Published in final edited form as:

Curr Opin Virol. 2015 June ; 12: 53–58. doi:10.1016/j.coviro.2015.02.004.

HCV Glycoprotein Structures: What to Expect from the Unexpected

Abdul Ghafoor Khan^a, Matthew T. Miller^a, and Joseph Marcotrigiano^a

^aCenter for Advanced Biotechnology and Medicine, Department of Chemistry and Chemical Biology, Rutgers University, 679 Hoes Lane West, Piscataway, New Jersey 08854

Abstract

Hepatitis C virus (HCV) is continuing to spread worldwide, adding 3 million new infections each year. Currently approved therapies are highly effective; however, access to them is limited due to the high cost of treatment. Therefore, a cost effective vaccine and alternative antivirals remain essential. HCV envelope glycoproteins, E1 and E2, heterodimerize on the virion surface and are the major determinant for virus pathogenicity and host immune response. Recent structural insights into amino-terminal domain of E1 and core of E2 have revealed unexpected folds not present in glycoproteins from related viruses. Here we discuss these structural findings with respect to their role in HCV entry and impact on potential vaccine design and new antivirals.

Introduction

Since its initial discovery in the late 1970s, hepatitis C virus (HCV) has been identified in all parts of the world, with seven major genotypes and more than 50 subtypes isolated. Currently, 3% of the human population is infected, making HCV a serious global health problem [1,2]. There is no vaccine, and it is estimated that an additional 3 to 4 million individuals become infected each year [3•]. Although the US Food and Drug Administration (FDA) recently approved several direct acting antivirals (DAA), including Telaprevir (VERTEX), Boceprevir (Merck), Harvoni (Gilead Sciences) and Viekira Pak (AbbVie), access to these medications is limited due to their high cost (over \$80,000 per course of treatment). Therefore, it is unlikely that treatment alone will halt the spread of the virus without an effective vaccine.

HCV is an enveloped virus with a single-stranded, positive sense RNA genome (Figure 1a). The virion particle carries two surface proteins, E1 and E2, which exist as a heterodimer. A unique feature of the HCV particle is its association with lipoproteins and lipids, resulting in an unusually low buoyant density [4]. E1 and E2 are heavily glycosylated, which is critical for proper folding, transport through the secretory pathway and escape from the host immune response. The process of viral entry is thought to involve a physical interaction between the E1/E2 heterodimer and host cell surface receptors. Many cellular receptors have

Corresponding Author: Joseph Marcotrigiano, Center for Advanced Biotechnology and Medicine, Department of Chemistry and Chemical Biology, Rutgers University, 679 Hoes Lane West, Piscataway, NJ 08854; jmarco@cabm.rutgers.edu; telephone: (732) 235-5309.

been implicated either directly or indirectly in HCV entry [5]. Convincing evidence suggests that glycosaminoglycans and low-density lipoprotein receptor are required for the initial attachment of the virus to host cells [6]. In addition, four receptors have been identified to function in entry, including scavenger receptor class B type 1 (SR-BI) [7], CD81 [8], Claudin-1 [9] and Occludin [10].

E1 and E2 are type I transmembrane proteins (Figure 1b&c). The ectodomains of E1 and E2 have been previously defined as the minimal deletions that result in secretion of properly folded protein [11]. Both E1 and E2 are heavily modified post translation with numerous N-linked glycans and intramolecular disulfide bonds [11–13]. The folding of these proteins requires ER chaperones, particularly calnexin [13]. For these reasons, overexpression of these proteins often results in misfolded, disulfide-linked aggregates, which has hindered biophysical and structural characterization.

A significant breakthrough in understanding the three dimensional organization of HCV glycoproteins is provided by recent crystal structures of the core ectodomain of E2 and amino-terminal domain of E1 [14–16••]. The structures of E1 and E2 reveal unexpected novel features and lacks the hallmarks of viral membrane fusion proteins, suggesting there may be a new entry mechanism for HCV. In this review we discuss the importance of these structures and their implication on HCV vaccine design.

E2 structure

Recently, two independent structures have provided the first structural insights into the core domain of E2 (PDB ID 4MWF and 4WEB) (Figure 2) [14••,15••]. Both groups obtained crystals by forming an E2-antibody fragment (Fab) complex and making deletions within E2 (Figure 1c). In 4MWF, a neutralizing, human antibody (AR3C) that prevents E2-CD81 interaction was used in conjunction with an E2 ectodomain, which did not contain hypervariable region 1 (HVR-1) and replaced HVR-2 with a flexible linker. In 4WEB, a non-neutralizing mouse monoclonal antibody (2A12) was used for co-crystallization with an E2 ectodomain lacking first 72 residues, which was shown to be disordered. This region includes conserved sequences implicated in binding to the cellular receptors (SR-BI and CD81) as well as several epitopes for neutralizing antibodies [17–21]. AR3C stabilizes the amino-terminal portion of the E2 ectodomain, enabling modeling of the CD81 binding site. In contrast, 2A12 recognizes a linear epitope at the carboxyl-terminus of the ectodomain, and does not interfere with E2 binding to CD81 or SR-BI.

The E2 core is a monomer with a compact, globular architecture consisting of immunoglobulin (IgG)-like fold and an additional novel domain. The novel domain comprises a central β sheet surrounded by loops, short helices and two β strands. In the core domain of HCV E2, many regions lack regular secondary structure and most of the N-linked glycans are largely disordered. Overall fold is formed by 4 β strands that are arranged in an inner buried and a solvent exposed, outer sheet. A loop connects the inner and outer sheets, and carries many key CD81 binding residues and is located adjacent to the front layer where additional CD81 contact residues are found. Electron microscopy [15••] and small angle X-ray scattering (SAXS) [14••] show that the full-length E2 ectodomain adopts a similar

compact globular fold. Furthermore, size-exclusion chromatography and SAXS data suggest that the E2 ectodomain does not undergo major structural or oligomeric changes upon exposure to low pH [14••].

The E2 core structures are largely similar with a root mean square deviation (RMSD) of less than 0.8Å for similar α -carbon positions; however, there are some interesting divergences, especially in the disulfide-bonding pattern (Figure 2). Despite the high sequence variability in E2, the cysteine residues are absolutely conserved across genotypes, suggesting an important structural and/or functional role. Both structures have three disulfide bonds in common formed between C5–C9, C7–C8 and C14–C16 that appear to be critical for the fold of the core domain as they link secondary structure elements. However, in 4MWF, C1 is bonded to C6, while C6 is free in 4WEB since C1 is absent. Notably, all of the disulfide bond discrepancies are in the regions where either the protein was modified or in flexible loops with poor electron density (mostly around HVR-3). The reasons for these discrepancies are currently unclear. The E2 protein sequences in the two structural studies were from different HCV genotypes; produced in HEK293T cells, which are known to yield functional envelope glycoproteins in form of HCVpp; and have various deletions. Both production methods yielded functional E2 as the full-length ectodomains can inhibit HCVpp or HCVcc infection [14••,15••]. Moreover, it is possible that E2 has different disulfide bonding patterns on the surface of the virion. E2 has been detected as non-covalently associated to E1 in the ER, whereas E1/E2 appear to be covalently bound on secreted virus particles [22–24].

E1 Structure

E1 forms a heterodimer with E2 on the virus surface; however, the exact functions of E1 remain elusive. Recently, the structure of the amino-terminal portion (residues 1–79) of the E1 (nE1) was determined by X-ray crystallography (Figures 1b and 3) [16••]. Like the E2 core structure, nE1 also shows an unexpected, novel architecture. Surprisingly, the protein is arranged as a disulfide linked, domain swapped homodimer. nE1 is composed primarily of β strands with a single long α -helix sandwiched by two and three antiparallel β strands. The first two β strands make a hairpin and are involved in domain swapped, homodimer formation. However, the possibility of nE1 adopting the monomeric conformation cannot be excluded. Interestingly, the fold has closest structural similarity with steroidogenic acute regulatory protein-related transfer domain, which binds to hydrophobic ligands such as sterol and lipid like molecules. This finding is particularly interesting as the HCV particle is associated with apolipoproteins. The most striking observation is that there are six monomers in the nE1 asymmetric unit, stabilized by a series of intra- and intermolecular disulfide bonds (Figure 3). As nE1 crystals were obtained at low pH, it is tempting to speculate that the current structure may represent a post-fusion form. Unfortunately, only a few residues of the putative fusion peptide in E1 are included in the construct making it impossible to authenticate the prediction [25•].

Membrane fusion

As a member of the *Flaviviridae* family, HCV was postulated to contain a class II fusion protein similar to other family members [26•]. All class II fusion proteins contain an elongated, three-domain architecture composed predominantly of β strands, and exist as homo or heterodimers with a hydrophobic, membrane fusion peptide buried at the dimer interface at neutral pH. Upon receptor binding and/or exposure to low pH, these proteins undergo self-rearrangement into stable trimers exposing the fusion peptide, resulting in viral and host membrane fusion. Although E2 was hypothesized to be the fusion protein, lack of an authentic fusion peptide and sequence similarity to class II fusion proteins raised doubts about this classification. Additionally, the E2 from the closely related pestivirus, bovine viral diarrhea virus (BVDV), lacks a membrane fusion peptide and class II fold despite possessing a multidomain elongated homodimer organization (Figure 4) [27,28].

The recent structural data for HCV E1 and E2 does not support the prediction that they are class II fusion proteins (Figure 4). E2 core and nE1 are only 32Å and 59Å in length, respectively, while envelope glycoproteins from other family members exist as extended, elongated dimers of over 140Å in length. Furthermore, the predicted putative fusion peptide in the E2 core is a part of the secondary structural elements within the hydrophobic core, ruling out membrane fusion activity. The E2 core also does not undergo major structural or oligomeric rearrangements upon exposure to low pH [14••]. These findings strongly suggest that E2 is not a fusion protein, however the fusion peptide may reside in the deleted carboxyl-terminal region, which are absent in these structures. Lastly, although nE1 exists as a dimer in solution, it lacks any structural resemblance to other viral fusion proteins. Therefore, the HCV membrane fusion might be a completely different mechanism that could be brought about by either: 1) E1 alone is sufficient and harbors the required fusion peptide outside of nE1; 2) E1/E2 heterodimer is responsible for fusion; 3) neither E1 nor E2 undergoes structural changes; however, a putative hydrophobic, fusion peptide exists in the carboxyl-terminus of E1 and/or E2 which is buried within heterodimer interface; or 4) lastly other factors are involved in HCV fusion process, such as the virion associated lipoproteins or the various cellular receptors and co-factors.

Impact of HCV glycoprotein structures and future prospects

E1 and E2 are primary determinants of viral pathogenicity and host immune response, emphasizing the potential value of the envelope glycoproteins in generating a vaccine. Indeed, prime-boost regimen with E1/E2 can elicit cross-reactive, neutralizing antibody responses in the chimpanzee model, indicating the existence of conserved immunogenic epitopes across genotypes [29•]. More recently, human volunteers challenged with E1/E2 recombinant proteins resulted in cross-neutralizing antibodies against all major genotypes providing a basis for the development of a vaccine, using the recombinant glycoproteins [30•].

The HCV glycoprotein structures are important breakthroughs for the development of a vaccine. However, the fact that majority of the E2 is composed of loops and flexible regions raises the concerns about the use of E2 alone for vaccine or drug design. Specifically, the

region between HVR-1 and HVR-2, which contains part of the AR3C epitope, represents an important neutralizing immunodominant face of the protein but appears to have high conformational flexibility. Nevertheless, structural information for the broadly neutralizing AR3C antibody is extremely useful. Moreover, since the E2 core maintains the native fold after the removal of the HVR-1 and -2, it is possible that future vaccine candidates could eliminate the inclusion of variable regions yet maintain the overall fold. In conclusion, current structural information provides the foundation to develop a cross protective vaccine and entry inhibitors.

Acknowledgments

This study was supported by the NIH through R01 (AI080659 and GM111959) and P50 (GM103368) to J.M.

References

1. Perz, JF.; Farrington, LA.; Pecoraro, C.; Hutin, YJF.; Armstrong, GL. 42nd Annual Meeting of the Infectious Disease Society of America; Boston, MA USA. 2004;
2. Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis.* 2005; 5:558–567. [PubMed: 16122679]
- 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. This reference describes the current epidemiology and provides valuable information regarding new infections. [PubMed: 23172780]
4. Catanese MT, Uryu K, Kopp M, Edwards TJ, Andrus L, Rice WJ, Silvestry M, Kuhn RJ, Rice CM. Ultrastructural analysis of hepatitis C virus particles. *Proc Natl Acad Sci U S A.* 2013; 110:9505–9510. [PubMed: 23690609]
5. Lindenbach BD, Rice CM. The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol.* 2013; 11:688–700. [PubMed: 24018384]
6. Helle F, Dubuisson J. Hepatitis C virus entry into host cells. *Cell Mol Life Sci.* 2008; 65:100–112. [PubMed: 17914604]
7. Scarselli E, Ansuini H, Cerino R, Roccasecca RM, Acali S, Filocamo G, Traboni C, Nicosia A, Cortese R, Vitelli A. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *Embo J.* 2002; 21:5017–5025. [PubMed: 12356718]
8. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R, Weiner AJ, Houghton M, Rosa D, Grandi G, et al. Binding of hepatitis C virus to CD81. *Science.* 1998; 282:938–941. [PubMed: 9794763]
9. Evans MJ, von Hahn T, Tscherne DM, Syder AJ, Panis M, Wölk B, Hatzioannou T, McKeating JA, Bieniasz PD, Rice CM. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature.* 2007; 446:801–805. [PubMed: 17325668]
10. Ploss A, Evans MJ, Gaysinskaya VA, Panis M, You H, de Jong YP, Rice CM. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature.* 2009
11. Michalak JP, Wychowski C, Choukhi A, Meunier JC, Ung S, Rice CM, Dubuisson J. Characterization of truncated forms of hepatitis C virus glycoproteins. *J Gen Virol.* 1997; 78 (Pt 9):2299–2306. [PubMed: 9292018]
12. Goffard A, Dubuisson J. Glycosylation of hepatitis C virus envelope proteins. *Biochimie.* 2003; 85:295–301. [PubMed: 12770768]
13. Dubuisson J, Rice CM. Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. *J Virol.* 1996; 70:778–786. [PubMed: 8551615]
- 14. Khan AG, Whidby J, Miller MT, Scarborough H, Zatorski AV, Cygan A, Price AA, Yost SA, Bohannon CD, Jacob J, et al. Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. *Nature.* 2014; 509:381–384. This reference provides the crystal structure of HCV E2 core, highlighting the novel fold and an unexpected architecture. [PubMed: 24553139]

- 15. Kong L, Giang E, Nieusma T, Kadam RU, Cogburn KE, Hua Y, Dai X, Stanfield RL, Burton DR, Ward AB, et al. Hepatitis C virus E2 envelope glycoprotein core structure. *Science*. 2013; 342:1090–1094. This reference provides the crystal structure of HCV E2 core, highlighting the novel fold and an unexpected architecture. [PubMed: 24288331]
- 16. El Omari K, Iourin O, Kadlec J, Sutton G, Harlos K, Grimes JM, Stuart DI. Unexpected structure for the N-terminal domain of hepatitis C virus envelope glycoprotein E1. *Nat Commun*. 2014; 5:4874. This reference provides the first N-terminal domain structure of HCV E1 protein. [PubMed: 25224686]
- 17. Wahid A, Dubuisson J. Virus-neutralizing antibodies to hepatitis C virus. *J Viral Hepat*. 2013; 20:369–376. [PubMed: 23647953]
- 18. Keck ZY, Xia J, Wang Y, Wang W, Krey T, Prentoe J, Carlsen T, Li AY, Patel AH, Lemon SM, et al. Human monoclonal antibodies to a novel cluster of conformational epitopes on HCV E2 with resistance to neutralization escape in a genotype 2a isolate. *PLoS Pathog*. 2012; 8:e1002653. [PubMed: 22511875]
- 19. Kong L, Giang E, Nieusma T, Robbins JB, Deller MC, Stanfield RL, Wilson IA, Law M. Structure of hepatitis C virus envelope glycoprotein E2 antigenic site 412 to 423 in complex with antibody AP33. *J Virol*. 2012; 86:13085–13088. [PubMed: 22973046]
- 20. Kong L, Giang E, Robbins JB, Stanfield RL, Burton DR, Wilson IA, Law M. Structural basis of hepatitis C virus neutralization by broadly neutralizing antibody HCV1. *Proc Natl Acad Sci U S A*. 2012; 109:9499–9504. [PubMed: 22623528]
- 21. Deng L, Zhong L, Struble E, Duan H, Ma L, Harman C, Yan H, Virata-Theimer ML, Zhao Z, Feinstone S, et al. Structural evidence for a bifurcated mode of action in the antibody-mediated neutralization of hepatitis C virus. *Proc Natl Acad Sci U S A*. 2013; 110:7418–7422. [PubMed: 23589879]
- 22. Drummer HE, Maerz A, Pournourios P. Cell surface expression of functional hepatitis C virus E1 and E2 glycoproteins. *FEBS Lett*. 2003; 546:385–390. [PubMed: 12832074]
- 23. Flint M, Logvinoff C, Rice CM, McKeating JA. Characterization of infectious retroviral pseudotype particles bearing hepatitis C virus glycoproteins. *J Virol*. 2004; 78:6875–6882. [PubMed: 15194763]
- 24. Vieyres G, Thomas X, Descamps V, Duverlie G, Patel AH, Dubuisson J. Characterization of the envelope glycoproteins associated with infectious hepatitis C virus. *J Virol*. 2010; 84:10159–10168. [PubMed: 20668082]
- 25. Sabahi A, Uprichard SL, Wimley WC, Dash S, Garry RF. Unexpected structural features of the hepatitis C virus envelope protein 2 ectodomain. *J Virol*. 2014; 88:10280–10288. [PubMed: 24991010]
- 26. Krey T, d'Alayer J, Kikuti CM, Saulnier A, Damier-Piolle L, Petitpas I, Johansson DX, Tawar RG, Baron B, Robert B, et al. The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. *PLoS Pathog*. 2010; 6:e1000762. This reference provide the first predictive model of HCV E2, based on that E2 was classified as class II fusion protein and contains similar organization as flavivirus E protein. [PubMed: 20174556]
- 27. Li Y, Wang J, Kanai R, Modis Y. Crystal structure of glycoprotein E2 from bovine viral diarrhea virus. *Proc Natl Acad Sci U S A*. 2013; 110:6805–6810. [PubMed: 23569276]
- 28. El Omari K, Iourin O, Harlos K, Grimes JM, Stuart DI. Structure of a pestivirus envelope glycoprotein E2 clarifies its role in cell entry. *Cell Rep*. 2013; 3:30–35. [PubMed: 23273918]
- 29. Houghton M. Prospects for prophylactic and therapeutic vaccines against the hepatitis C viruses. *Immunol Rev*. 2011; 239:99–108. This reference provides the importance of use of E1/E2 recombinant proteins for potential vaccine studies. [PubMed: 21198667]
- 30. Law JL, Chen C, Wong J, Hockman D, Santer DM, Frey SE, Belshe RB, Wakita T, Bukh J, Jones CT, et al. A hepatitis C virus (HCV) vaccine comprising envelope glycoproteins gpE1/gpE2 derived from a single isolate elicits broad cross-genotype neutralizing antibodies in humans. *PLoS One*. 2013; 8:e59776. This reference provide further encouragement to use HCV glycoproteins for cross-protective immunity in human volunteers. [PubMed: 23527266]

Figure 1.

Schematic representation of HCV polyprotein (a), highlighting the E1 (b) and E2 (c) domain organization. The PDB ID and construct design for each crystal structure are given. Location of cysteines (black bars) and N-linked glycosylation sites (Y) are indicated for each protein at corresponding positions. The hypervariable regions (HVRs), stem and transmembrane (TM) are labeled accordingly. Note in E1, one glycan is not conserved and is indicated in light gray color.

Figure 2. E2 core structure. (a) Superposition of E2 core from PDB ID 4WMF and 4WEB. (b) Cartoon diagram of individual core structures. (c) Topology diagram of the core domain, highlighting secondary structure and disulfide bonds (yellow).

Figure 3. Structure of nE1. (a) Cartoon diagram of the nE1 monomer colored by rainbow from amino to carboxy terminus. (b) Cartoon diagram showing the six copies in asymmetric unit (AU). The disulfide bonds are represented by spheres colored brown for intramolecular, grey for intermolecular with another molecule in the AU, and black for intermolecular between two AUs.

Figure 4.

A comparison of glycoproteins from *flaviviridae* family. Cartoon diagrams of the ectodomains from Flavivirus E, BVDV E2, HCV nE1 and E2 core. The drawings are on the same scale with the length of each protein provided.