1 Introduction
General Introduction: The Hepatitis C Virus

Global burden of hepatitis C virus infection

HCV Natural history and clinical outcome

Hepatitis C virus (HCV) causes acute and chronic hepatitis, which can eventually lead to permanent liver damage and hepatocellular carcinoma (HCC). Patients with acute HCV infection are usually asymptomatic and spontaneous clearance occurs in 15-45% of cases over a period of 6 months. Approximately 55-85% develop a chronic infection and, 10-20% of these patients will develop liver cirrhosis and are at risk for HCC (1). HCV is transmitted via the blood-borne route. Transfusion of unscreened blood or blood products and unsafe medical procedures are the principal risk factors involved in the transmission in developing countries and some developed countries. Other risks of infection include tattooing, piercing and sharing needles between individuals who use injectable drugs. Sexual transmission has been considered low risk; however, unsafe and/or violent sex can increase the risk of exposure (2). HCV infection is rarely diagnosed during the acute phase, since most of the patients are asymptomatic. Therefore, the diagnosis of acute HCV is challenging and based on the detection of the HCV RNA (viral genome), which appears in circulation 1-2 weeks after the primary infection. The persistence of the HCV RNA, presenting with a plateau or fluctuating viremia and detection of antibodies against HCV (anti-HCV) in the serum for more than six months, is defined as a chronic infection (3, 4). Liver fibrosis is the consequence of chronic infection and inflammation leading to disruption of hepatic architecture and impairment of liver microcirculation and cellular functions (5). Chronic hepatitis C is the most common cause of cirrhosis and occurs in 5-25% of patients with chronic HCV infection over a period of 25-30 years (6). Some environmental and host factors can increase the risk and/or accelerate the natural course of HCV-related disease. These factors include: daily alcohol consumption, infection at an older age (> 40 years), male gender, the level of inflammation, comorbidities such as immunosuppression or metabolic conditions like non-alcoholic steatohepatitis, obesity and insulin resistance (7). Finally, it has recently been demonstrated that HCV may also replicate in the liver in the absence of detectable virus in the blood, a condition referred to as “occult hepatitis C”, with lower potential for progressive disease (8).
Several mechanisms have been implicated in viral clearance and persistence. However, no consensus is defined about the parameters that can accurately predict spontaneous HCV resolution (5). A higher genetic diversity of the infecting virus is correlated with an inefficient immune response to control viral replication, resulting in chronic infection (9). Likewise, a lower viral load may predict a higher rate of clearance (10) and the co-infection with hepatitis B virus (HBV) or human immunodeficiency virus (HIV) also favors HCV viral resolution due to viral interference (11,12). Host factors related to spontaneous resolution are an efficient and robust CD4^+ and CD8^+ T cell response during the acute phase of infection (13), polymorphisms of the interleukin 28B gene and female gender (5).

**HCV epidemiology**

The prevalence of HCV infection has been estimated from population-based studies on the seroprevalence of antibodies to HCV (anti-HCV) reported in the scientific literature. A systematic review in 2013 found that, between 1990 and 2005, the prevalence and number of people with anti-HCV antibodies increased from 2.3% (95% uncertainty interval [UI]: 2.1%-2.5%) to 2.8% (95% UI: 2.6%-3.1%), corresponding to approximately 185 million infected individuals in 2005. The review was based on the meta-analysis of 232 papers reporting on HCV seroprevalence (14). However, a more recent systematic review based on 4,901 studies from 87 countries and some unpublished reports, projected a lower HCV prevalence of 110 million anti-HCV positive individuals (95% UI: 92-149 million) and a chronic HCV prevalence of 80 million individuals (95% UI: 64-103 million). The latter estimate is more in line with the updated prevalence reported by the World Health Organization (WHO) (15,16). The distribution of HCV infection is highly variable among individual countries ranging from <1% to >10% (Figure 1) (17). The highest prevalence has been reported in Africa, especially in Egypt and Cameroon (>10%), followed by the Middle East (18,19). The Americas, Australia, Northern and Western Europe are considered areas with low prevalence (20). In absolute numbers, the countries with the highest number of HCV-infected individuals are China with approximately 30 million infected individuals, followed by India (18 million), Egypt (11 million) and Pakistan and Indonesia (approximately 9.5 million each) (17).
Figure 1. HCV Epidemiology. The estimated prevalence of HCV infection and the global distribution of HCV genotypes are presented. Genotype 7, which was only recently identified in patients from Central Africa, is not included. Adapted with permission from Macmillan Publishers Ltd: Nature Reviews Gastroenterology and Hepatology form (20). License number 4231501427667.

In developed countries in North America (21–23), Northern and Western Europe (24), Australia and Japan, the HCV prevalence is low (<2%) (20). The patterns of HCV prevalence in developing countries are highly heterogeneous. The highest prevalence of HCV has been reported in Egypt (15% anti-HCV positivity in adults), followed by Pakistan and Iran (10 million infected individuals) (19,25,26). In the Caribbean area and Latin America, prevalence ranges from 0.5% to 2.3% (23,27).

Virus structure and replication cycle

HCV genome organization

Hepatitis C virus (HCV) is classified by the International Committee of Virus Taxonomy (ICVT) into the Flaviviridae family, genus Hepacivirus. The HCV genome is a positive-sense, single-stranded RNA molecule (+ssRNA) of approximately 9.6 kilobases (kb). The genome contains a single open reading frame (ORF) that encodes a polyprotein of 3,008 to 3,100 amino acids (aa), depending on the virus genotype (28). The polyprotein is processed co- and post-translationally by viral and cellular proteases into four structural proteins and six non-structural proteins (Figure 2). The structural proteins are located at the N-terminal end of the polyprotein and include the capsid protein Core, the glycoproteins E1 and E2 and the viroporin P7. The non-structural proteins
NS2, NS3, NS4A, NS4B, NS5A and NS5B, all with diverse biochemical functions described below, are located at the C-terminal end (Figure 2) (29). An additional protein has been identified as the alternative reading frame protein (ARFP), which is synthesized by an ORF overlapping at the coding sequence of Core at nucleotide +1 (30).

The ORF is flanked at the 5’- and 3’-ends by two highly conserved untranslated regions (UTRs) that regulate virus replication. The 5’-UTR region corresponds to 341 nucleotides (nt) located upstream of the start codon and is composed of 4 domains numbered I to IV (32). It forms a secondary structure known as the internal ribosomal entry site (IRES) that is essential for translation of the viral genome via a cap-independent mechanism, contrary to the mechanism that is normally used for translation of messenger RNAs from the cell (33). The 3’-UTR region corresponds to a sequence of 230 nt and has a tripartite structure consisting of a short and variable region of 40 nt, a poly-uracil sequence and a 98 nt conserved element that is essential for viral replication (Figure 2) (34).

**HCV structural proteins**

The cleavage of the polyprotein between the Core and E1 sequence by a signal peptidase yields an immature 191 aa long Core protein. Further C-terminal processing by the intramembrane cleaving protease SPP (Signal Peptide Peptidase) yields the mature 21 kDa Core protein of 177 aa (35,36). Mature HCV Core is a homodimeric membrane protein stabilized through disulfide bond formation at the Cysteine-128 residue and is responsible for capsid formation (37). Three domains have been identified in the HCV core protein (191aa long), based on predicted structural and functional characteristics. Domain I (aa1-aa120), corresponding to the N-terminal region, is a highly basic domain that is involved in the recruitment of the viral RNA during formation of new virions and homo-dimerization and therefore important in nucleocapsid assembly. Domain II, located between aa 120 and aa 177, is a hydrophobic region predicted to form one or two β-helices that are involved in the association of HCV Core with the endoplasmic reticulum (ER) membrane and lipid droplets. Domain III, corresponding to the C-terminal (aa177-aa191) of the protein, is a highly hydrophobic region that serves as a signal sequence for the targeting of the E1 protein to the ER (38).
The envelope glycoproteins E1 (160 aa) and E2 (360 aa) are type I transmembrane proteins with an N-terminal ectodomain and a short C-terminal transmembrane domain (TMD) of approximately 30 aa. E1 and E2 play pivotal roles in different steps of the HCV life cycle, including the assembly of viral particles, virus entry and fusion with the endosomal membrane in the host cell (39). P7 is a 63 aa integral membrane polypeptide that forms hexamers or heptamers with cation channel activity and therefore belongs to the viroporin family. P7, comprising two transmembrane α-helices connected by a positively charged cytosolic loop, facilitates virus production during assembly of new virions. P7 is not required for RNA replication in vitro, but is essential for the assembly and release of infectious HCV particles in vitro and in vivo (40) (Figure 2).
**HCV non-structural proteins.**

NS2 is a viral cysteine autoprotease that cleaves the NS2/NS3 junction of the polyprotein. Once processed it is a protein of 217 aa and has a molecular weight of approximately 23 kDa (Figure 2). NS2 possesses a hydrophobic N-terminal subdomain as well as a C-terminal cytoplasmic domain. Its catalytic activity resides in the C-terminal domain between aa 94-217 (41). In addition to its protease activity, NS2 plays a central role in virus assembly due the interaction with E1 and E2 and the non-structural proteins NS3 and NS5A (42,43). NS3 is a 70 kDa multifunctional protein with serine protease activity located in the N-terminal domain (aa 1–180) and a nucleoside-triphosphatase (NTPase)/RNA helicase function in the C-terminal domain (aa 181–631) (Figure 2). Both enzyme activities have been well characterized and require the binding of NS4A (54 aa) that acts as a cofactor of the non-covalent complex NS3/NS4A (39,44). NS4B is a hydrophobic 27 kDa protein of 261 aa. It is an integral membrane protein comprising an N-terminal part (aa 1 to ~69), a central part harboring four predicted transmembrane domains (aa ~70 to ~190), and a C-terminal part (aa ~191 to 261). The N-terminal part contains two amphipathic α-helices (AH), AH1 and AH2, extending from aa 3-35 and 42-66, respectively, and upon oligomerization AH2 can cross the membranes. NS4B induces the formation of a membranous web, a specific membrane alteration consisting of locally-confined membranous vesicles that serves as a scaffold for the HCV replication complex at the ER (39). NS5A is a 447 aa membrane-associated phosphoprotein that plays an important role in regulating HCV RNA replication and particle formation. HCV NS5A can be found in both basally phosphorylated (56 kDa) and hyper-phosphorylated (58 kDa) forms. It possesses 3 domains (D1-D3) involved in different stages of viral replication. HCV NS5A has attracted considerable interest because of its role in modulating the response to interferon-alpha (IFN-α) therapy due to the presence of a region termed “interferon sensitivity determining region” (ISDR) (45). Finally, NS5B (591 aa) is a 68 kDa protein and is the viral RNA-dependent RNA polymerase (RdRp). It is responsible for HCV genome replication via synthesis of a complementary negative-strand RNA using the genome as a template and the subsequent synthesis of genomic positive-strand RNA from this negative-strand RNA template (46) (Figure 2).
HCV replication

HCV is an enveloped virus of approximately 40-70 nanometers (nm) that circulates in the infected host associated with low-density lipoproteins (LDL) and very-low-density lipoproteins (VLDL). Both types of particles appear to represent the infectious fraction. Additionally, HCV can also circulate bound to immunoglobulins and as free virions. The HCV Core protein and the envelope glycoproteins E1 and E2 are the principal protein components of the virion. E1 and E2 are anchored to a host cell-derived double-layer lipid envelope that surrounds the nucleocapsid composed of multiple copies of the Core protein and the genomic RNA. E1 and E2 play an important role in the attachment of the virion to its receptors and co-receptors, which will be described below (47,48).

The first steps of the HCV viral lifecycle are the attachment to and entry into the host cell (Figure 3). HCV enters by clathrin-mediated endocytosis and the hepatocytes are the main target cells. However, the infection of B cells, dendritic cells and other types of immune cells have also been reported (49–51). The tetraspanin protein CD81, which is found at the surface of many cell types, including hepatocytes, the LDL receptor (LDLR) and the scavenger receptor class B type I (SR-BI) have been proposed to act as HCV receptors (Figure 3) (52–54). Claudin-1 (CLDN1) was identified as HCV co-receptor and was found to be essential for HCV entry into hepatocytes (55). Both, CD81 and SR-BI bind to E2 and are necessary, but not sufficient for HCV entry. In addition, HCV E2 also can bind to DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) and L-SIGN (Liver/Lymph node-specific intercellular adhesion molecule-3-grabbing integrin). L-SIGN is a calcium-dependent lectin expressed on liver sinusoidal endothelial cells that may facilitate the infection process by trapping the virus for subsequent interaction with hepatocytes (29).

The translation of the HCV viral genome occurs with the formation of a complex of the IRES and the 40S ribosomal subunit. This is followed by the assembly of a 48S complex at the AUG initiation codon after the association of eukaryotic translation initiation factor 3 (eIF3) and the ternary complex (eIF2•Met-tRNA•GTP) (56). As described earlier, translation of the HCV ORF produces a polyprotein that is co- and post-translationally processed at the ER (Figure 3).
The replication of the HCV genome occurs in a membrane-associated replication complex, composed of viral proteins NS3, NS4A, NS4B, NS5A and NS5B, the replicating RNA and cellular membranes derived from the ER, Golgi apparatus, mitochondria and lysosomes (Figure 3). The membranes are modified in specific ways and serve as physical support to organize the RNA replication complex and to protect the viral RNA from double-strand RNA-mediated host defenses or RNA interference (29).

**Figure 3. HCV life cycle**: HCV lipoviroparticles attach and enter target hepatocytes via interaction with CD81 and SR-B1 and subsequent receptor-mediated endocytosis (Step 1 and 2). Released viral RNA is translated at the ER producing a single polyprotein precursor that is cleaved by host and viral proteases (Steps 3 and 4). The RNA is replicated by the viral RdR-polymerase (NS5B) via a negative-strand intermediate at the membranous web (Step 5). Newly synthesized positive-strand RNA is encapsidated by the viral nucleocapsid core in close proximity to lipid droplets and envelope glycoproteins are acquired through budding into the ER lumen (Step 6). Lipoviroparticles mature in the ER through interactions with lipoproteins (Step 7) and exit the cell by via the cellular Golgi apparatus (Step 8).

The final steps of HCV replication are the packaging, assembly and particle release (Figure 3). HCV particle assembly requires the spatial and temporal synchronization of the structural proteins and the replication complexes to facilitate the budding of an enveloped nucleocapsid. An environment rich in lipid droplets (LDs) is considered
essential for HCV assembly, hence the association of HCV with non-alcoholic steatohepatitis (NASH). LDs are intracellular lipid deposits of cholesterol esters and triacylglycerides and inhibition of the synthesis of these lipids can block HCV assembly (57). The HCV Core protein attaches to LDs through the D2 domain resulting in the accumulation of Core around LDs. The nascent particle matures further in a post-ER step, yielding the characteristic low-density infectious particle. The exit of the HCV particles occurs through the secretory pathway and depends on classical host factors of the secretory pathway (58) (Figure 3).

**HCV genetic diversity and distribution of genotypes**

HCV has a high genetic diversity resulting from the high rate of replication (estimated to generate $10^{12}$ new viral particles per day) and the absence of proofreading activity of the viral RNA polymerase (7). After the publication of the first complete genome sequence of HCV (47,59), it became clear that HCV isolates from different individuals showed substantial genetic diversity. This variation was subsequently organized as genotypes, subtypes and quasispecies. The complete coding region sequences available at the National Center for Biotechnology Information (NCBI) genome data base and the Los Alamos HCV data base reveal seven major phylogenetic groups corresponding to genotypes 1 to 7 (Table 1), that vary over 30% in nucleotide sequence (60). These genotypes are subdivided into 67 subtypes, indicated by a letter following the genotype number (Figure 4 and Table 1). The qualification as a subtype requires a complete or nearly complete coding region sequence difference of at least 15% from other sequences. Additionally, 20 provisionally assigned subtypes, and 21 unassigned subtypes have been reported at the web site of the ICTV (Figure 1 and 4) (61).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Subtypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1a, 1b, 1c, 1e, 1g, 1h and 1l</td>
</tr>
<tr>
<td>2</td>
<td>2a, 2b, 2c, 2d, 2e, 2i, 2j, 2k, 2m, 2q and 2r</td>
</tr>
<tr>
<td>3</td>
<td>3a, 3b, 3g, 3h, 3l, and 3k</td>
</tr>
<tr>
<td>4</td>
<td>4a, 4b, 4c, 4d, 4f, 4g, 4k, 4l, 4m, 4n, 4o, 4p, 4q, 4r, 4t, 4v, 4w</td>
</tr>
<tr>
<td>5</td>
<td>5a</td>
</tr>
<tr>
<td>6</td>
<td>6a to 6w and the subtype 6xa</td>
</tr>
<tr>
<td>7</td>
<td>7a</td>
</tr>
</tbody>
</table>
The seven HCV genotypes represent a diverse global distribution, which reflects differences in the epidemiology, transmission modes, ethnic groups and social-economic levels of the region (Figure 1). The genotypes 1, 2 and 3 are the most common across the world, whereas genotypes 4 to 7 are generally confined to specific geographical regions (20). Genotype 1 has the broadest geographical distribution and has been identified in North America (21,62), Northern and Western Europe (63), South America (20), Asia and Australia (64). Genotype 2 is found in Eastern, Southern and Northern Europe, South America and Asia (20). However, studies from Poland, Estonia and Greece have reported an increase in genotype 3 (subtype 3a) and decrease in genotype 2 over time (65–67). In addition, genotype 3 is predominant in Pakistan, South Asia and Australia (20). Genotype 4 is found predominantly in Africa and the Middle East (68), although the subtypes 4c and 4d have also been reported in Spain (69); genotype 5 is almost exclusively found in South Africa; genotype 6 is endemic in Southeast Asia and highly prevalent in Hong Kong and Southern China (20) and genotype 7 subtype 7a was recently identified in patients from Central Africa (70).

Hepatitis C virus: new treatment options and future perspectives

Until recently, HCV therapy was based on interferon type I and ribavirin requiring up to 48 weeks of co-therapy. However, interferon is poorly tolerated due its side effects and its low efficacy: the sustained virological response (SVR) is often less than 50%. Subsequently, the introduction of two NS3/4A protease inhibitors used in combination with interferon marked the start of the era of Direct-Acting Antivirals (DAA). Since then, additional therapies have become available comprising interferon-free DAA regimes curing more than 90% of infected patients. The non-structural proteins are the targets for currently approved DAAs, including NS3/4A protease inhibitors (PI), NS5A inhibitors and NS5B nucleoside (NA) and non-nucleoside (NNA) analogues (71). At present, six interferon-free DAA regimes are approved for HCV treatment, including combinations of DAAs in fixed-dose pills (Table 2). All of these treatments require less than 24 months of treatment depending on the clinical status of the patient and the HCV genotype (72).
Figure 4. HCV Genetic variability: Phylogenetic tree of 129 representative complete coding region sequences. Up to two representatives of each confirmed genotype/subtype were aligned and a neighbor joining tree constructed using maximum composite likelihood nucleotide distances between coding regions using MEGAS. Sequences were chosen to illustrate the maximum diversity within a subtype. Tips are labeled by accession number and subtype (*; unassigned subtype). For genotypes 1, 2, 3, 4, and 6, the lowest common branch shared by all subtypes and supported by 100% of bootstrap replicates (n = 1000) is indicated by ·. Adapted with permission from Elsevier Publishers Ltd from (61).

Table 2. Direct-Acting Antivirals (DAA) for HCV treatment.

<table>
<thead>
<tr>
<th>Regimen</th>
<th>DAA treatment/combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Daclatasvir + Asunaprevir</td>
</tr>
<tr>
<td>2</td>
<td>Daclatasvir + Sofosbuvir ± Ribavirin</td>
</tr>
<tr>
<td>3</td>
<td>Ledipasvir/Sofosbuvir</td>
</tr>
<tr>
<td>4</td>
<td>Paritaprevir/Ritonavir/Ombitasvir + Dasabuvir ± Ribavirin</td>
</tr>
<tr>
<td>5</td>
<td>Simeprevir + Sofosbuvir</td>
</tr>
<tr>
<td>6</td>
<td>Sofosbuvir + Ribavirin</td>
</tr>
</tbody>
</table>
In contrast to DAAs that target viral proteins, host-targeting agents (HTAs) have been developed and studied to interfere with cellular factors involved in HCV viral life cycle. By acting through a complementary mechanism of action and by exhibiting a generally higher barrier to resistance, HTAs offer a promising option to prevent and treat viral resistance. Indeed, given their complementary mechanism of action, HTAs and DAAs can act synergistically to reduce viral loads. Some of the HTAs act as 1) virus entry inhibitors to prevent initiation of viral infection and viral dissemination, e.g. monoclonal antibodies that target cell entry receptors as CD81, SR-BI and co-receptor CLDN1 (73,74); 2) translation inhibitors to prevent subsequent viral replication following viral endocytosis and fusion. Because HCV RNA is translated in an IRES-mediated manner, microRNA-122 (miR-122) plays an important role in HCV translation. miR-122, one of the most abundant liver-expressed miRNAs, binds to the HCV genome and enhances viral translation and replication. Sequestering miR-122 using Miravirsen, a locked nucleic acid-modified oligonucleotide complementary to the 5`-end of miR-122, showed prolonged and dose-dependent reduction in HCV viremia in patients without evidence of long-term safety issues (75); 3) assembly inhibitors, e.g. Celgosivir and 4) biological response modifiers, like agonists of the Toll-like receptors (TLR) 7 and TLR9 (76).

Despite the advances in treatment options, the ability of HCV to develop resistance to antiviral drugs is quite high due to its high replication and mutation rate and lack of proof-reading activity. Therefore, the genetic diversity of HCV is currently the biggest challenge for the development and implementation of successful DAA and HTAs regimes. The best example in this regard is the existence of resistance-associated variants (RAVs) of HCV that correspond to viral sequences with preexisting polymorphisms that can reduce the efficacy of the DAAs and HTAs. RAVs can emerge from the viral population as the dominant species during treatment (72).

DAAs do not offer protective immunity, which may limit the use of DAA therapy as a prevention strategy. Therefore, not only HCV treatment is important, but also the development of a vaccine to control and avoid new infections is required. Vaccine development for HCV has been challenging because of the high sequence variability within the protein coding regions, the evolution of quasispecies that can exhaust the immune response and the mechanisms used by HCV for the evasion of the immune system. Several HCV components have been used as a target for vaccine development.
and to produce a neutralizing antibody response, like the glycoproteins E1 and E2, and the Core-E1-E2 DNA sequence (77). However, the efficacy is low and this area is still in development.

Major obstacles still exist for the successful elimination of HCV infection: 1) improving public health surveillance, 2) increasing awareness in the infected population 3) provide the infected people with proper health care, and 4) improving access to effective treatments. In USA, the National Academy of Sciences released a plan for eliminating HBV and HCV as a public health problem. Likewise, the WHO launched a similar plan to eradicate HBV and HCV before the year 2030.

**Conclusions**

Despite the huge progress in our understanding of HCV pathogenesis in the recent years, some aspects still need more attention. This is particularly true for the host-virus interaction in HCV infection and the adaptation of host cells to HCV infection. In the next chapter (Chapter 2) we will review the current literature on the interaction of the virus with the host cell (hepatocyte).
References


33. Beales LP, Rowlands DJ, Holzenburg A. The internal ribosome entry site (IRES) of hepatitis C virus visualized by electron microscopy. RNA. 2001;7(5):661–70.


Introduction

Protein 2 of the Hepatitis C Virus Reveal Its Key Role as Organizer of Virion Assembly. PLoS Pathog. 2010;6(12).


