ISSN 1007-9327 (print)
ISSN 2219-2840 (online)

World Journal of *Gastroenterology*

World J Gastroenterol 2018 August 14; 24(30): 3313-3468

Published by Baishideng Publishing Group Inc

World Journal of Gastroenterology

Contents

Weekly Volume 24 Number 30 August 14, 2018

REVIEW

MINIREVIEWS

ORIGINAL ARTICLE

Basic Study

Retrospective Study

World Journal of Gastroenterology

Contents

Weekly Volume 24 Number 30 August 14, 2018

ABOUT COVER

Editorial board member of *World Journal of Gastroenterology*, Herwig Cerwenka, MD, Professor, Department of Surgery, Medical University of Graz, Graz A-8036, Austria

AIMS AND SCOPE

World Journal of Gastroenterology (*World J Gastroenterol*, *WJG*, print ISSN 1007-9327, online ISSN 2219-2840, DOI: 10.3748) is a peer-reviewed open access journal. *WJG* was established on October 1, 1995. It is published weekly on the 7th, 14th, 21st, and 28th each month. The *WJG* Editorial Board consists of 642 experts in gastroenterology and hepatology from 59 countries.

The primary task of *WJG* is to rapidly publish high-quality original articles, reviews, and commentaries in the fields of gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, hepatobiliary surgery, gastrointestinal oncology, gastrointestinal radiation oncology, gastrointestinal imaging, gastrointestinal interventional therapy, gastrointestinal infectious diseases, gastrointestinal pharmacology, gastrointestinal pathophysiology, gastrointestinal pathology, evidence-based medicine in gastroenterology, pancreatology, gastrointestinal laboratory medicine, gastrointestinal molecular biology, gastrointestinal immunology, gastrointestinal microbiology, gastrointestinal genetics, gastrointestinal translational medicine, gastrointestinal diagnostics, and gastrointestinal therapeutics. *WJG* is dedicated to become an influential and prestigious journal in gastroenterology and hepatology, to promote the development of above disciplines, and to improve the diagnostic and therapeutic skill and expertise of clinicians.

INDEXING/ABSTRACTING

World Journal of Gastroenterology (*WJG*) is now indexed in Current Contents®/Clinical Medicine, Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports®, Index Medicus, MEDLINE, PubMed, PubMed Central and Directory of Open Access Journals. The 2018 edition of Journal Citation Reports® cites the 2017 impact factor for *WJG* as 3.300 (5-year impact factor: 3.387), ranking *WJG* as 35th among 80 journals in gastroenterology and hepatology (quartile in category Q2).

NAME OF JOURNAL
World Journal of Gastroenterology

ISSN
ISSN 1007-9327 (print)
ISSN 2219-2840 (online)

LAUNCH DATE
October 1, 1995

FREQUENCY
Weekly

EDITORIAL OFFICE
Ze-Mao Gong, Director
World Journal of Gastroenterology
Baishideng Publishing Group Inc
7901 Stoneridge Drive, Suite 501,
Pleasanton, CA 94588, USA
Telephone: +1-925-2238242
Fax: +1-925-2238243
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.f6publishing.com/helpdesk
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
7901 Stoneridge Drive, Suite 501,
Pleasanton, CA 94588, USA
Telephone: +1-925-2238242
Fax: +1-925-2238243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.f6publishing.com/helpdesk
http://www.wjgnet.com

PUBLICATION DATE
August 14, 2018

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/bpg/gerinfo/204

ONLINE SUBMISSION
http://www.f6publishing.com

World Journal of Gastroenterology

Submit a Manuscript: http://www.f6publishing.com

World J Gastroenterol 2018 August 14; 24(30): 3374-3383

DOI: 10.3748/wjg.v24.i30.3374

ISSN 1007-9327 (print) ISSN 2219-2840 (online)

MINIREVIEWS

Form confers function: Case of the 3'X region of the hepatitis C virus genome

Mariola Dutkiewicz, Jerzy Ciesiołka

Mariola Dutkiewicz, Jerzy Ciesiołka, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan 61-704, Poland

ORCID number: Mariola Dutkiewicz (0000-0003-1098-0308); Jerzy Ciesiołka (0000-0001-7274-0751).

Author contributions: All authors equally contributed to this paper with conception and design of the study, literature review and analysis, drafting and critical revision and editing, and final approval of the final version.

Supported by the Polish Ministry of Science and Higher Education under the KNOW program and Foundation for Polish Science, No. POMOST/2013-8/5.

Conflict-of-interest statement: No potential conflicts of interest. No financial support.

Manuscript source: Invited manuscript

Correspondence to: Mariola Dutkiewicz, PhD, Associate Professor, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, Poznan 61-704, Poland. mariolad@ibch.poznan.pl
Telephone: +48-61-8528503
Fax: +48-61-8528532

Received: April 28, 2018
Peer-review started: April 28, 2018
First decision: June 11, 2018
Revised: June 25, 2018
Accepted: June 30, 2018
Article in press: June 30, 2018
Published online: August 14, 2018

Abstract

At the 3' end of genomic hepatitis C virus (HCV) RNA there is a highly conserved untranslated region, the 3' X-tail, which forms part of the 3'UTR. This region plays key functions in regulation of critical processes of the viral life cycle. The 3'X region is essential for viral replication and infectivity. It is also responsible for regulation of switching between translation and transcription of the viral RNA. There is some evidence indicating the contribution of the 3'X region to the translation efficiency of the viral polyprotein and to the encapsidation process. Several different secondary structure models of the 3'X region, based on computer predictions and experimental structure probing, have been proposed. It is likely that the 3'X region adopts more than one structural form in infected cells and that a specific equilibrium between the various forms regulates several aspects of the viral life cycle. The most intriguing explanations of the structural heterogeneity problem of the 3'X region came with the discovery of its involvement in long-range RNA-RNA interactions and the potential for homodimer formation. This article summarizes current knowledge on the structure and function of the 3'X region of hepatitis C genomic RNA, reviews previous opinions, presents new hypotheses and summarizes the questions that still remain unanswered.

Key words: Hepatitis C virus; 3'UTR; 3'X-tail; 3'X region; 3'X RNA; RNA structure

Core tip: Several different secondary structure models of the 3'X region have been proposed. It is likely that the 3' X region adopts more than one structural form in infected cells and that a specific equilibrium between the various forms regulates several aspects of the viral life cycle. This article summarizes current knowledge of the structure and function of the 3'X region of hepatitis C genomic RNA,

reviews previous opinions, presents new hypotheses and summarizes the questions that still remain unanswered.

Dutkiewicz M, Ciesiołka J. Form confers function: Case of the 3' X region of the hepatitis C virus genome. *World J Gastroenterol* 2018; 24(30): 3374-3383 Available from: URL: http://www.wjgnet.com/1007-9327/full/v24/i30/3374.htm DOI: http://dx.doi.org/10.3748/wjg.v24.i30.3374

INTRODUCTION

Hepatitis C virus (HCV) belongs to the family *Flaviviridae*, a member of the genus *hepaciviruses*. HCV was identified in 1989 as one of the viruses capable of causing viral hepatitis, in addition to the previously known hepatitis A and B viruses (HAV, HBV)[1]. There are seven major HCV genotypes that differ in virulence and their geographic distribution[2]. Currently, it is estimated that worldwide more than 185 million people are infected with HCV, which represents 2.8% of the world's population[3]. Infection usually takes place in the absence of obvious clinical symptoms, and the resulting inflammation of the liver often progresses to become chronic, usually lasting for years. In time, the chronic inflammation can lead to cirrhosis of the liver and ultimately liver failure or to the development of primary liver cancer[4]. Although the virus replicates mainly in hepatocytes, it also occurs in peripheral blood mononuclear and central nervous system microglia[5,6]. In people with simultaneous infection with HIV (human immunodeficiency virus), HCV replication was also observed in other tissues[7]. Sometimes HCV enters cells of the immune system, leading to very long-lasting effects which are extremely difficult to treat effectively[8].

While progress has been made in the treatment of some other common viral infections, HCV infection remains an important health problem in the world. Current standard treatment methods have the desired therapeutic effect on 40%-60% of patients[4]. The new highly effective drug, grazoprevir, is able to cure patients with 93% effectivity[9], unfortunately, the treatment is still very expensive and out of reach for the majority of the infected individuals, many of whom live in Third World countries. So far, no HCV vaccine has been developed, which is due to the high genetic variability of the virus, comparable to the genetic variability of HIV. In the absence of widely accessible conventional drugs and vaccines, numerous attempts have been made to design inhibitors of viral proteins, inhibitory oligomers of the antisense and ribozyme type, and more recently also of RNA interference tools directed against viral RNA[4,10,11].

HCV is a small, enveloped virus with a diameter of 40-60 nm, whose genome is a single-stranded, positive-sense RNA of about 9.6 kb in length[6,10,12]. In the viral replication process, the positive-sense RNA strand is transcribed into negative-sense counterpart, the replication intermediate, which serves as the template for the RNA synthesis of progeny genomes. The HCV genome has one very long open reading frame (ORF). It encodes a precursor polyprotein, which is digested in a series of cleavage processes to finally produce proteins: C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B[13]. The coding sequence of the HCV genome is flanked by two untranslated regions: the 5'UTR and the 3'UTR (5' and 3' untranslated regions). Both these regions play key functions in regulating HCV life cycle and determine its expression level. The 5'UTR contains a highly structured regulatory element, an IRES (internal ribosome entry site), that enables cap-independent translation. The 3' UTR is engaged in the replication process and in the regulation of translation. In addition, control of other processes as, for example, assembly of virions and switching between different developmental phases takes place with the participation of the structural RNA elements present at the very 3' end of the HCV genome.

This article presents the current state of knowledge about the structure and functions of the most terminal section of the 3'UTR of hepatitis C virus, the 3'X-tail.

STRUCTURE OF THE 3'UTR REGION

HCV 3'UTR has a variable length of 170 to 250 nucleotides (Figure 1). Three characteristic sections have been recognized in this region: immediately after the stop codon there is a variable region about 25-130 nucleotides in length, characterized by high sequence heterogeneity between various genotypes, but conserved within the same genotype of the virus. Next there is a poly-pyrimidine segment of varying length, independent of the type, or even a subtype of virus, ranging from about 30 to 130 nucleotides. At the very end of the genome, there is an 3'X region, discovered 6 years after cloning of the virus, which is 98 nucleotides long and is an almost absolutely conserved sequence[6].

Within the first, variable region of 3'UTR, two sequence motifs are present that are found in all genotypes of HCV. These are: the ACACUCC section, which represents a seed-region for miR-122[14], and the UG dinucleotide located at the very end of the region, directly upstream the poly (U/UC) section[6]. The stop codon is located in the apical loop of the stem-loop motif, named 5BSL3.4 or SL9360, which is created partly from the terminal nucleotides of the ORF encoding the NS5B protein and partly from non-coding nucleotides. The poly-pyrimidine section can be divided into a two poly (UC) parts uneven in length, which are separated by the one poly (U) region. The poly-pyrimidine segment is heterogeneous, not only in terms of length, but also in nucleotide sequence. In genotypes 2a, 3a, 3b there are several conserved adenosine residues in this region that are missing in genotypes 1b and 2b[6]. Individual guanosine residues are also observed on rare occasions.

3'X-tail

Initially, it was suspected that at the 3' terminus of the HCV genome a poly (U) or poly (A) sequence was

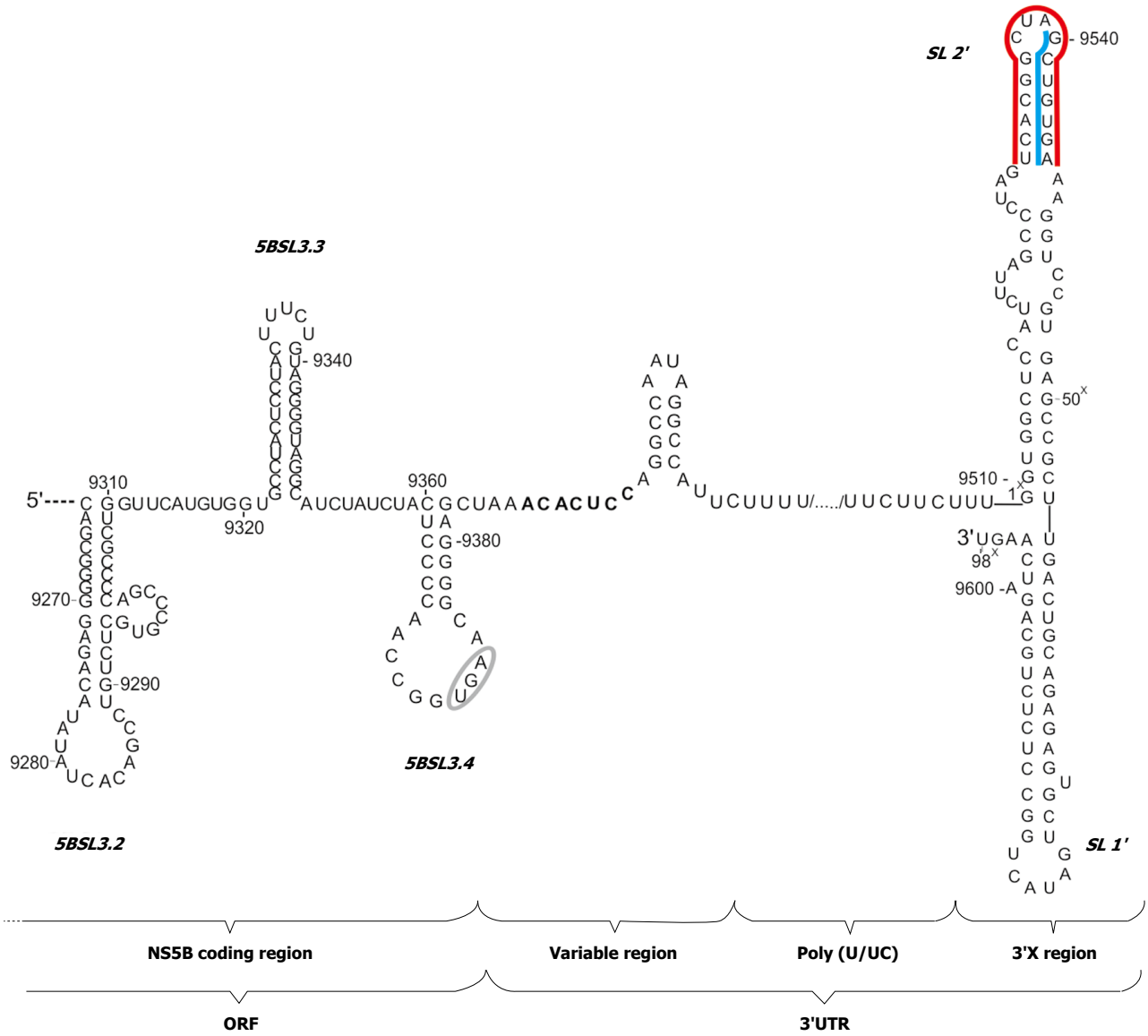

Figure 1 Secondary structure model of the 3'UTR of hepatitis C virus genome with adjacent 3' terminal sequence of the coding region. The stop codon is indicated by a gray ellipse. The regions involved in the kissing-interactions are marked with blue lines; the region involved in dimerization is indicated by a red line; the seed region for miR122 is highlighted in bold.

present. The existence of the 98-nt long 3'X region was discovered by Tanaka and colleagues in 1995[15]. Almost simultanously the presence of the 3'X at the end of the HCV genome identified Kolykhalov *et al*[16]. Comparison of the sequence of this RNA segment in different viral isolates indicated 96%-100% sequence conservation of the 3'X region, with only single substitutions in the 3' terminal 46-nt sequence[6]. This is an unusual feature for this dynamically changing virus, that suggests its extremely important function.

Several different secondary structure models of the 3'X region were proposed, based on computer predictions and experimental structure probing by chemical modification, enzyme digestion, RNA cleavage induced by Pb^{2+} ions, NMR (nuclear magnetic resonance) and SAXS (small angle X-ray scattering). A stable structure of the SL1 hairpin was proposed for the 3' part of the 3'X RNA, which is common to different structural models (Figure 2)[17-21]. The 52-nucleotide segment making up the 5' part of the 3'X region could not be assigned an unambiguous structure based on the experimental results obtained. The proposed models were only partially confirmed by the results of experimental studies[6,17,18]. Poor ordering of this fragment or formation of more than one structural form have been suggested[17,18,20].

One of the first structural models of the 3'X region suggested the presence of three hairpin motifs: SL1, SL2, and SL3 (Figure 2A)[17,18]. Another structural model suggested a set of four hairpins: SL1, SL2a, SL2b and SL3, where SL1 and SL3 did not differ from the first model, but two shorter hairpins replaced the SL2 motif[20]. The main reason for the proposed change was the observation of strong DMS modifications, as well as Pb^{2+} ion-induced cleavages, in the middle of the SL2 double-stranded stem (C44-C45), which indicated a single-stranded or highly flexible region there. In this four

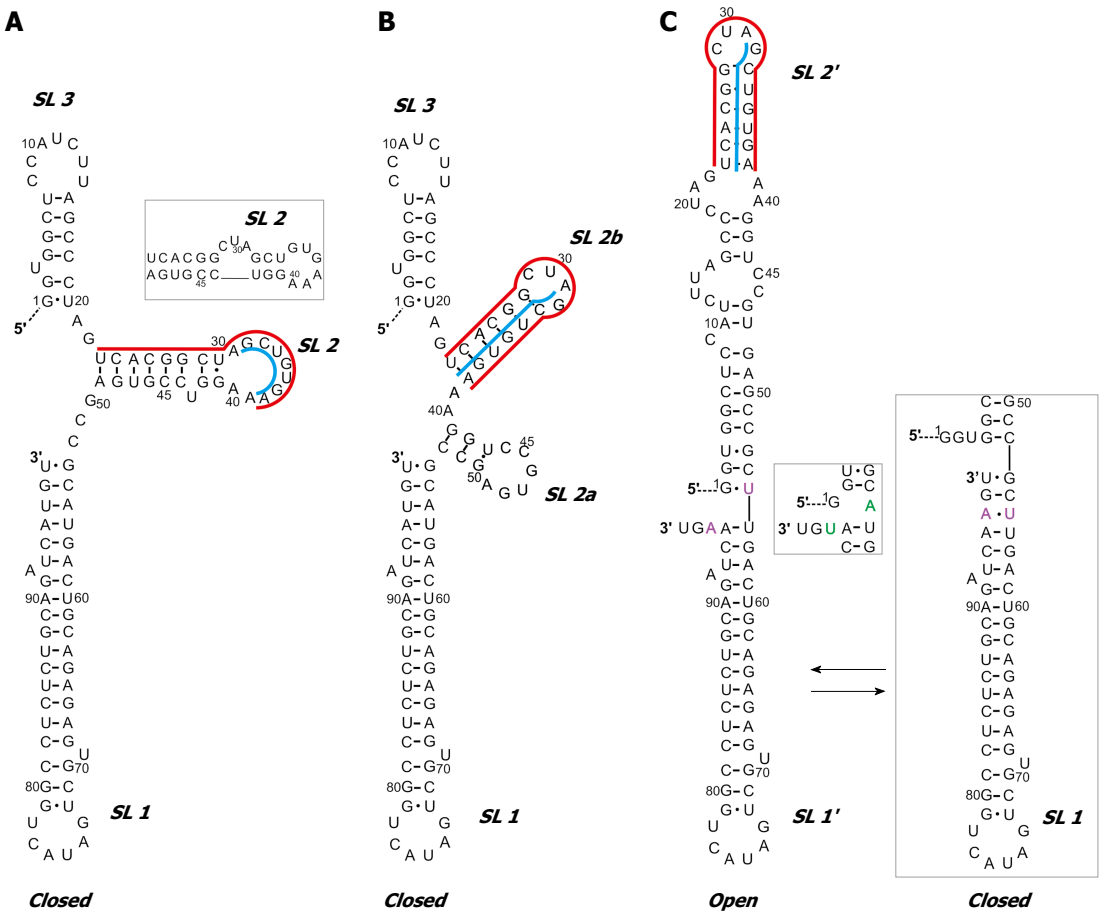

Figure 2 Diverse secondary structure models proposed for the 3'X region of hepatitis C virus genome. A: The 3xSL model[17,18]; B: The 4xSL model[20]; C: The 2xSL model[21]. The region involved in the kissing-interactions is indicated with blue line, the region involved in dimerization is marked with red line; possible alternative folding of separate fragments are displayed as gray rectangles.

stem-loop model (4xSL), the reactive cytidine residues are located in the apical loop of the short SL2b motif (Figure 2B). In addition, it was noted that the SL2a and SL2b hairpins have the potential to create a pseudoknot after rearrangement of their base pairing[20]. However, discovery of functionally important long-range kissing interactions between a sequence located in the apical loop of the SL2 hairpin and the upstream sequence in the region encoding NS5B[22] (see next section), seemed to support the three- (3xSL), but not the four- (4xSL) stem-loop model for the 3'X region. Thus, the 3xSL structure model of the 3'X-tail became the favored idea during the following decade.

Recently, based on NMR and SAXS studies, a two-hairpin model (2xSL) has been proposed for the 3'X region, consisting of the SL1 and SL2' (named elsewhere also SL2/3) (Figure 2C)[21,23,24]. In addition, in this model the SL1 hairpin may fold in two different ways: the closed structure - SL1 or the open one - SL1', differing by the three terminal base-pairs of the hairpin being paired or unpaired, respectively (Figure 2C)[21,23,25]. In the 2xSL model the previously identified flexible cytosine residues (C44-C45) are located within the internal loop of the SL2' motif. The closed conformation of SL1 is associated with long-range kissing interactions with another part of the genome, while the open conformation, SL1' with the 3' overhang is associated with dimerization (described in the next section).

The role of selected metal ions was investigated regarding its influence on the structure of the 3'X-tail but it seems that neither magnesium, nor sodium ion concentration determine its folding, within the range of normal physiological conditions[26]. However, at higher ionic strength, extended homodimers are preferentially formed over 2xSL monomers[24]. A chaperone role of the viral C protein (core protein), has also been suggested[23,27]. Long-range RNA-RNA interactions with 5' sequences in the genome or with a second genomic RNA molecule seems to influence the structure of the 3'X region more than the presence of specific metal ions.

It is very likely that the 3'X region can adopt more than one structural form in infected cells and that a specific equilibrium between these forms regulates several processes of the viral life cycle. These different structural forms may be favored by distinct viral genotypes what can help to explain their differential virulence and drug resistance[28,29].

Long range RNA-RNA interactions

Previous investigation of genomic HCV RNA and a con-

struct containing only 5'UTR and 3'UTR did not show any interaction of the X region with other regions of the molecules studied. This suggested the structural independence of these two regions of viral RNA from each other[17]. Later tertiary[30-32] interactions at the 3' end of the HCV genomic strand were proposed by Friebe *et al*[22]. The kissing interactions between the absolutely conserved "k" segment of the X region: 32^{X}-GCUGUGA-38^{X} and the "k'" segment of the NS5B coding sequence: 9281-UCACAGC-9287 do not require a protein chaperone. These sequence stretches are located within apical parts of SL elements called SL2 or SL2' and 5BSL3.2, which is one of the domains in the CRE (*cis*-acting replication element) (Figure 1)[22,28,31,33]. It is easy to imagine that kissing interactions could be initiated by any complementary stretch of nucleotides located within two apical loops of RNA hairpins. This scenario was previously suggested for the 3xSL model of the 3'X-tail[22,28,31,33]. However, recent studies with the use of mutagenesis, NMR and SAXS methods, indicate that before kissing-interactions are formed, the "k"- sequence in the X region is involved in base pairing within the SL2' element[21,24]. How the 5BSL3.2 element is able to induce the conformational transition from SL2' to SL2, or more globally: from 2xSL to 3xSL form of the 3'X, remains unclear. The next great challenge is to elucidate step by step how this transition occurs.

Two replicon systems, Con1b and JF-H1, have been investigated with a SHAPE method, which is based on a chemical modification of single-stranded RNA residues. The replicons are constructed on the basis of two different viral genotypes 1b and 2a, respectively. The experimental results showed that the proposed kissing interactions were detectable only for replicon JF-H1 (genotype 2a). The results obtained for replicon Con1b were in agreement with data obtained for the genotype 1a (strain H) of the virus and favors the open conformation of SL1'. In this open conformation the very 3' end of the SL' remains single-stranded, and that is associated with an increase in the efficiency of RNA synthesis initiation. In fact, subtype 1b is more virulent and resistant to interferon-based therapy than other genotypes, including subtype 2a[29,34]. This suggests how virus virulence and drug response is significantly influenced by the long-range kissing interactions, which likely cause changes in the base-pairing character of the very 3' terminal nucleotides.

Dimerization

Another intriguing set of tertiary interactions were proposed for the 3'X region of HCV genome by Ivanyi-Nagy *et al*[23]. Primarily *in vitro* investigations showed that the apical part of the SL2 hairpin in the 3xSL model, 29^{X}-CUAG-32^{X}, is able to interact with the respective palindromic sequence in the second 3'X RNA molecule thus inducing the formation of a homodimer[35]. The 16-nt palindromic sequence, called also DLS (dimerization leading stretch) is absolutely conserved among all HCV genotypes[36]. The homodimer, consisting of two isolated 3'X-RNA molecules, was characterized *in vitro* by NMR and SAXS. It was proposed that the resulting homoduplex could involve shorter (SL2) or extended (SL2') sequence fragments (Figure 3)[21,23-25,35]. The core protein supports the formation and stabilizes the extended homodimer[24].

The dimerization of the HCV genome still remains unproved *in vivo* and its function remains to be elucidated. It was suggested that the dimerization could be helpful in ensuring that only full-length progeny RNA molecules are encapsidated[24]. Moreover, the unwinding of the 3'-end of the genome, could greatly facilitate the minus RNA synthesis[37]. Masante *et al*[38] suggested that the homodimeric genome operates as a preferred template for the HCV polymerase (NS5B). Additionally, dimerization might enhance the rate of RNA recombination between two homologue RNA strands (resumed in[39]). The equilibrium between 2xSL monomers and dimers would likely also be tuned by the local concentration of RNA and the presence of core protein[24].

FUNCTIONS OF THE 3'UTR REGION

Hepatitis C virus can only infect humans and chimpanzees; there is no experimental model of its infectivity among small animals. For this reason, very few studies on the spread of the virus have been carried out *in vivo*. An experiment was carried out by Yanagi *et al*[40], in which a number of viral constructs containing deletions within the 3'UTR were injected into the liver of the chimpanzee at time intervals, and then the animal was examined for the presence of HCV RNA, anti-HCV antibodies and liver enzymes in its serum. Viral mutants lacking the entire 3'X region or parts thereof (nt: 1-50 and 57-98) were not able to replicate. Also viral infection was not observed when a mutant containing no poly (U/UC) segment was used. Only the construct devoid of 24 nucleotides within the variable region turned out to be infectious, indicating that this region is not essential for the viral life cycle under these conditions[40]. A similar experiment was carried out by Kolykhalov *et al*[41], and the obtained results were in line with previous observations of Yanagi *et al*[40].

One of the possible explanations of 3'X region function was its influence on genomic RNA stability. This has been shown *in vitro* but is not equally important for all genotypes of the virus[42].

Replication process

The information on the involvement of individual parts of the 3'UTR in the life cycle of the virus presented in the previous section was confirmed in research with the replicating Huh7 and HeLa cell systems[19,22,42,43]. Namely, constructs devoid of the entire poly (U/UC) section, were not able to replicate[42,43], and the minimum length of the polypyrimidine segment was 50 nt in one study[43], and only 26 nt in the other[42]. In contrast, the deletion of the

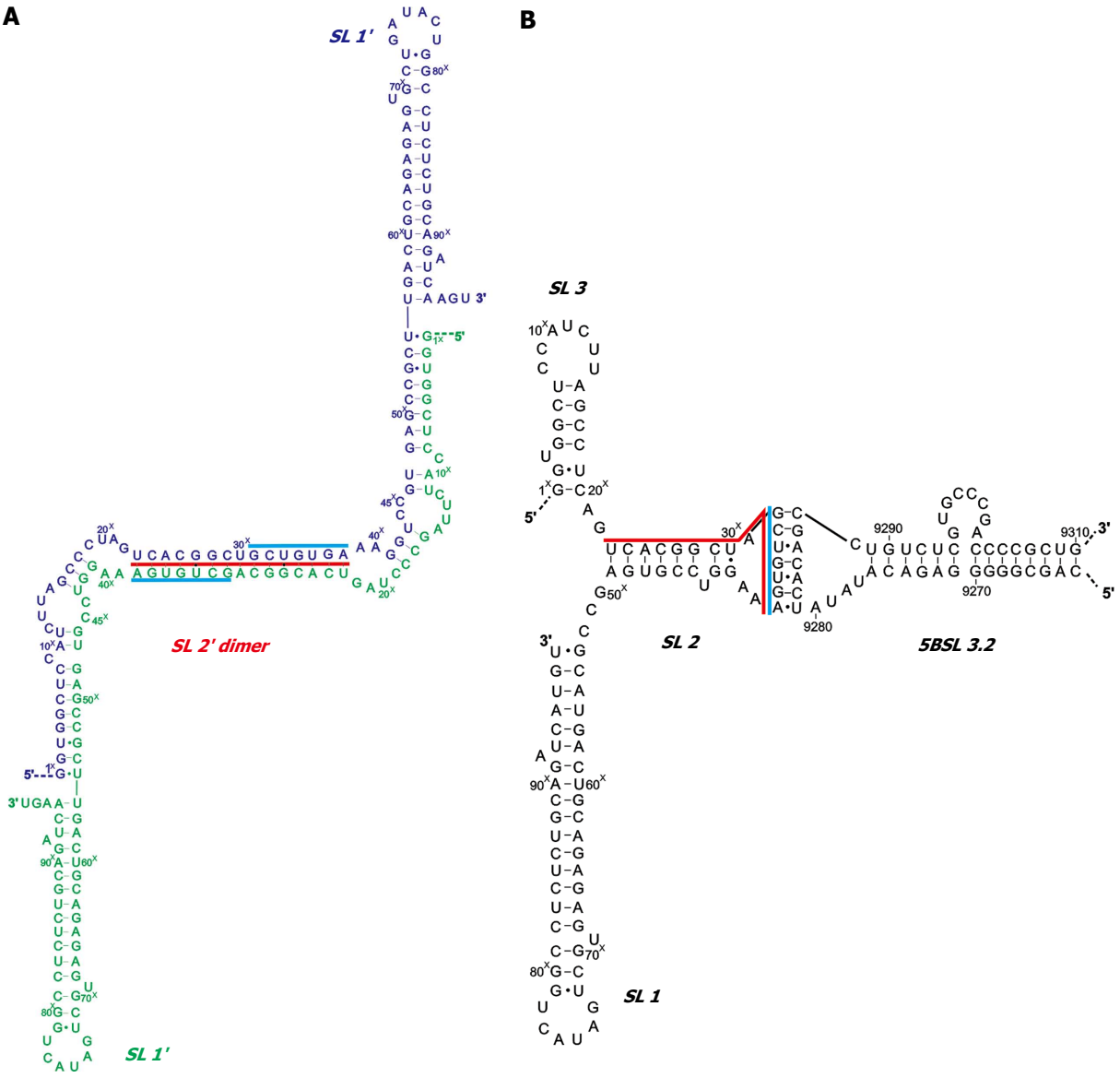

Figure 3 Long-range RNA-RNA interactions proposed for the 3'X region of hepatitis C virus genome. A: The homodimeric interactions between two 3'X regions embedded into two RNA molecules, model according to Cantero-Camacho *et al*[24]; B: The kissing-interactions with SL5B3.2[22,28,31,33]. Nucleotide sequence involved in dimerization is additionally indicated with red line, while those involved in the kissing-interactions are marked with blue lines.

3'X region or any one of its parts, SL1, SL2 or SL3, led to a complete failure to replicate[42,43] and only few point mutations in this region were tolerated[19,43]. In turn, the removal of the variable region from the 3'UTR of the viral genome only reduced the efficiency of the process leading to decrease in the rate of replication[22,42,43]. In addition, it has been shown that the region directly upstream the stop codon plays a key role in viral replication[22].

The site of the NS5B polymerase attachment within region 3'X has been mapped to be within the SL2 sequence and within the SL1[44]. They are protected against digestion with RNase T1 at guanosine residues at positions 41, 42, 50 and 53. The direct interaction of the 3'X RNA with the NS5B protein has been found in studies conducted *in vitro*[6,12,44-46]. The specificity of the viral polymerase to the model RNA template is relatively low and the presence of the 3'X region is not always necessary for RNA synthesis[6,12,37,45]. However, in the case of matrices containing genomic or subgenomic RNA, the presence of the 3'X region is necessary for the efficiency and the specificity of the process. It has been shown that the lack of this sequence, or a part thereof, almost completely inhibits replication, and in the case of deletion of nucleotides in positions 31-40, the product is too long[12,44].

The NS5B polymerase catalytic center probably only interacts with a single-stranded RNA fragment. The initiation of replication seems to take place in the SL1 loop, 21 nucleotides from the 3' end of the 3'X RNA[12]. However, another research group, Kim *et al*[46] indicates that the process begins near the 3' terminus, in a region rich in purines. However, in studies carried out in a cellular system, it has been shown that the presence

of the 3' terminal GU dinucleotide is preferred in the reaction[19]. Similar preferences of the replicase for the U at the 3'-end of the template RNA were observed by Shimm *et al*[47] in the *in vitro* system. Also, Kao *et al*[37] postulate that the enzyme requires a stable secondary template structure and at least one unpaired cytidine residue at its 3' end to initiate RNA synthesis. Butcher *et al*[48] proposed a replication initiation scenario, one common to various polymerases, in which the synthesis of a new strand of RNA begins with a nucleotide complementary to the penultimate of the 3' end of the nucleotide of the template molecule. Secondly, a complementary nucleotide is added to the remainder of the template at the 3' end, and only after the synthesis of this dinucleotide-primer is the complementary strand of viral RNA synthesized. Many observations suggest that NS5B is a non-specific enzyme, *i.e*. it can recognize more than one nucleotide sequence. The presence of specific Ⅰ-, Ⅱ- and tertiary structures at the 3'-end of the template RNA may allow modulation of the specificity of the enzyme for better yield and/or greater precision in the selection of the origin. In addition, both viral and host proteins can be involved in the replication process[6,49].

In summary, it should be noted that this replication phase of the life cycle of the virus has not yet been precisely understood. In addition to interacting with the viral polymerase, the 3'X region also interacts with other proteins one of them being the NS3 viral protein[49]. Acting as a helicase, it is probably a very important component of the replication complex, because the NS5B polymerase tends to detach from the RNA template when it encounters very stable RNA secondary structures[50]. The specific interaction between NS3 and the HCV 3'UTR probably involves a large part of this region containing the sequence 3'X and the poly (U/UC) section, since none of these elements alone is sufficient to form a stable complex with this protein[49].

Translation process

One of the earliest known cellular proteins interacting with HCV 3'UTR was the PTB protein. The interaction of PTB with the 3'X region may be related to the regulation of the translation process. The proposed site of PTB binding is 21 nucleotides from the 5' end of the 3'X region (SL3), containing a part of the consensus sequence recognized by this protein[18]. Similar results are presented by Tsuchihara *et al*[51] suggesting that the region involved in the interaction with PTB is 19 nucleotides from the 5' end of the 3'X region and extends 7 nucleotides upstream this sequence. The mutagenesis data showed that both secondary structure and the nucleotide sequence in the SL2 and SL3 regions are important for this interaction[18,51].

The 3'UTR has been ascribed as a translation enhancer[52-56]. One of mechanisms proposed by which enhancing the process is achieved, is the interaction between the 5' and 3' ends of the HCV genome, mediated by the PTB protein and another, hypothetical Y protein[57]. The effect of enhancing translation through genome cyclization is observed in many viruses[58]. However, not all results indicate PTB enhances HCV translation, sometimes the protein seems to be even inhibitory[59]. There is also no consensus on the impact of the 3'UTR on the efficiency of translation. For example, Murakami *et al*[59] observed that in an *in vitro* system, deletion of SL3 in the 3'X region and/or the poly (U/UC) section resulted in an increase in the amount of protein produced. A possible explanation for this observation would be that these sequence fragments, by interacting with other protein factors, inhibit the translation process. In other studies, however, it was observed that the presence of 3'UTR had no effect on the increase of the product amount and the efficiency of the polyprotein cleavage process[43,60]. It appears the choice of the experimental model plays a key role in this kind of research. For example, while in the lysate from rabbit reticulocytes no effect of 3'UTR on translation efficiency is observed[42,53], in HeLa cells, hepatocytes and in an *in vivo* mouse model the presence of the 3'UTR stimulates this process[53]. Recently, the role of 3'UTR in enhancing translational efficiency has been reported both in the rabbit reticulocyte lysate and in Huh7 cells[61].

Finally, it has been shown that the region X interacts with the ribosomal proteins L22, L3, S3 and mL3[62]. Similar viral RNA - L22 protein interactions have been observed for EBV (Epstein-Barr virus) and HPV-1 (human type 1 papilloma virus), showing a positive effect of L22 on translational efficiency[62]. Meanwhile, in the case of Qβ virus, ribosomal host proteins are involved in the formation of the viral replication complex[62]. It is unclear whether the interaction of ribosomal proteins with the HCV X region is related to the replication or translation process, or to both of these processes. The demonstrated ability of HCV protein NS5B to bind to the ribosome[63] is reminiscent of the strategy used by the Qβ virus, which uses ribosomal proteins to build its replication complex.

CONCLUSION

The most intriguing feature of the 3'X region is its extremely high sequence conservation. The maintenance of a 98-nt long stretch of RNA at over 95% conservation places severe constraints on such a fast-mutating RNA virus. Apparently almost any change to this stretch result in progeny incapable of reproducing which rapidly disappear from the population.

The high sequence conservation might be explained by existence of different structural forms of the 3'X region. Within one structure there is usually a way to neutralize point mutations by adaptive mutations in another structure, thus restoring important base-pairing. Simultaneous compensation of mutations in two (or more) different structures/forms is much less probable. However, instead of several different structural forms of the X-region, the long-range RNA-RNA interactions with the involvement of that region could explain its high

sequence conservation.

Evidence indicates the 3'X region is involved in a number of interactions including kissing 5BSL3.2 (required for replication), an interaction with the second RNA molecule carrying the X-region sequence (homodimerization), binding of the NS5B protein (for initiation of RNA synthesis), an interaction with NS3 helicase (for elongation of RNA synthesis), binding of PTB (for proposed translation regulation). Is it possible that one structural form of the 3'X region supports such different interactions? The answer is: yes, it is possible, but unlikely. On the other hand, it is easy to imagine that different structural forms of this region are responsible for different processes and interactions. For instance: 2xSL - structure for dimerization and enhanced RNA synthesis, 3xSL - structure for kissing interactions that supports translation.

Hopefully in-depth *in vivo* structure mapping of the 3'X-region at the various replication stages, in different cellular compartments will answer these questions to give us a better understanding of the virus and lead to additional means to control it.

ACKNOWLEDGMENTS

The authors thank Erik Wade for careful reading of the manuscript and valuable comments.

REFERENCES

1 **Choo KB**, Liew LN, Wang HC. Production of transgenic mice carrying the human hepatitis B virus or the human papillomavirus DNA sequences in Taiwan: analysis of physical structure and hereditary mode. *Proc Natl Sci Counc Repub China B* 1989; **13**: 314-318 [PMID: 2561574]

2 **Bartenshlager R**. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Current Topics in Microbiology and Immunology* 2013; **369**: V-VI [DOI: 10.1007/978-3-642-27340-7]

3 **Mohd Hanafiah K**, Groeger J, Flaxman AD, Wiersma ST. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* 2013; **57**: 1333-1342 [PMID: 23172780 DOI: 10.1002/hep.26141]

4 **Dev A**, Patel K, McHutchison JG. New therapies for chronic hepatitis C virus infection. *Curr Gastroenterol Rep* 2004; **6**: 77-86 [PMID: 14720458 DOI: 10.1007/s11894-004-0030-5]

5 **Rijnbrand RC**, Lemon SM. Internal ribosome entry site-mediated translation in hepatitis C virus replication. *Curr Top Microbiol Immunol* 2000; **242**: 85-116 [PMID: 10592657 DOI: 10.1007/978-3-642-59605-6_5]

6 **Shi ST**, Lai MM. Hepatitis C viral RNA: challenges and promises. *Cell Mol Life Sci* 2001; **58**: 1276-1295 [PMID: 11577984 DOI: 10.1007/PL00000939]

7 **Laskus T**, Radkowski M, Wang LF, Vargas H, Rakela J. The presence of active hepatitis C virus replication in lymphoid tissue in patients coinfected with human immunodeficiency virus type 1. *J Infect Dis* 1998; **178**: 1189-1192 [PMID: 9806058 DOI: 10.1086/515682]

8 **Conca P**, Cafaro G, De Renzo A, Coppola A, Cimino E, Tarantino G. Successful and Safe Long-Term Standard Antiviral Therapy in a Patient with „Explosive" Immune Response in Course of HCV-Related Liver Cirrhosis. *Int J Mol Sci* 2015; **16**: 14075-14085 [PMID: 26101866 DOI: 10.3390/ijms160614075]

9 **Bell AM**, Wagner JL, Barber KE, Stover KR. Elbasvir/Grazoprevir: A Review of the Latest Agent in the Fight against Hepatitis C. *Int J Hepatol* 2016; **2016**: 3852126 [PMID: 27403342 DOI: 10.1155/2016/3852126]

10 **Gómez J**, Nadal A, Sabariegos R, Beguiristain N, Martell M, Piron M. Three properties of the hepatitis C virus RNA genome related to antiviral strategies based on RNA-therapeutics: variability, structural conformation and tRNA mimicry. *Curr Pharm Des* 2004; **10**: 3741-3756 [PMID: 15579068 DOI: 10.2174/1381612043382675]

11 **Randall G**, Rice CM. Interfering with hepatitis C virus RNA replication. *Virus Res* 2004; **102**: 19-25 [PMID: 15068876 DOI: 10.1016/j.virusres.2004.01.011]

12 **Oh JW**, Ito T, Lai MM. A recombinant hepatitis C virus RNA-dependent RNA polymerase capable of copying the full-length viral RNA. *J Virol* 1999; **73**: 7694-7702 [PMID: 10438859]

13 **Penin F**, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM. Structural biology of hepatitis C virus. *Hepatology* 2004; **39**: 5-19 [PMID: 14752815 DOI: 10.1002/hep.20032]

14 **Gerresheim GK**, Dünnes N, Nieder-Röhrmann A, Shalamova LA, Fricke M, Hofacker I, Höner Zu Siederdissen C, Marz M, Niepmann M. microRNA-122 target sites in the hepatitis C virus RNA NS5B coding region and 3' untranslated region: function in replication and influence of RNA secondary structure. *Cell Mol Life Sci* 2017; **74**: 747-760 [PMID: 27677491 DOI: 10.1007/s00018-016-2377-9]

15 **Tanaka T**, Kato N, Cho MJ, Shimotohno K. A novel sequence found at the 3' terminus of hepatitis C virus genome. *Biochem Biophys Res Commun* 1995; **215**: 744-749 [PMID: 7488017 DOI: 10.1006/bbrc.1995.2526]

16 **Kolykhalov AA**, Feinstone SM, Rice CM. Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *J Virol* 1996; **70**: 3363-3371 [PMID: 8648666]

17 **Blight KJ**, Rice CM. Secondary structure determination of the conserved 98-base sequence at the 3' terminus of hepatitis C virus genome RNA. *J Virol* 1997; **71**: 7345-7352 [PMID: 9311812]

18 **Ito T**, Lai MM. Determination of the secondary structure of and cellular protein binding to the 3'-untranslated region of the hepatitis C virus RNA genome. *J Virol* 1997; **71**: 8698-8706 [PMID: 9343228]

19 **Yi M**, Lemon SM. Structure-function analysis of the 3' stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. *RNA* 2003; **9**: 331-345 [PMID: 12592007 DOI: 10.1261/rna.2144203]

20 **Dutkiewicz M**, Ciesiolka J. Structural characterization of the highly conserved 98-base sequence at the 3' end of HCV RNA genome and the complementary sequence located at the 5' end of the replicative viral strand. *Nucleic Acids Res* 2005; **33**: 693-703 [PMID: 15681619 DOI: 10.1093/nar/gki218]

21 **Cantero-Camacho Á**, Gallego J. The conserved 3'X terminal domain of hepatitis C virus genomic RNA forms a two-stem structure that promotes viral RNA dimerization. *Nucleic Acids Res* 2015; **43**: 8529-8539 [PMID: 26240378 DOI: 10.1093/nar/gkv786]

22 **Friebe P**, Boudet J, Simorre JP, Bartenschlager R. Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication. *J Virol* 2005; **79**: 380-392 [PMID: 15596831 DOI: 10.1128/jvi.79.1.380-392.2005]

23 **Ivanyi-Nagy R**, Kanevsky I, Gabus C, Lavergne JP, Ficheux D, Penin F, Fossé P, Darlix JL. Analysis of hepatitis C virus RNA dimerization and core-RNA interactions. *Nucleic Acids Res* 2006; **34**: 2618-2633 [PMID: 16707664 DOI: 10.1093/nar/gkl240]

24 **Cantero-Camacho Á**, Fan L, Wang YX, Gallego J. Three-dimensional structure of the 3'X-tail of hepatitis C virus RNA in monomeric and dimeric states. *RNA* 2017; **23**: 1465-1476 [PMID: 28630140 DOI: 10.1261/rna.060632.117]

25 **Cristofari G**, Ivanyi-Nagy R, Gabus C, Boulant S, Lavergne JP, Penin F, Darlix JL. The hepatitis C virus Core protein is a potent nucleic acid chaperone that directs dimerization of the viral (+) strand RNA in vitro. *Nucleic Acids Res* 2004; **32**: 2623-2631 [PMID: 15141033 DOI: 10.1093/nar/gkh579]

26 **Kranawetter C**, Brady S, Sun L, Schroeder M, Chen SJ, Heng X. Nuclear Magnetic Resonance Study of RNA Structures at the 3'-End of the Hepatitis C Virus Genome. *Biochemistry* 2017; **56**: 4972-4984 [PMID: 28829576 DOI: 10.1021/acs.biochem.7b00573]

27 **Sharma KK**, de Rocquigny H, Darlix JL, Lavergne JP, Pénin F,

Lessinger JM, Mély Y. Analysis of the RNA chaperoning activity of the hepatitis C virus core protein on the conserved 3'X region of the viral genome. *Nucleic Acids Res* 2012; **40**: 2540-2553 [PMID: 22127859 DOI: 10.1093/nar/gkr1140]

28 **Tuplin A**, Struthers M, Simmonds P, Evans DJ. A twist in the tail: SHAPE mapping of long-range interactions and structural rearrangements of RNA elements involved in HCV replication. *Nucleic Acids Res* 2012; **40**: 6908-6921 [PMID: 22561372 DOI: 10.1093/nar/gks370]

29 **Carter LDB**, Aronsohn A. Overcoming injustice: A roadmap to improve access to hepatitis C virus therapy for our medicaid patients. *Hepatology* 2017; **65**: 1735-1740 [PMID: 28160311 DOI: 10.1002/hep.29095]

30 **Diviney S**, Tuplin A, Struthers M, Armstrong V, Elliott RM, Simmonds P, Evans DJ. A hepatitis C virus cis-acting replication element forms a long-range RNA-RNA interaction with upstream RNA sequences in NS5B. *J Virol* 2008; **82**: 9008-9022 [PMID: 18614633 DOI: 10.1128/JVI.02326-07]

31 **You S**, Rice CM. 3' RNA elements in hepatitis C virus replication: kissing partners and long poly(U). *J Virol* 2008; **82**: 184-195 [PMID: 17942554 DOI: 10.1128/JVI.01796-07]

32 **You S**, Stump DD, Branch AD, Rice CM. A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for hepatitis C virus RNA replication. *J Virol* 2004; **78**: 1352-1366 [PMID: 14722290 DOI: 10.1128/jvi.78.3.1352-1366.2004]

33 **Tuplin A**, Struthers M, Cook J, Bentley K, Evans DJ. Inhibition of HCV translation by disrupting the structure and interactions of the viral CRE and 3' X-tail. *Nucleic Acids Res* 2015; **43**: 2914-2926 [PMID: 25712095 DOI: 10.1093/nar/gkv142]

34 **Irshad M**, Mankotia DS, Irshad K. An insight into the diagnosis and pathogenesis of hepatitis C virus infection. *World J Gastroenterol* 2013; **19**: 7896-7909 [PMID: 24307784 DOI: 10.3748/wjg.v19.i44.7896]

35 **Shetty S**, Kim S, Shimakami T, Lemon SM, Mihailescu MR. Hepatitis C virus genomic RNA dimerization is mediated via a kissing complex intermediate. *RNA* 2010; **16**: 913-925 [PMID: 20360391 DOI: 10.1261/rna.1960410]

36 **Yamada N**, Tanihara K, Takada A, Yorihuzi T, Tsutsumi M, Shimomura H, Tsuji T, Date T. Genetic organization and diversity of the 3' noncoding region of the hepatitis C virus genome. *Virology* 1996; **223**: 255-261 [PMID: 8806561 DOI: 10.1006/viro.1996.0476]

37 **Kao CC**, Yang X, Kline A, Wang QM, Barket D, Heinz BA. Template requirements for RNA synthesis by a recombinant hepatitis C virus RNA-dependent RNA polymerase. *J Virol* 2000; **74**: 11121-11128 [PMID: 11070008 DOI: 10.1128/JVI.74.23.11121-11128.2000]

38 **Masante C**, Jaubert C, Palau W, Plissonneau J, Besnard L, Ventura M, Di Primo C. Mutations of the SL2 dimerization sequence of the hepatitis C genome abrogate viral replication. *Cell Mol Life Sci* 2015; **72**: 3375-3385 [PMID: 25822205 DOI: 10.1007/s00018-015-1893-3]

39 **Romero-López C**, Barroso-delJesus A, Berzal-Herranz A. The chaperone-like activity of the hepatitis C virus IRES and CRE elements regulates genome dimerization. *Sci Rep* 2017; **7**: 43415 [PMID: 28233845 DOI: 10.1038/srep43415]

40 **Yanagi M**, St Claire M, Emerson SU, Purcell RH, Bukh J. In vivo analysis of the 3' untranslated region of the hepatitis C virus after in vitro mutagenesis of an infectious cDNA clone. *Proc Natl Acad Sci U S A* 1999; **96**: 2291-2295 [PMID: 10051634 DOI: 10.1073/pnas.96.5.2291]

41 **Kolykhalov AA**, Mihalik K, Feinstone SM, Rice CM. Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo. *J Virol* 2000; **74**: 2046-2051 [PMID: 10644379 DOI: 10.1128/JVI.74.4.2046-2051.2000]

42 **Yi M**, Lemon SM. 3' nontranslated RNA signals required for replication of hepatitis C virus RNA. *J Virol* 2003; **77**: 3557-3568 [PMID: 12610131 DOI: 10.1128/JVI.77.6.3557-3568.2003]

43 **Friebe P**, Bartenschlager R. Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. *J Virol* 2002; **76**: 5326-5338 [PMID: 11991961 DOI: 10.1128/JVI.76.11.5326-5338.2002]

44 **Oh JW**, Sheu GT, Lai MM. Template requirement and initiation site selection by hepatitis C virus polymerase on a minimal viral RNA template. *J Biol Chem* 2000; **275**: 17710-17717 [PMID: 10749880 DOI: 10.1074/jbc.M908781199]

45 **Reigadas S**, Ventura M, Sarih-Cottin L, Castroviejo M, Litvak S, Astier-Gin T. HCV RNA-dependent RNA polymerase replicates in vitro the 3' terminal region of the minus-strand viral RNA more efficiently than the 3' terminal region of the plus RNA. *Eur J Biochem* 2001; **268**: 5857-5867 [PMID: 11722573 DOI: 10.1046/j.0014-2956.2001.02532.x]

46 **Kim M**, Kim H, Cho SP, Min MK. Template requirements for de novo RNA synthesis by hepatitis C virus nonstructural protein 5B polymerase on the viral X RNA. *J Virol* 2002; **76**: 6944-6956 [PMID: 12072495 DOI: 10.1128/JVI.76.14.6944-6956.2002]

47 **Shim JH**, Larson G, Wu JZ, Hong Z. Selection of 3'-template bases and initiating nucleotides by hepatitis C virus NS5B RNA-dependent RNA polymerase. *J Virol* 2002; **76**: 7030-7039 [PMID: 12072503 DOI: 10.1128/JVI.76.14.7030-7039.2002]

48 **Butcher SJ**, Grimes JM, Makeyev EV, Bamford DH, Stuart DI. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* 2001; **410**: 235-240 [PMID: 11242087 DOI: 10.1038/35065653]

49 **Banerjee R**, Dasgupta A. Specific interaction of hepatitis C virus protease/helicase NS3 with the 3'-terminal sequences of viral positive- and negative-strand RNA. *J Virol* 2001; **75**: 1708-1721 [PMID: 11160669 DOI: 10.1128/jvi.75.4.1708-1721.2001]

50 **DeMarini DJ**, Johnston VK, Konduri M, Gutshall LL, Sarisky RT. Intracellular hepatitis C virus RNA-dependent RNA polymerase activity. *J Virol Methods* 2003; **113**: 65-68 [PMID: 14500128 DOI: 10.1016/S0166-0934(03)00226-X]

51 **Tsuchihara K**, Tanaka T, Hijikata M, Kuge S, Toyoda H, Nomoto A, Yamamoto N, Shimotohno K. Specific interaction of polypyrimidine tract-binding protein with the extreme 3'-terminal structure of the hepatitis C virus genome, the 3'X. *J Virol* 1997; **71**: 6720-6726 [PMID: 9261396]

52 **Ito T**, Tahara SM, Lai MM. The 3'-untranslated region of hepatitis C virus RNA enhances translation from an internal ribosomal entry site. *J Virol* 1998; **72**: 8789-8796 [PMID: 9765423]

53 **McCaffrey AP**, Ohashi K, Meuse L, Shen S, Lancaster AM, Lukavsky PJ, Sarnow P, Kay MA. Determinants of hepatitis C translational initiation in vitro, in cultured cells and mice. *Mol Ther* 2002; **5**: 676-684 [PMID: 12027551 DOI: 10.1006/mthe.2002.0600]

54 **Bradrick SS**, Walters RW, Gromeier M. The hepatitis C virus 3'-untranslated region or a poly(A) tract promote efficient translation subsequent to the initiation phase. *Nucleic Acids Res* 2006; **34**: 1293-1303 [PMID: 16510853 DOI: 10.1093/nar/gkl019]

55 **Song Y**, Friebe P, Tzima E, Jünemann C, Bartenschlager R, Niepmann M. The hepatitis C virus RNA 3'-untranslated region strongly enhances translation directed by the internal ribosome entry site. *J Virol* 2006; **80**: 11579-11588 [PMID: 16971433 DOI: 10.1128/jvi.00675-06]

56 **Bung C**, Bochkaeva Z, Terenin I, Zinovkin R, Shatsky IN, Niepmann M. Influence of the hepatitis C virus 3'-untranslated region on IRES-dependent and cap-dependent translation initiation. *FEBS Lett* 2010; **584**: 837-842 [PMID: 20079737 DOI: 10.1016/j.febslet.2010.01.015]

57 **Ito T**, Lai MM. An internal polypyrimidine-tract-binding protein-binding site in the hepatitis C virus RNA attenuates translation, which is relieved by the 3'-untranslated sequence. *Virology* 1999; **254**: 288-296 [PMID: 9986795 DOI: 10.1006/viro.1998.9541]

58 **Villordo SM**, Gamarnik AV. Genome cyclization as strategy for flavivirus RNA replication. *Virus Res* 2009; **139**: 230-239 [PMID: 18703097 DOI: 10.1016/j.virusres.2008.07.016]

59 **Murakami K**, Abe M, Kageyama T, Kamoshita N, Nomoto A. Down-regulation of translation driven by hepatitis C virus internal ribosomal entry site by the 3' untranslated region of RNA. *Arch Virol* 2001; **146**: 729-741 [PMID: 11402859 DOI: 10.1007/s007050170142]

60 **Imbert I**, Dimitrova M, Kien F, Kieny MP, Schuster C. Hepatitis C virus IRES efficiency is unaffected by the genomic RNA 3'NTR even in the presence of viral structural or non-structural proteins. *J Gen Virol* 2003; **84**: 1549-1557 [PMID: 12771425 DOI: 10.1099/vir.0.18907-0]

61 **Wang H**, Shen XT, Ye R, Lan SY, Xiang L, Yuan ZH. Roles of the polypyrimidine tract and 3' noncoding region of hepatitis C virus RNA in the internal ribosome entry site-mediated translation. *Arch Virol* 2005; **150**: 1085-1099 [PMID: 15747050 DOI: 10.1007/s00705-005-0491-3]

62 **Wood J**, Frederickson RM, Fields S, Patel AH. Hepatitis C virus 3' X region interacts with human ribosomal proteins. *J Virol* 2001; **75**: 1348-1358 [PMID: 11152508 DOI: 10.1128/jvi.75.3.1348-1358.2001]

63 **Tanaka T**, Sugiyama K, Ikeda M, Naganuma A, Nozaki A, Saito M, Shimotohno K, Kato N. Hepatitis C virus NS5B RNA replicase specifically binds ribosomes. *Microbiol Immunol* 2000; **44**: 543-550 [PMID: 10941940 DOI: 10.1111/j.1348-0421.2000.tb02532.x]

P- Reviewer: Lei YC, Tamori A, Tarantino G, Wang L
S- Editor: Wang XJ **L- Editor**: A **E- Editor**: Yin SY

Published by **Baishideng Publishing Group Inc**
7901 Stoneridge Drive, Suite 501, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.f6publishing.com/helpdesk
http://www.wjgnet.com

ISSN 1007-9327