Molecular biology and clinical implication of hepatitis C virus

A.C. Lyra[1], X. Fan[2] and A.M. Di Bisceglie[2]

[1]Serviço de Gastro-Hepatologia e Laboratório de Biologia Molecular, Hospital São Rafael e Universidade Federal da Bahia, Salvador, BA, Brasil
[2]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Saint Louis University School of Medicine, St. Louis, MO, USA

Abstract

Correspondence

A.C. Lyra

Rua Socrates Guanaes Gomes, 84/401

40283-320 Salvador, BA

Brasil

Fax: +55-71-452-9589/350-4664

E-mail: aclyra@atarde.com.br

. . . .

Hepatitis C virus (HCV) was first described in 1989 as the putative viral agent of non-A non-B hepatitis. It is a member of the Flaviviridae family and has been recognized as the major causative agent of chronic liver disease, including chronic active hepatitis, cirrhosis and hepatocellular carcinoma. HCV is a positive RNA virus with a genome containing approximately 9500 nucleotides. It has an open reading frame that encodes a large polyprotein of about 3000 amino acids and is characterized by extensive genetic diversity. HCV has been classified into at least 6 major genotypes with many subtypes and circulates within an infected individual as a number of closely related but distinct variants known as quasispecies. This article reviews aspects of the molecular biology of HCV and their clinical implication.

Introduction

Hepatitis C virus (HCV) was first described in 1989 as the putative viral agent of non-A non-B hepatitis (1). It is a member of the Flaviviridae family and has been recognized as the major causative agent of chronic liver disease, including chronic active hepatitis, cirrhosis and hepatocellular carcinoma (2). The overall prevalence of anti-HCV antibodies in the United States is 1.4% (3.9 million people). The estimated number of actively infected individuals is 2.7 million. In Brazil, the overall prevalence of anti-HCV antibodies ranges from 1.5 to 1.7% among the general population and blood donors.

Genomic organization of hepatitis C virus

HCV is a positive RNA virus with a genome containing approximately 9500 nucleotides. It has an open reading frame that encodes a large polyprotein of about 3000 amino acids. When cleaved by viral and host enzymes this polyprotein produces 10 polypeptides. The genomic organization of HCV consists of a 5' untranslated region (5' UTR) of about 340 base pairs and downstream to the 5' UTR there are regions that encode the following proteins: "core", envelope 1, envelope 2, p7, non-structural region 2, non-structural region 3 (NS3), non-structural regions 4A and 4B, and non-structural regions 5A and 5B (NS5A and NS5B) (Figure 1). Downstream to this coding region there is another untranslated region of approximately 200 nucleotides (3' UTR).

The 5' UTR is highly conserved and therefore is the target for diagnostic assays. It has a complex secondary structure and is the putative internal ribosome entry site, with an

RNA sequence that directs the ribosomes to initiate translation (3).

The HCV core has RNA binding capacity and appears to be associated with the HCV genome. It has been suggested that this protein induces hepatic steatosis in transgenic mice and is an important factor in the development of hepatocellular carcinoma in patients with chronic HCV infection (4). Serum quantification of the core protein by ELISA correlates with serum levels of HCV RNA.

HCV envelope 1 and envelope 2 proteins have molecular masses of approximately 31 and 70 kDa, respectively. They interact with each other and may be a heterodimer. The envelope 1 region is used for genotyping for clinical purposes. The first 81 nucleotides of the envelope 2 region are thought to be the hypervariable region 1 (HVR1) of HCV (5). HVR1 appears to induce the production of neutralizing antibodies and could function as a decoy to help the virus escape the immune system. HCV envelope 2 is thought to have binding sites for CD81, a protein expressed in the membranes of mammalian cells that might be implicated in the entry of HCV into the hepatocyte (6-8). NS3 protein has protease and helicase activities (9). Its helicase activity participates in the process of RNA unwinding during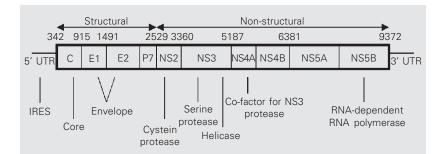 replication. The NS5A region has a particular nucleotide sequence that may determine the sensitivity of the virus to interferon therapy (interferon sensitivity determination region) (10,11). Nevertheless, other investigators have reported discordant results and this issue is controversial (12,13). An *in vitro* study did not find a relationship between interferon sensitivity determination region and sensitivity to interferon (14).

The NS5B region has motifs characteristic of an RNA-dependent RNA polymerase, an enzyme responsible for viral replication (15).

Genetic heterogeneity of hepatitis C virus

As is the case for other RNA viruses, HCV has a high genetic heterogeneity. The estimated mutation rate in the human organism is 1.92×10^{-3} nucleotide site^{-1} year^{-1} (16). This genetic diversity is not evenly distributed in the viral genome. The non-coding regions are relatively conserved, while the envelope regions, especially HVR1, have the highest mutation rate. This genetic diversity resulted in the classification of HCV into at least 6 major genotypes (1 to 6) and many subtypes, based on the analysis of the NS5 region (17-23). Different genotypes may differ by 30 to 35% of their viral genomes, while subtypes may differ by 15-20% of their nucleotide sequences.

Several studies have suggested that HCV genotype 1b might be associated with greater severity of liver disease than other HCV genotypes, including the development of hepatocellular carcinoma (24,25). However, other investigators have not confirmed these data, and this issue remains unresolved (26,27). On the other hand, it has been well established that patients chronically infected with HCV genotype 1 isolates respond less favorably to therapy with interferon or pegylated interferon alone or in combination with ribavirin compared to patients infected with HCV genotype 2 or 3 (28,29).

It is interesting to note that HCV circulates

Figure 1. The genomic organization of hepatitis C virus. The two ends contain untranslated regions (UTR), which are shown as lines (5′ UTR and 3′ UTR). Downstream from the 5′ UTR there is an open reading frame encoding a single large polyprotein that is processed into structural and non-structural (NS) proteins by viral and host enzymes. Nucleotide positions are reported according to the HCV J4 strain (GenBank accession number D10749). IRES = interferon sensitivity determination region.

as a population of different but closely related genomes, known as quasispecies, whose sequences differ only by a few nucleotides (30). These variants may be generated by nucleotide substitutions occurring due to lack of fidelity of the HCV RNA-dependent RNA polymerase, as observed with other RNA viruses. The variants present at any time may be determined by a combination of viral fitness and response to immune pressure.

The existence of a quasispecies population appears to hamper the development of vaccines for HCV and to favor the perpetuation of the virus in the human organism. In a study of 12 patients with acute hepatitis C, Farci et al. (31) observed that individuals who developed chronic infection had a significantly greater HVR1 genetic diversity after 8 to 11 weeks compared to patients who eliminated the virus during the acute phase.

The influence of HCV quasispecies on the response to therapy with alpha interferon has been analyzed by many investigators (32-36). Although the results were not uniform, a higher HVR1 complexity and diversity before treatment appeared to be more frequently observed among non-responders.

The role of HCV quasispecies in the severity of recurrent hepatitis C after liver transplantation has been evaluated in several studies.

Gretch et al. (37) studied 3 patients with severe recurrent hepatitis C and 2 with mild recurrence of the disease. HVR1 genetic diversity was higher in the mild group compared to the severe group. In another study with a larger number of patients, the same group of investigators obtained similar results (38). On the other hand, Pessoa et al. (39) obtained contradictory findings. In a study of 6 patients with fibrosing cholestatic HCV after transplant and 6 patients with mild recurrence of the disease, these investigators found a greater HVR1 diversity among patients with severe recurrent hepatitis C. We studied 11 patients who underwent transplantation for HCV-related cirrhosis and developed recurrent hepatitis C (40). Patients were divided into two groups according to the stage of hepatic fibrosis detected in follow-up biopsies after transplantation. The mild group had stage 1 or 2 fibrosis while the severe group had stage 3 or 4. After liver transplantation, a more complex HCV HVR1 quasispecies population was associated with mild recurrence of the disease. Among those patients in whom recurrent hepatitis C led to severe disease, the number of HVR1 non-synonymous substitutions was greater than the number of synonymous substitutions. Interestingly, the regions flanking HVR1 had a similar pattern of genetic evolution in patients from the mild and severe groups. This suggests that the difference in the alterations observed in HVR1 did not occur sporadically. The highest degree of change in viral amino acid sequences occurred within the first 36 months after transplantation.

It has been suggested that partially mismatched primers may preferentially amplify different HVR1 sequences in a heterogeneous virus population. This suggests a possible mechanism of bias during the amplification of HVR1 sequences and may be responsible for some conflicting data regarding the evolutionary or clinical implications of HCV quasispecies (41).

After liver transplantation, we have demonstrated that the quasispecies nature of two putative regions of CD81-binding sites located in the E2 region appears to remain conserved (Lyra AC, Fan X and Di Bisceglie AM, unpublished data).

When liver transplantation is carried out in a situation in which both donor and recipient are infected with hepatitis C, only one viral strain, either the strain from the donor or the strain from the recipient, predominates and persists up to 4 years after surgery. The two strains are not detected at the same time. This suggests that hepatitis C co-infection after liver transplantation is not a frequent finding (42).

Conclusions

The application of sophisticated techniques of molecular biology over the last 10 years has led to remarkable advances in the knowledge of the molecular structure of HCV. Studies on HCV replication and pathogenesis have been hampered by the lack of reliable and efficient cell culture systems. Recently, efficient initiation of HCV RNA replication in cell-based systems was reported for full-length RNAs derived from genotype 1b isolates, and later for full-length RNAs derived from genotype 1a strain H77 (14,43,44). These cell-based systems should be useful for basic replication studies and for the development of antiviral agents. New therapies for treating hepatitis C are in various stages of preclinical and clinical development. These include nucleic acid-based approaches (antisense and ribozymes), small molecule inhibitors of essential HCV-encoded enzymes (protease, helicase, and polymerase), immune modulation, and immunotherapy.

References

1. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW & Houghton M (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, 244: 359-362.

2. Di Bisceglie AM (2000). Natural history of hepatitis C: its impact on clinical management. *Hepatology*, 31: 1014-1018.

3. Tsukiyama-Kohara K, Iizuka N, Kohara M & Nomoto A (1992). Internal ribosome entry site within hepatitis C virus RNA. *Journal of Virology*, 66: 1476-1483.

4. Moriya K, Fujie H, Shintani Y, Yotsuyanagi H, Tsutsumi T, Ishibashi K, Matsuura Y, Kimura S, Miyamura T & Koike K (1998). The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nature Medicine*, 4: 1065-1067.

5. Weiner AJ, Brauer MJ, Rosenblatt J, Richman KH, Tung J, Crawford K, Bonino F, Saracco G, Choo QL & Houghton M (1991). Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology*, 180: 842-848.

6. Pileri P, Uematsu Y, Campagnoli S et al. (1998). Binding of hepatitis C virus to CD81. *Science*, 282: 938-941.

7. Flint M, Maidens C, Loomis-Price LD, Shotton C, Dubuisson J, Monk P, Higginbottom A, Levy S & McKeating JA (1999). Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81. *Journal of Virology*, 73: 6235-6244.

8. Levy S, Todd SC & Maecker HT (1998). CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. *Annual Review of Immunology*, 16: 89-109.

9. Kim DW, Gwack Y, Han JH & Choe J (1995). C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochemical and Biophysical Research Communications*, 215: 160-166.

10. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Izumi N, Marumo F & Sato C (1995). Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *Journal of Clinical Investigation*, 96: 224-230.

11. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Ogura Y, Izumi N, Marumo F & Sato C (1996). Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *New England Journal of Medicine*, 334: 77-81.

12. Pawlotsky JM, Germanidis G, Neumann AU, Pellerin M, Frainais PO & Dhumeaux D (1998). Interferon resistance of hepatitis C virus genotype 1b: relationship to nonstructural 5A gene quasispecies mutations. *Journal of Virology*, 72: 2795-2805.

13. Ibarrola N, Moreno-Monteagudo JA, Saiz M, Garcia-Monzon C, Sobrino F, Garcia-Buey L, Lo Iacono O, Moreno-Otero R & Martinez-Salas E (1999). Response to retreatment with interferon-alpha plus ribavirin in chronic hepatitis C patients is independent of the NS5A gene nucleotide sequence. *American Journal of Gastroenterology*, 94: 2487-2495.

14. Blight KJ, Kolykhalov AA & Rice CM (2000). Efficient initiation of HCV RNA replication in cell culture. *Science*, 290: 1972-1974.

15. Behrens SE, Tomei L & De Francesco R (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO Journal*, 15: 12-22.

16. Ogata N, Alter HJ, Miller RH & Purcell RH (1991). Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proceedings of the National Academy of Sciences, USA*, 88: 3392-3396.

17. Simmonds P, Holmes EC, Cha TA, Chan SW, McOmish F, Irvine B, Beall E, Yap PL, Kolberg J & Urdea MS (1993). Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *Journal of General Virology*, 74: 2391-2399.

18. Okamoto H, Kurai K, Okada S, Yamamoto K, Lizuka H, Tanaka T, Fukuda S, Tsuda F & Mishiro S (1992). Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology*, 188: 331-341.

19. Houghton M, Weiner A, Han J, Kuo G & Choo QL (1991). Molecular biology of the hepatitis C viruses: implications for diagnosis, development and control of viral disease. *Hepatology*, 14: 381-388.

20. Simmonds P, Alberti A, Alter HJ et al. (1994). A proposed system for the nomenclature of hepatitis C viral genotypes. *Hepatology*, 19: 1321-1324 (Letter).

21. Ohno O, Mizokami M, Wu RR, Saleh MG, Ohba K, Orito E, Mukaide M, Williams R & Lau JY (1997). New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. *Journal of Clinical Microbiology*, 35: 201-207.

22. Davidson F, Simmonds P, Ferguson JC, Jarvis LM, Dow BC, Follett EA, Seed CR, Krusius T, Lin C & Medgyesi GA (1995). Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *Journal of General Virology*, 76: 1197-1204.

23. Simmonds P, Smith DB, McOmish F, Yap PL, Kolberg J, Urdea MS & Holmes EC (1994). Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. *Journal of General Virology*, 75: 1053-1061.

24. Bruno S, Silini E, Crosignani A et al. (1997). Hepatitis C virus genotypes and risk of hepatocellular carcinoma in cirrhosis: a prospective study. *Hepatology*, 25: 754-758.

25. Tanaka H, Tsukuma H, Yamano H, Okubo Y, Inoue A, Kasahara A & Hayashi N (1998). Hepatitis C virus 1b(II) infection and development of chronic hepatitis, liver cirrhosis and hepatocellular carcinoma: a case-control study in Japan. *Journal of Epidemiology*, 8: 244-249.

26. Benvegnu L, Pontisso P, Cavalletto D, Noventa F, Chemello L & Alberti A (1997). Lack of correlation between hepatitis C virus genotypes and clinical course of hepatitis C virus-related cirrhosis. *Hepatology*, 25: 211-215.

27. Reid AE, Koziel MJ, Aiza I, Jeffers L, Reddy R, Schiff E, Lau JY, Dienstag JL & Liang TJ (1999). Hepatitis C virus genotypes and viremia and hepatocellular carcinoma in the United States. *American Journal of Gastroenterology*, 94: 1619-1626.

28. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, Goodman ZD, Ling MH, Cort S & Albrecht JK (1998). Interferon alpha-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *New England Journal of Medicine*, 339: 1485-1492.

29. Poynard T, Marcellin P, Lee SS et al. (1998). Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. International Hepatitis Interventional Therapy Group (IHIT). *Lancet*, 352: 1426-1432.

30. Martell M, Esteban JI, Quer J, Genesca J, Weiner A, Esteban R, Guardia J & Gomez J (1992). Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *Journal of Virology*, 66: 3225-3229.

31. Farci P, Shimoda A, Coiana A et al. (2000). The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science*, 288: 339-344.

32. Wilson JJ, Polyak SJ, Day TD & Gretch DR (1995). Characterization of simple and complex hepatitis C virus quasispecies by heteroduplex gel shift analysis: correlation with nucleotide sequencing. *Journal of General Virology*, 76: 1763-1771.

33. Kanazawa Y, Hayashi N, Mita E, Li T, Hagiwara H, Kasahara A, Fusamoto H & Kamada T (1994). Influence of viral quasispecies on effectiveness of interferon therapy in chronic hepatitis C patients. *Hepatology*, 20: 1121-1130.

34. Okada S, Akahane Y, Suzuki H, Okamoto H & Mishiro S (1992). The degree of variability in the amino terminal region of the E2/NS1 protein of hepatitis C virus correlates with responsiveness to interferon therapy in viremic patients. *Hepatology*, 16: 619-624.

35. Polyak SJ, Faulkner G, Carithers Jr RL, Corey L & Gretch DR (1997). Assessment of hepatitis C virus quasispecies heterogeneity by gel shift analysis: correlation with response to interferon therapy. *Journal of Infectious Diseases*, 175: 1101-1107.

36. Hassoba HM, Bzowej N, Berenguer M, Kim M, Zhou S, Phung Y, Grant R, Pessoa MG & Wright TL (1999). Evolution of viral quasispecies in interferon-treated patients with chronic hepatitis C virus infection. *Journal of Hepatology*, 31: 618-627.

37. Gretch DR, Polyak SJ, Wilson JJ, Carithers Jr RL, Perkins JD & Corey L (1996). Tracking hepatitis C virus quasispecies major and minor variants in symptomatic and asymptomatic liver transplant recipients. *Journal of Virology*, 70: 7622-7631.

38. Sullivan DG, Wilson JJ, Carithers Jr RL, Perkins JD & Gretch DR (1998). Multigene tracking of hepatitis C virus quasispecies after liver transplantation: correlation of genetic diversification in the envelope region with asymptomatic or mild disease patterns. *Journal of Virology*, 72: 10036-10043.

39. Pessoa MG, Bzowej N, Berenguer M, Phung Y, Kim M, Ferrell L, Hassoba H & Wright TL (1999). Evolution of hepatitis C virus quasispecies in patients with severe cholestatic hepatitis after liver transplantation. *Hepatology*, 30: 1513-1520.

40. Lyra AC, Fan X, Lang DM, Yusim K, Ramrakhiani S, Brunt EM, Korber B, Perelson AS & Di Bisceglie AM (2002). Evolution of hepatitis C viral quasispecies after liver transplantation. *Gastroenterology*, 123: 1485-1493.

41. Fan X, Lyra AC, Tan D, Xu Y & Di Bisceglie AM (2001). Differential amplification of hypervariable region 1 of hepatitis C virus by partially mismatched primers. *Biochemical and Biophysical Research Communications*, 284: 694-697.

42. Fan X, Lang DM, Xu Y, Lyra AC, Yusim K, Everhart JE, Korber BT, Perelson AS & Di Bisceglie AM (2003). Liver transplantation with hepatitis C virus-infected graft: interaction between donor and recipient viral strains. *Hepatology*, 38: 25-33.

43. 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.

44. Blight KJ, McKeating JA, Marcotrigiano J & Rice CM (2003). Efficient replication of hepatitis C virus genotype 1a RNAs in cell culture. *Journal of Virology*, 77: 3181-3190.