B). Similar results were observed with patients' sera titrated with purified core protein (Fig. 3C). Purified core protein was electro-transferred on to nitrocellulose membrane from SDS-PAGE. Several strips of the membrane containing purified core protein were used to titrate out anti-core antibodies in three different positive sera. Finally a western blot without a signal for purified core protein, when titrated sera were used, confirmed the absence of anti-core antibodies. The positive signal observed for F1 in western blot analysis with HCV infected sera and its absence with uninfected sera suggests that F1 protein is plausibly synthesized in the natural course of HCV infection in Indian patients as well.

An overview of the results amply indicates that F1 protein is also synthesized in the natural course of HCV in Indian patients. Phylogenetic analysis of F1 of various HCV isolates revealed that aminoacid sequences of F1 are genotype specific.

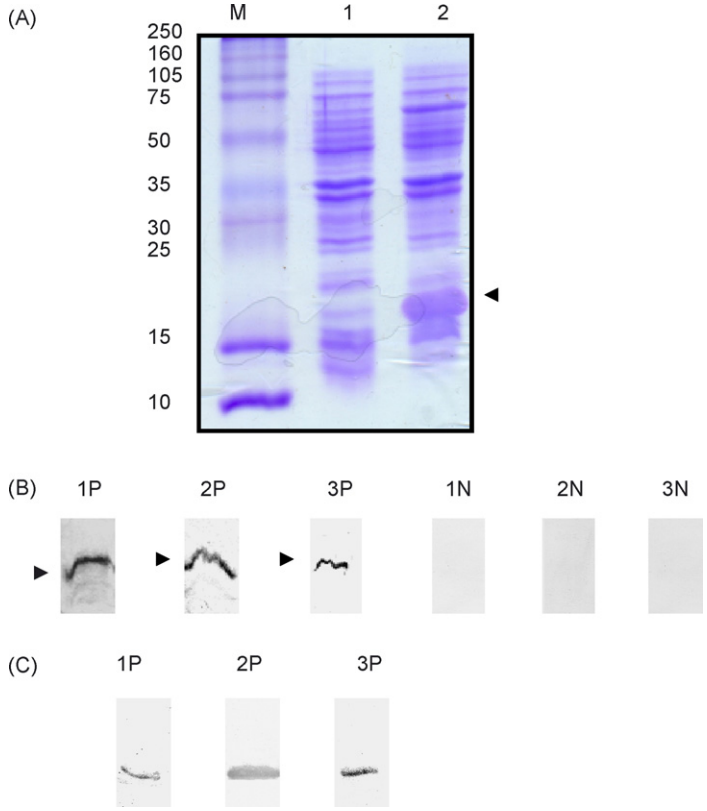

Fig. 3. Expression and western blot analysis of tF1 protein of HCV. (A) Expression of tF1. M, protein marker. Lane 1, pET21b-tF1 before IPTG induction. Lane 2, pET21b-tF1 after IPTG induction showing ~17 kD induced protein. (B) Western blot analysis of tF1. 1P, 2P, 3P: Three different positive sera were used as primary antibody. 1N, 2N, 3N: Three different negative sera were used as negative controls. (C) Western blot analysis of tF1 using patients' sera titrated with purified C120. 1P, 2P, 3P: Three different positive sera used as primary antibody after titration with purified core protein. N represents negative sera and P for positive sera.

Establishment of the presence of antibodies to F1, in this investigation, emphasizes the need for further studies dealing with the role of F1 in HCV pathogenesis.

Acknowledgement

MSK is highly grateful to the CSIR, Govt. of India, New Delhi, for the award of fellowship. Authors extend their thanks to Prof. T. Papi Reddy, former Head, Department of Genetics, O.U, for his helpful suggestions in improving the manuscript.

References

Alter, H.J., Seeff, L.B., 2000. Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. Semin Liver Dis. 20, 17–35.

Baranov, P.V., Henderson, C.M., Anderson, C.B., Gesteland, R.F., Atkins, J.F., Howard, M.T., 2005. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. Virology 332, 498–510.

Barbara, J.A., Contreras, M., 1991. Non-A, non-B hepatitis and the anti-HCV assay. Vox Sang. 60, 1–7.

Baril, M., Brakier-Gingras, L., 2005. Translation of the F protein of hepatitis C virus is initiated at a non-AUG codon in a +1 reading frame relative to the polyprotein. Nucleic Acids Res. 33, 1474–1486.

Basu, A., Steele, R., Ray, R., Ray, R.B., 2004. Functional properties of a 16 kDa protein translated from an alternative open reading frame of the core encoding genomic region of hepatitis C virus. J. Gen. Virol. 85, 2299–2306.

Boulant, S., Becchi, M., Penin, F., Lavergne, J.P., 2003. Unusual multiple recoding events leading to alternative forms of hepatitis C virus core protein from genotype 1b. J Biol Chem. 278, 45785–45792.

Choi, J., Xu, Z., Ou, J.H., 2003. Triple decoding of hepatitis C virus RNA by programmed translational frameshifting. Mol. Cell Biol. 23, 1489–1497.

Grakoui, A., Wychowski, C., Lin, C., Feinstone, S.M., Rice, C.M., 1993. Expression and identification of hepatitis C virus polyprotein cleavage products. J. Virol. 67, 1385–1395.

Kolykhalov, A.A., Agapov, E.V., Blight, K.J., Mihalik, K., Feinstone, S.M., Rice, C.M., 1997. Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. Science 277, 570–574.

Lo, S.Y., Selby, M., Tong, M., Ou, J.H., 1994. Comparative studies of the core gene products of two different hepatitis C virus isolates: two alternative forms determined by a single amino acid substitution. Virology 199, 124–131.

Lo, S.Y., Masiarz, F., Hwang, S.B., Lai, M.M., Ou, J.H., 1995. Differential subcellular localization of hepatitis C virus core gene products. Virology 213, 455–461.

Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L., Bartenschlager, R., 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 285, 110–113.

Rijnbrand, R.C., Lemon, S.M., 2000. Internal ribosome entry site-mediated translation in hepatitis C virus replication. Curr. Top. Microbiol. Immunol. 242, 85–116.

Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., 1998. Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. Arch. Virol. 143, 2493–2503.

Roussel, J., Pillez, A., Montpellier, C., Duverlie, G., Cahour, A., Dubuisson, J., Wychowski, C., 2003. Characterization of the expression of the hepatitis C virus F protein. J. Gen. Virol. 84, 1751–1759.

Simmonds, P., 1999. Viral heterogeneity of hepatitis C virus. J. Hepatol. 31, 54–60.

Suzuki, R., Suzuki, T., Ishii, K., Matsuura, Y., Miyamura, T., 1999. Processing and functions of hepatitis C virus proteins. Intervirology 42, 145–152.

Suzuki, T., Aizaki, H., Murakami, K., Shoji, I., Wakita, T., 2007. Molecular biology of hepatitis C virus. J. Gastroenterol. 42, 411–423.

Varaklioti, A., Vassilaki, N., Georgopoulou, U., Mavromara, P., 2002. Alternate translation occurs within the core coding region of the hepatitis C viral genome. J. Biol. Chem. 277, 17713–17721.

Walewski, J.L., Keller, T.R., Stump, D.D., Branch, A.D., 2001. Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. RNA 7, 710–721.

Xu, Z., Choi, J., Yen, T.S., Lu, W., Strohecker, A., Govindarajan, S., Chien, D., Selby, M.J., Ou, J., 2001. Synthesis of a novel hepatitis C virus protein by ribosomal frame shift. EMBO J. 20, 3840–3848.

Xu, Z., Choi, J., Lu, W., Ou, J.H., 2003. Hepatitis C virus f protein is a short lived protein associated with the endoplasmic reticulum. J. Virol. 77, 1578–1583.

Yanagi, M., Purcell, R.H., Emerson, S.U., Bukh, J., 1999. Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic 1a and 2a chimeras. Virology 262, 250–263.