Targets and Tools: Recent Advances in the Development of Anti-HCV Nucleic Acids

C. Romero-López, F. J. Sánchez-Luque and A. Berzal-Herranz*

Instituto de Parasitología y Biomedicina "López-Neyra", CSIC, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento s/n, Armilla 18100, Granada, Spain

Abstract: Hepatitis C virus (HCV), the major etiological agent of transfusion-associated non-A, non-B hepatitis, is a severe health problem affecting up to 3% of the world population. Since its identification in 1989, enormous efforts have been made to characterize the viral cycle. However, many details regarding the virus' penetration of hepatocytes, its replication and translation, and the assembling of virions remain unknown, mostly because of a lack of an efficient culture system. This has also hampered the development of fully effective antiviral drugs. Current treatments based on the combination of interferon and ribavirin trigger a sustained virological response in only 40% of infected individuals, thus the development of alternative therapeutic strategies is a major research goal. Nucleic acid based therapeutic agents may be of some potential in hepatitis C treatment. In recent years, much effort has gone into the improvement of DNA and RNA molecules as specific gene silencing tools. This review summarizes the state of the art in the development of new HCV therapies, paying special attention to those involving antisense oligonucleotides, aptamers, ribozymes, decoys and siRNA inhibitors. The identification of potential viral targets is also discussed.

Keywords: HCV, hepatitis C, gene silencing, RNA-based inhibitors, ribozyme, antisense, aptamer, siRNA.

GENERAL FEATURES

Hepatitis C virus (HCV) infection is a major global health problem. More than 3% of the world population is affected (WHO data), although the incidence varies from one region to another. Once the virus has entered the body it infects the liver cells and begins to replicate. Some 15-25% of infected individuals develop effective barriers that finally resolve the infection, but the remainder suffers the long-term replication of the virus [1]. These patients usually remain asymptomatic for years before showing fibrosis that may, in some 30% of these cases, gradually progress to cirrhosis and even hepatocellular carcinoma. At this stage, the only treatment possible is a liver transplant; indeed, hepatitis C patients figure strongly on liver transplant waiting lists [2].

Current therapies, based on the combination of interferon (IFN) and ribavirin, are effective in only 40% of patients. The high mutation rate of the HCV genome and the virus' ability to escape the immune system are partially responsible for this failure. Thus the identification of new targets and the search for fully effective antiviral compounds are major goals of HCV research.

HCV belongs to the *Hepaciviruses* (family Flaviviridae), which include yellow fever virus (the classic flavivirus), bovine diarrhea virus (BVDV, a pestivirus) and GB virus. The HCV genome shows great variability which allows the identification of six genotypes [3] differing from one other by up to 30% in their nucleotide sequences (reviewed in [4]). Further subtypes and isolates of these different genotypes varying in their virulence have also been described. Viral genotype clearly affects the success of interferon therapy, although no clear correlation with virulence exists.

In addition, the HCV population present in an infected patient is structured in terms of *quasispecies*. This term defines the closely related sequences of a heterogeneous viral population infecting a single individual [5]. Genomic sequence diversity is mainly due to the high viral replication rate and the mutations introduced by viral RNA-dependent RNA polymerase (RdRp) during the replication of the viral genome [6, 7]. Quasispecies structure has been associated with the failure of infected people to clear the virus and the subsequent development of a chronic infection [8]. It has been reported that the viral populations infecting immunosuppressed patients show a reduced probability of mutation compared to that seen in non-immunocompromised individuals [9-11]. A strong host immune response may encourage the appearance of resistant variants.

HCV virions have a diameter of 55-65 nm [12] and can be detected in patient serum either complexed to immunoglobulins or low density lipoproteins, or as free viral particles [13]. The genetic material is bound to the core protein, forming an icosaedric capsid. This is in turn covered by a lipid envelope. A heterodimeric complex consisting of the viral glycoproteins E1 and E2, sticks out through this envelope like spikes [14].

The HCV genome is a 9600 nucleotide-long, single stranded positive RNA molecule (Fig. (**1**); [15-17]). It contains two highly conserved untranslatable regions (UTR) flanking the coding sequence, which play an essential role in viral translation and replication [18-21]. The viral proteins are synthesized as a single precursor that is then co- and post-translationally processed by both viral and cellular

*Address correspondence to this author at the Instituto de Parasitología y Biomedicina "López-Neyra" CSIC, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento s/n, Armilla 18100, Granada, Spain; Tel: +34 958 18 16 48; Fax: +34 958 18 16 32; E-mail: aberzalh@ipb.csic.es

1871-5265/06 $50.00+.00

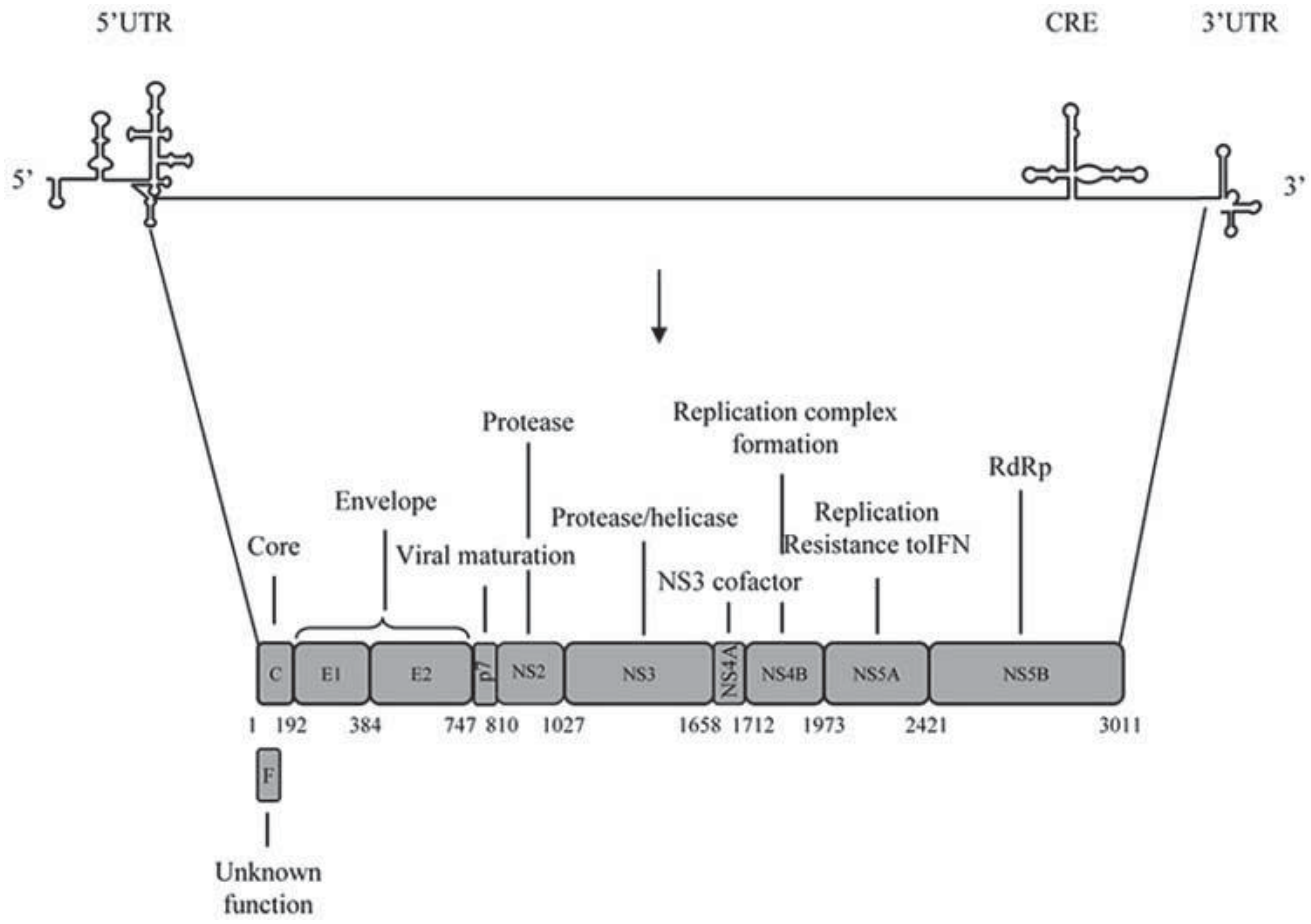

Fig. (1). Genetic organization and protein products of HCV. Above: the HCV genome, showing the 5' and 3' untranslated regions. The proposed secondary structure for the CRE domain (*cis*-response element) is also shown. IRES-dependent translation generates a 3011 amino acid-long polyprotein product (below) which is subsequently processed into structural and non-structural proteins. The proteolysis products and their functions are indicated.

proteases. The structural proteins include the capsid protein (C), p7 and the envelope proteins E1 and E2. The non-structural products are involved in polyprotein processing (NS2, NS3, NS4A) and replication (NS4B, NS5A and NS5B).

The success of HCV infection is conditioned by how well the virus avoids the immune host response. The initial defense reaction depends on the viral pathogen-associated molecular pattern (PAMP). For HCV, the PAMP includes the genomic RNA, the NS5A protein, the core protein, and even entire viral particles, all of which can be presented by dendritic cells (for a review, see [22]). PAMP presentation initiates the production of interferon and cytokines [23] in an autoregulatory fashion ([24]; reviewed in [22]). It has been reported that the NS3/4A protein acts as an antagonist of certain key factors in the induction of the immune response. This mainly involves a reduction in IFN production, thus inducing alterations in the pathways and complexes whose functioning is dependent on it. For example, antigen presentation is greatly reduced, meaning T cells cannot be stimulated. NS3/4A has also been related to a prolonged blocking of pro-apoptotic factor activity, providing a molecular link between HCV infection and the development of hepatocellular carcinoma [25]. Similarly, the NS5A protein may activate IL-8 production, which interferes with IFN-mediated response pathways, and in conjunction with E2 it may block protein kinase R (PKR), an enzyme involved in the amplification of the immune response. The intensity of these effects differs depending on the genetic diversity of the infecting viral pool. Finally, HCV genotypes 1a and 1b have a smaller number of UU and AU sites than genotypes 2 or 3, impeding RNase L-mediated degradation of the HCV genome (this enzyme specifically cleaves nucleic acids at these sites; [26]). This provides another possibility for variation in terms of resistance to IFN since RNase L is activated by this molecule.

The identification of HCV in 1989 by Houghton and co-workers [15] led to numerous studies aimed at deciphering its molecular biology, but advances in our knowledge regarding the viral cycle have been hindered by the lack of an efficient and stable culture system. However, recent advances in replicon and pseudoparticle production methods [27, 28] have allowed a more detailed understanding of HCV spread and replication to be grasped. In addition, the latest virion production strategies provide the starting point for the

identification and evaluation of new viral targets and the development of novel antiviral drugs [29-31].

This review discusses the main features and functions of the HCV targets that are being explored, and the therapeutic agents presently in use. New treatments based on the use of nucleic acids, potentially very powerful anti-HCV agents, are discussed.

VIRAL TARGETS

The high mutation rate of the HCV genome is a key determinant in the appearance of variants resistant to many of the antiviral compounds tested to date. Over recent years, much effort has been spent on the identification of new therapeutic targets and the development of effective anti-HCV drugs. The most common drug targets under study are the 5' and 3' untranslated regions (5'UTR, 3'UTR) and the NS3 and NS5B proteins; these are important in the viral cycle and their conservation rates are high. This section summarizes the main features and roles of these elements in the viral cycle.

THE 5' UNTRANSLATED REGION

This is the most commonly chosen target for the design of HCV inhibitors based on viral nucleic acid chemistry. This 341 nt-long region is one of the most conserved of the entire HCV genome: sequence identity is estimated at around 85% for all viral isolates [32, 33]. Domains essential for the initiation of viral replication [20] and translation [18, 19] have been identified within the 5' UTR. The synthesis of the HCV polyprotein is mediated by an internal ribosome entry site domain (IRES; from nt 40 to 370) in a cap-independent manner (Fig. (**2**); [34]). Thus, the 5' sequence coding for the capsid protein participates in the initiation of translation and may modulate the efficiency of this IRES [35]. Moreover, it has recently been reported that the N-terminal core domain might interact with the 5'UTR region and regulate viral translation [36, 37].

The secondary structure of the 5'UTR region was originally proposed by Lemon and co-workers [38]. However, further structural and functional analyses prompted its redefinition to include several new motifs and interactions. The most commonly accepted conformation is shown in Fig. (**2**). The 5'UTR region is folded into several stem-loop motifs that define different functional domains. The establishment of tertiary interactions among them has been described. Its maintenance is critical for IRES activity [39].

- *Domain I* of the 5'UTR region is crucial for the replication of the HCV genome [20]. It modulates viral polyprotein translation but it is not essential for IRES-mediated functions [40, 41].

- *The linker region between domains I and II* appears to be an unstructured, single-stranded sequence, but in fact it takes part in long-range RNA-RNA interactions involving the nucleotides at positions G_{432}-G_{438} [42]. The resulting duplex favors the creation of a large loop-like structure in which the IRES region is enclosed. Its formation might, in some way, modulate IRES-dependent protein synthesis. This is in good agreement with previous reports describing that mutations at this level affect IRES-dependent protein synthesis [43].

- *Domain II* of the 5'UTR region consists of two helical segments separated by a highly conserved internal loop [44, 45] thought to participate in long-range interactions with other viral RNA domains [39]. This loop may also provide a site for the association of translation initiation factors [46, 47]. The nucleotides of the apical loop of domain II define an essential location where protein and RNA ligands can attach [39, 48-50].

- *Domain III* of the 5'UTR region is the anchorage site for the eIF3 factor and the 40S ribosomal subunit [50, 51]. It is composed of several subdomains that fold into stem-loop structures able to interact with other domains of the IRES [39]. The IIIabc subdomain consists of three stem-loop elements that form a four-way junction involved in binding with the 40S ribosomal subunit. This is needed for eIF3 factor recruitment [51]. The subdomains IIId and IIIe form the main anchoring sites for the 40S ribosomal subunit [51-53]. Because of their importance in IRES-dependent translation, they have been the subject of much research and are considered excellent candidates for HCV targeting. NMR studies have revealed the three dimensional structure of subdomains IIId and IIIe, clarifying their critical role in IRES conformation and the maintenance of its biological function [53, 54]. Subdomain IIId adopts a stem-loop structure interrupted by an internal E-loop closed by a highly conserved apical loop which folds into a U-turn structure. This exposes the nucleotides and favors their interaction with other viral sequences and proteins. The subdomain IIIe contains a highly conserved GAUA tetraloop associated with an efficient viral polyprotein synthesis [53]. An important feature of domain III is the formation of a conserved pseudoknot structure involving subdomain IIIf and a pyrimidine-rich region (positions 325-330; [55]). This pseudoknot has been proposed to interact directly with the 40S ribosomal subunit [50, 51]. It is also present in related viruses and has been associated with the regulation of translation initiation [55, 56].

- *Domain IV* of the 5'UTR region contains the AUG initiation codon at nucleotide 342 that is included in an apical loop enclosing a helical motif, the stability of which is inversely related to IRES translational efficiency [57].

As mentioned above, the maintenance of the IRES secondary and tertiary structure is critical in IRES-dependent translation. The different domains may behave as initiation factors able to capture proteins necessary for the synthesis of the viral precursor polyprotein. The 5'UTR region recruits eIF3 and the 40S ribosomal subunit to form a ribonucleoprotein complex that unleashes the formation of active 80S ribosomal systems, and therefore promotes polyprotein translation [50, 58].

The essential role of this region in the viral cycle and its high conservation rate make it a very attractive therapeutic target.

THE 3' UNTRANSLATED REGION

This region is a 200-240 nt-long sequence located at the 3' end of the HCV RNA genome. Its secondary structure is maintained across different viral isolates (Fig. (**3**); [59]). Mutational analysis has been used to decipher the biological role of the 3' UTR [60]. Enzymatic and chemical analyses

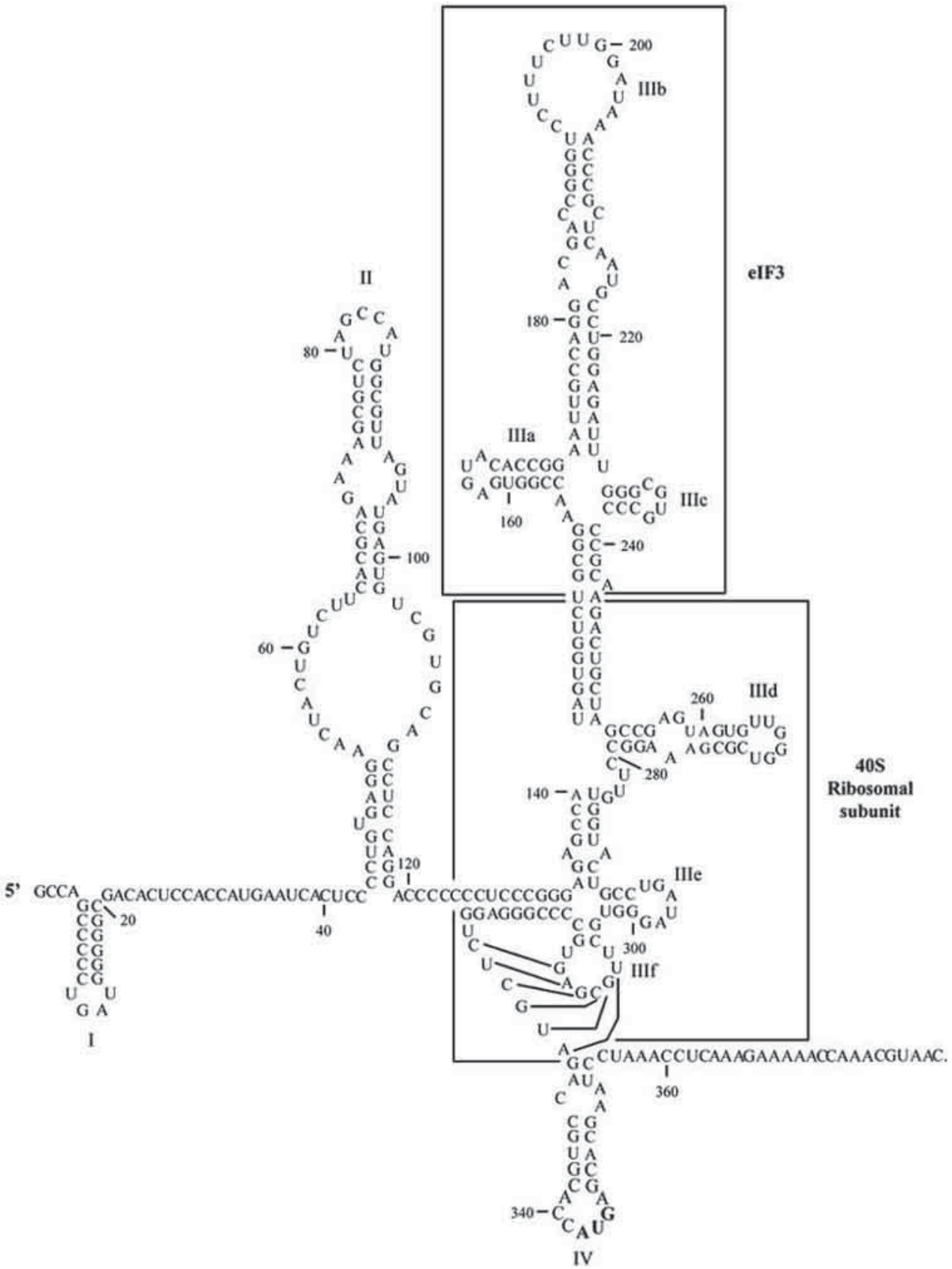

Fig. (2). Proposed secondary structure of the 5'UTR domain of HCV. Domains involved in the interaction with eIF3 factor and ribosomal subunit 40S are marked in boxes. The translation start codon is shown in bold.

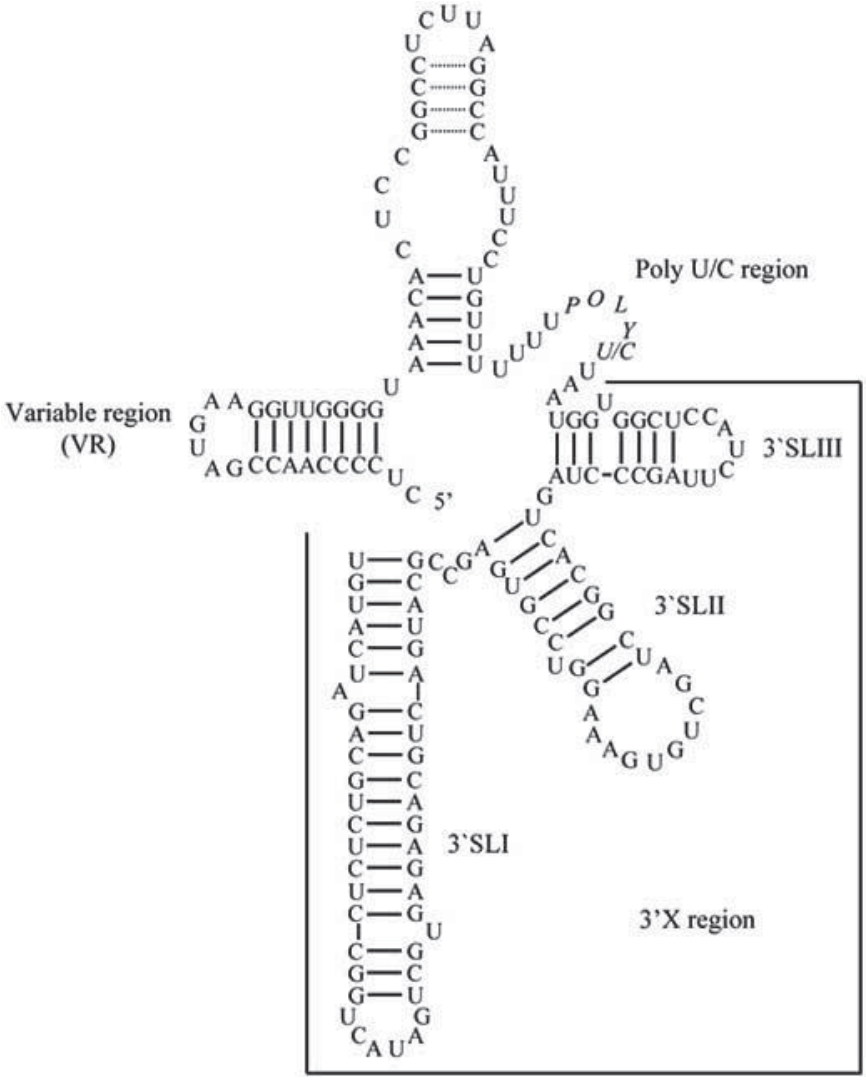

Fig. (3). Structural model proposed for the HCV 3'UTR segment. The different domains of the entire secondary structure are indicated.

have identified three different domains: the VR region, the polyU/C tract and 3'X region [61, 62].

The 3'UTR begins with a short and highly variable sequence, although it also contains a conserved motif and UG dinucleotide [63]. This domain is named VR (variable region) and it has been shown its implication in viral RNA synthesis using subgenomic replicon systems [21].

A polyU/C-rich tract of variable length and composition follows the VR sequence. This can interact with host proteins and regulate viral translation [64, 65]. Some of these host proteins have been identified, e.g., polypyrimidine tract-binding protein (PTB), heterogeneous nuclear ribonucleoprotein C (hnRNP C) and glyceraldehyde-3-phosphate dehydrogenase (GADPH), among others [66-68].

Finally, the 3'UTR region contains a 98 nt-long domain known as the 3'X region – one of the most conserved motifs within this region. It folds into three different stem-loop motifs – SL1, SL2 and SL3 [59] – and is involved in viral RNA replication [21, 69-71] and virion infectivity *in vivo* [60]. Its interaction with human ribosomal complexes has also been reported, suggesting a potential role in the modulation of IRES-dependent translation [72]. It is important to note that this domain is exclusive to HCV (reviewed in [73]) and is an excellent candidate target for new antiviral drugs.

THE NS3 PROTEIN

One of the most extensively studied and best understood targets in anti-HCV therapy is the NS3 protein - a multifunctional molecule composed of two domains with serine protease and helicase activity [74, 75].

Protease activity is one of the preferred targets for the designers of new drugs. Protease has a key role in HCV infection since it is responsible for polyprotein processing during virus maturation. The host proteases first cleave the precursor to release the structural proteins. Secondly, virally-encoded NS2/NS3 proteases yield the mature NS3 protein [76] which initiates the cleavage of the remaining target sites in an Mg^{2+}-dependent reaction [77]: first, cotranslationally at the NS3/NS4A junction, and then in the order NS5A-5B, NS4A-4B and NS4B-5A [78]. The *cis* cleavage is governed

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.