

Production and Characterization of the Ectodomain of E2 Envelope Glycoprotein of Hepatitis C Virus Folded in the Presence of Full-length E1 glycoprotein

Sara Ortega-Atienza^{1#}, Laura Lombana¹ Julián Gómez-Gutiérrez¹, Belén Yélamos¹, Darrell L. Peterson² and Francisco Gavilanes¹.

¹Department of Biochemistry and Molecular Biology, Faculty of Chemistry, University Complutense of Madrid, 28040 Madrid and ²Department of Biochemistry and Molecular Biology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia, 23298

* Corresponding author: F. Gavilanes, Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain. Phone: (+34) 91 394 42 66. Fax: (+34) 91 394 41 59. E-mail: pacog@bbm1.ucm.es

[#]Present address: Department of Pathology and Laboratory Medicine, Brown University, Providence, RI, 02912 USA

Abstract

Hepatitis C virus (HCV) envelope glycoproteins, E1 and E2, are involved in the first steps of virus infection. The E2 ectodomain can be produced as an isolated form (E2₆₆₁). However, there is some concern about its proper conformation and the role that E1 can play as a chaperone for the folding of E2. In order to verify this fact we have expressed a chimeric protein (E1tmbE2) based on the full-length E1 sequence followed by the E2 ectodomain using the baculovirus-insect cells system. The E2 ectodomain is folded in the presence of the E1, proteolytically processed by cellular proteases and secreted to cell culture media (E2_{661p}), while the E1 protein is retained into the cell due to its transmembrane sequence. The purification of E2_{661p} from culture media was facilitated by a His tag introduced in its amino terminus. Both E2₆₆₁ and E2_{661p} glycoproteins shared very similar structural features, monitored by spectroscopic and antigenic studies. Moreover, their functional properties, tested by means of CD81 binding, were almost indistinguishable, indicating that the E2 ectodomain constitutes an independent folding unit.

Keywords: Hepatitis C virus envelope proteins; protein structure; E2 glycoprotein; E2 ectodomain

Highlights: 1. The ectodomain of E2 was folded in a native conformation in the presence of E1. 2. Its properties are the same than those of the ectodomain folded in the absence of E1. 3. E2 ectodomain behaves as an independent folding unit.

Abbreviations: Hepatitis C Virus (HCV), HCV E2 envelope protein ectodomain folded in the absence of E1 ($E2_{661}$), HCV E2 envelope protein ectodomain folded in the presence of E1 ($E2_{661p}$), cell-cultured HCV (HCVcc), CD81 Large Extracellular Loop (CD81-LEL), Convex Constraint Analysis (CCA).

ACCEPTED MANUSCRIPT

Introduction

Hepatitis C virus (HCV) is a positive sense, single-stranded RNA virus belonging to the *Hepacivirus* genus of the *Flaviviridae* family. HCV infection is a major cause of liver disease, having a high propensity for establishing a chronic infection [1]. In the search of a future first vaccine for HCV, a possible strategy is to abolish the initial steps of the HCV infection. In this regard, HCV entry is a highly orchestrated process mediated by viral envelope glycoproteins E1 and E2 and several host factors including heparan sulfate, the tetraspanin CD81, the scavenger receptor BI, the tight junction proteins claudin-1 and occludin, the Niemann-Pic C1-like 1 cholesterol absorption receptor and other factors (reviewed in [2]).

HCV envelope glycoproteins E1 and E2 are components of the viral spike which play an essential role in virus entry into host cells by interacting with cell surface receptors and inducing fusion between the viral and cellular membranes [3, 4]. E2 is the primary glycoprotein responsible for receptor interaction, while little is known about the exact role of E1 protein. They are type I transmembrane glycoproteins, highly glycosylated, with a large N-terminal ectodomain and a C-terminal transmembrane domain. There are reports which indicate that E1 and E2 may assist each other for correct folding in such a way that E2 assists E1 by stabilizing a semi-native conformation meanwhile E1 drives E2 towards a productive folding pathway [5-9]. These observations would indicate that HCV envelope glycoproteins cooperate to form a functional complex. However, E2 protein acquires a three dimensional structure compatible with a native conformation when it is expressed as an isolated form [6, 9] which can block HCV infection *in vitro* [10]. Moreover, recent studies

carried out with E2 ectodomain (residues 384-661 of the polyprotein) produced, either in insect [11] or in mammalian cells [10], have shown that E2 can act as an independent folding unit. Structural characterization of the E2 ectodomain by means of circular dichroism [10-12] and infrared spectroscopy [12], indicates that E2 is mainly constituted by β -sheet and non-ordered structures. Based on gene organization HCV E2 envelope glycoprotein has been classified as a class II fusion protein [13]. One feature of these polypeptides is that they fold as a heterodimer with the previous glycoprotein in the precursor polyprotein. Thus, the first protein in the tandem (E1) would act as a chaperone for the folding of the second one (E2) which is the fusion protein. However, very recently the structure of the core ectodomain of E2 bound to an antigen-binding fragment has been described [14, 15]. The structure differs clearly from the extended conformation of class II fusion proteins being organized as a compact domain composed mainly of β -strands.

The fact that E1 may have some role as a chaperone for the folding of E2 can be exploited to produce E2 ectodomain in a higher proportion. To this end, we have designed a chimeric construct in which full-length E1 sequence is followed by the E2 ectodomain (E1tmbE2 construct). The expression of this construct by using the baculovirus-insect cells system allowed the folding of the E2 ectodomain in the presence of the E1 protein, and after proteolytic processing, its secretion and purification from cell culture medium. In this paper we describe the characterization of this protein, named as E2₆₆₁p, and compare its structural and functional properties with those of the E2 ectodomain folded in the absence of E1 (E2₆₆₁) [10-12].

Materials and Methods

Plasmid constructs. Construct of pAcGP67-A-E2₆₆₁ (residues 384-661) plasmid has been previously described [11]. pAcGP67-A-E1tmbE2 plasmid was constructed using pAcGP67-A-E1E2₇₄₆ plasmid as template. This plasmid comprises E1 (residues 192-383) and E2 (residues 384-746) complete sequences, and it has a His tag in the amino terminal region of E1. Two PCR reactions were carried out in parallel using the same template. For the first one we used the following primers: 5'-G CGA CAT GTC CAT CAC CAT CAC CAT CAC GAA ACC CAC GTC ACC GGG-3' (forward), which introduces a His tag in the N-terminus of E2 ectodomain and *PciI* restriction site and 5'-GC GGA TCC GCG GCC GCT TCA CTC GGA CCT GTC CCT GT-3' (reverse), which hybridizes with 3' end of E2 ectodomain. The second reaction was performed with these two primers: 5'-CAA GAT GGT AAG CGC TAT TG-3' (forward; nucleotides 4185-4204 of pAcGP67-A plasmid), which hybridizes with GP67 sequence of the plasmid, that is located upstream E1 sequence, and 5'-A AGC TTA CAC ATG TGT TTC CGC GTC GAC GCC-3' (reverse), which allows to introduce the first four aminoacids of E2 after E1 sequence and a *PciI* restriction site. These two PCR amplified fragments were digested with *PciI* restriction enzyme, and ligated. A second PCR reaction was carried out to amplify the cDNA ligated with border primers. Amplified cDNA was digested with adequate restriction enzymes (*EcoRI/NotI*) and ligated in pAcGP67-A plasmid digested with the same enzymes. HCV envelope proteins sequence corresponds to genotype 1a, strain PT.

Protein expression and purification. The recombinant E1tmbE2 construct was expressed in Tni insect cells infected with recombinant baculovirus

obtained after transfection with the pAcGP67A-E1tmbE2 plasmid and the linearized viral DNA. Insect cells were grown in X-Press serum-free media prior to infection with high titer virus ($>10^8$ pfu/ml) at a multiplicity of infection of 5-10. Four days after infection, the cells were harvested and centrifuged at 5000 rpm for 10 min and the supernatant dialyzed against 10 mM Tris-HCl pH 7.0, 50 mM NaCl. The chimera E1tmbE2 was proteolitically processed inside the Tni insect cells. The secreted E2₆₆₁p glycoprotein was purified from the supernatant by Ni²⁺-nitrilotriacetic acid agarose chromatography, according to methods previously described for the expression and isolation of E2₆₆₁ [11].

Protein analysis. Protein concentration was determined spectrophotometrically from the absorbance at 280 nm and the extinction coefficient calculated from the amino acid analysis. The absorption spectra were recorded on a Beckman DU-640 spectrophotometer. The amino acid analysis of hydrolyzed aliquots was performed on a Beckman 6300 automatic analyzer. Automated Edman protein degradation of E2₆₆₁p was performed using an Applied gas-phase sequencer (model 494).

Circular Dichroism. CD measurements were carried out on a Jasco spectropolarimeter, model J-715. All the measurements were carried out at room temperature. Far-UV CD spectra were measured at a protein concentrations of 0.1 mg/ml using protein dialyzed against 30 mM MOPS pH 7, 100 mM NaCl. The circular cuvette pathlength was 1 mm. Five scans were averaged for each measurements and the contribution of the buffer was always subtracted. The spectra were calculated by using 110 as the mean residue molecular mass and the results are expressed in terms of residue

molar ellipticity in degree-cm²·dmol⁻¹. The secondary structure of the protein was evaluated by computer fit of the dichroism spectra according to Convex Constraint Analysis (CCA) [16]. This method relies on an algorithm that calculates the contribution of the secondary structure elements that give rise to the original spectral curve without referring to spectra from model systems. The secondary structure was also predicted from the amino acid sequence by the GOR IV method [17].

Fluorescence spectroscopy. Emission spectra were obtained at 25°C using an SLM AMINCO 8000C spectrofluorimeter, fitted with a 450-W xenon arc. Excitation and emission slit widths were set at 4 nm. The protein concentration was 0.05 mg/ml and a 0.2 x 1 cm cuvette was used. Buffer was 30 mM MOPS, pH 7.0, 100 mM NaCl. Excitation was performed at 275 or 295 nm, and the emission spectra were recorded over the range 285-450 nm. The contribution of the buffer was always subtracted. The tyrosine contribution to the emission spectra was calculated by subtracting the emission spectrum measured at $\lambda_{\text{exc}} = 295$ nm multiplied by a factor from that measured at $\lambda_{\text{exc}} = 275$ nm. The factor was obtained from the ratio between the fluorescence intensities measured with $\lambda_{\text{exc}} = 275$ and $\lambda_{\text{exc}} = 295$ nm at wavelengths above 380 nm, where there is no tyrosine contribution.

SDS-PAGE. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli using 15% polyacrylamide gels [18]. Samples were subjected to gel electrophoresis under either nonreducing or reducing conditions (with 5% (v/v) β -mercaptoethanol) and the proteins were stained with Coomassie brilliant blue

R-250. The molecular mass of the protein bands was estimated by comparison with protein markers of known molecular mass (Prestained SDS-PAGE Standards, Bio-Rad).

Western blotting. After SDS-PAGE, proteins were transferred to nitrocellulose membranes (Hybond-ECL; Amersham) in 48 mM Tris-HCl pH 9.0, containing 39 mM glycine, 0.0375% SDS and 20% (v/v) methanol, for 1 h at 1 mA/cm², by using a V20-SDB apparatus (Scie-Plas). To detect proteins, membranes were incubated for 16 h at 4°C either with a HRP-conjugated monoclonal anti-His antibody (Sigma) at a 1:3000 dilution or with a polyclonal goat anti-E2 antibody (US Biologicals) at a 1:500 dilution, or with the A4 monoclonal anti-E1 antibody (kindly provided by Dr. Jean Dubuisson, Institut Pasteur de Lille, France) at a 1:1000 dilution. In the case of incubations with anti-E2 and anti-E1 antibodies, membranes were washed and incubated for 1 h at room temperature with the corresponding secondary antibodies, HRP-conjugated rabbit anti-goat (Jackson Immunoresearch) or goat anti-mouse IgG (Sigma) at a 1:3000 dilution. Membranes were developed with ECL reagents (Amersham Life Sciences) and exposure to photographic film.

Enzyme Linked Immunosorbent Assay (ELISA). 96 wells microtitre plates (Nunc F96 maxisorp) were coated overnight at 4°C with 100 ng/well of purified recombinant protein diluted to 1 µg/ml in PBS. Unbound antigen was discarded, and the wells were blocked with 3% non-fat dry milk in PBS for 60 min at room temperature. After washing the wells were incubated at 37°C for 2 h with human serum. Six HCV-positive human sera and ten HCV-negative human sera were used at a dilution of 1:100-1:500. The plates were washed

three times with PBS/0.05% Tween 20 and incubated at 37 °C for 1 h with HRP-conjugated anti-human IgG (Fc) (Sigma) diluted at 1:10000. Bound antibodies were detected by adding 100 mM sodium citrate pH 5.0, 4% Methanol buffer containing H₂O₂ and the substrate *o*-phenylenediamine dihydrochloride (OPD) (Merck). The optical density at 490 nm was measured using an ELISA Expert 96 microplate reader (ASYS Hitech). Sera from infected and control patients were provided by Dr. Fernando Vivanco (Fundación Jiménez Díaz, Madrid, Spain).

For H53 ELISA assays, the same protocol was followed but with these exceptions: 96 well plates were coated with 1 µg /well of purified protein. E2 was detected with H53 antibody at a 1:400 dilution and a goat anti-mouse-HRP conjugated antibody (Sigma) at a 1:3000 dilution. The monoclonal antibody H53 is conformation-dependent and was kindly provided by Dr. Jean Dubuisson (Intitut Pasteur de Lille, France).

LEL-CD81 binding assay A 100 µl aliquot of a 3 µg/ml solution of glutathione S-transferase (GST) fusion protein expressing the second extracellular region (EC2) of the LEL of human CD81 (GST-hLEL) [19] was used to coat each well of a Nunc F96 Maxisorp, 96 well-plate and incubated overnight at room temperature. Unbound protein was discarded, and the wells were blocked with 2% non-fat dry milk in PBS/0.05% Tween 20 for 2 h at room temperature. After washing the wells with PBS/0.05% Tween 20, 100 µl of soluble E2 protein at different concentrations (0-10 µg/ml) was incubated with bound LEL-CD81 protein at room temperature for 2 h. LEL-CD81 bound E2 protein was detected using H53 antibody followed by a HRP-conjugated goat anti-mouse IgG (Sigma) at a 1:3000 dilution for 1 h at room temperature.

Detection was performed by adding 100 mM sodium citrate pH 5.0, 4% Methanol buffer containing H₂O₂ and the substrate OPD (Merck). The optical density at 490 nm was measured.

ACCEPTED MANUSCRIPT

Results and Discussion

E2_{661p} expression and purification. Previous studies using S2 Drosophila cells have demonstrated that a truncated form of E2 can be processed and secreted after expression of a polypeptide containing the HCV envelope genome lacking the E2 transmembrane domain (E1-E2 Δ TM) [12]. Following a similar procedure we have designed the chimera E1tmbE2 (Fig. 1). After the full-length E1 sequence this polypeptide contains an N-terminal His-tagged E2 ectodomain, in order to facilitate the purification of the processed E2_{661p}, that should be almost identical to the previously described E2₆₆₁ [11].

The intra and extracellular expression of E1 and E2 proteins by Tni infected cells was evaluated using anti-E1, anti-E2 and anti-His antibodies by western blot (Fig. 2a). The theoretical masses of E1 and E2_{661p} based on the amino acid sequences are 22 and 32 kDa respectively. It can be observed that the E1tmbE2 polypeptide was proteolytically processed by cellular proteases to generate the E2 ectodomain. This domain was partially secreted to the extracellular medium, while the E1 protein was retained into the infected cells. The secreted E2 ectodomain folded in the presence of E1 was named as E2_{661p} in order to distinguish it from E2₆₆₁, folded in the absence of E1 [11]. Intracellular E1 and E2 fractions presented complex patterns of bands, probably due to unfolded and non-processed proteins and even proteolytically cleaved proteins with different glycosylation degrees. E2_{661p} protein secreted to the extracellular medium was subsequently purified using affinity chromatography on a Ni²⁺-nitrilotriacetic acid-agarose column. The protein that elutes with 200 mM imidazole had a purity higher than 95% upon densitometry of stained SDS gels (Fig. 2b). The purified protein was recognized by monoclonal anti-His and

goat anti-E2 antibodies (Fig. 2b). Following this procedure, approximately 1 mg of E2_{661p} glycoprotein was obtained from 1 L of culture media. This low yield of purification compared to that of E2₆₆₁ (5-6 mg/ L culture medium [11]) could be due to the lower secretion levels after improper processing of the chimeric protein and/or interaction of E2 with E1 retained in the insect cells.

Biochemical characterization of E2_{661p}. The Edman degradation of purified E2_{661p} confirmed that the amino-terminal sequence was ETHVHHHHHH, indicating that the E1tmbE2 polypeptide was properly processed after residue 383 of E1. The sequence ETHV corresponds to residues 384-387 of E2.

SDS-PAGE of recombinant E2_{661p} in the presence of reducing agents showed a single band with a molecular mass of 48 kDa (Fig. 3a), coincident with that of E2₆₆₁ but higher than the theoretical mass based on the amino acid sequence. Based on the fact that E2₆₆₁ can be partially deglycosylated by N-glycosydase F [11] it can be suggested that both E2₆₆₁ and E2_{661p} proteins have a similar glycosylation pattern, and possibly an analogous folding process. In the absence of reducing agents, E2₆₆₁ and E2_{661p} migrate in SDS-PAGE as at least four bands corresponding to monomers, dimers, trimers and tetramers (Fig. 3b), suggesting that oligomers are constituted by disulfide bonds linked monomers. The trend in oligomer distribution is maintained in the different preparations obtained. Nevertheless, the amount of monomers observed in the case of E2_{661p} is higher than in E2₆₆₁ protein (Fig. 3b). This result could be explained because E2_{661p} is more diluted in the extracellular medium and, hence, intermolecular interactions should occur with less probability. A second possible explanation could be that small conformational changes in the regions implied in oligomerization were taking

place when E2 is folding in the presence of E1, leading to lower interaction between monomers. If this were the case, E1 would have a chaperone-like role in the folding of E2.

Spectroscopic characterization of E2_{661p}. The spectroscopic characterization of E2_{661p} was carried out by means of circular dichroism and fluorescence spectroscopies. The spectra were taken in parallel with E2₆₆₁ in order to check if folding in the presence of E1 induces conformational changes in E2 ectodomain. The far-UV CD spectrum of both proteins showed a minimum at 206 nm (Fig. 4a). Deconvolution of these spectra using the program CCA [16] yielded the percentages of secondary structure elements showed in Table 1. As observed, both proteins showed a very similar secondary structure pattern, with a low helical content being the β -sheet the main ordered structure. The predictive GOR IV method [17], which is based upon the propensity of each amino acid to adopt a particular secondary structure yielded basically similar results (theoretical, Table 1). Also, these results agree well with the recently published structure of the core of E2 ectodomain [14, 15]. Moreover, these results show that there are no changes in terms of secondary structure when the E2 ectodomain is folded in the presence of E1.

Fluorescence emission spectrum of E2_{661p} showed a maximum at 332 nm upon excitation at 275 nm, which is 2 nm lower than that exhibited by E2₆₆₁ which is centered at 334 nm (Fig. 4b). In both cases, the difference between the fluorescence spectra obtained upon excitation at 275 and 295 nm, the latter being normalized, indicated that the contribution of tyrosine residues to the emission fluorescence upon excitation at 275 nm was very

low, approximately 5 % (Fig. 4b). Thus, we can conclude that in both proteins the Trp residues are located in a very similar environment of relatively low hydrophobicity. Nevertheless, as denoted by the 2 nm shift of the maximum to lower wavelength, E2_{661p} fluorophores could be located in a barely more non polar microenvironment than the E2₆₆₁ ones, as a result of a slightly more compact folding of E2 ectodomain in the presence of E1. This observation would be in accordance with the lower tendency of E2_{661p} to aggregate, since a more compact conformation diminishes the surface exposition of hydrophobic domains and, thus, reduces the probability of intermolecular interactions.

Antigenic characterization of E2_{661p}. A panel of six HCV-positive and ten HCV-negative human sera was used to assess the antigenic properties of E2_{661p}, following the ELISA described in Methods. E2_{661p} and E2₆₆₁ proteins were able to bind to antibodies present in the six HCV-positive human sera (Fig. 5, sera 1, 2, 3, 4, 5 and 6) at different levels, whereas they were not recognized by immunoglobulins from HCV-negative sera (Fig. 5, sera 7-8; only two of the ten negative sera are depicted). The reactivity pattern was very similar in both cases. However, it must be noted that the mean signal obtained with E2_{661p} is about 30% lower than the observed with E2₆₆₁. This result could indicate that the protein folded in presence of E1 has a lower exposition of its epitopes, probably as a consequence of its slightly more compact conformation.

Moreover, the reactivities of both proteins with the anti-E2 conformation-dependent H53 monoclonal antibody were almost identical (Fig.

5, inset). Again, these results suggest that both E2₆₆₁ and E2_{661p} possess very similar conformations.

Binding of E2_{661p} to CD81-LEL. The specific interaction with the human tetraspanin CD81-LEL region has been regularly used to monitor the functional properties of HCV E2 glycoprotein [20, 21]. Using a recombinant human CD81-LEL (residues 116-202) fusion protein with GST (Glutathione S-Transferase), an ELISA binding assay has been performed with both E2₆₆₁ and E2_{661p} proteins. The results shown in Fig. 6 indicate that both recombinant glycoproteins bind to CD81-LEL in a concentration dependent manner, with almost identical reactivity. Although the interaction of CD81 is specific for E2, recent studies using E1-cysteine mutants in cell-cultured HCV (HCVcc) indicated that E1 glycoprotein has a modulating effect in the binding of E2 to CD81 [22]. However, this effect must be restricted to the context of the E1E2 heterodimer, since the isolated E2 ectodomain has the same CD81-binding capacity when it folds either independently (E2₆₆₁) or in the presence of E1 (E2_{661p}).

To sum up, the results shown herein, carried out with isolated and soluble recombinant proteins, highlight that the presence of E1 is indeed not necessary for the correct folding of the E2 ectodomain, since E2₆₆₁ and E2_{661p} presented very similar structural, antigenic and functional characteristics. Although there are reports in the literature that also support the independent folding of E2 ectodomain, most of them are based on mammalian cell culture transient expression assays with viral and non-viral vectors. Thus, using Sindbis and vaccinia virus vectors, Michalak et al. [6] showed that the folding of E2 is independent of E1, but E2 is required for the proper folding of E1. The effect of

co-expression of E1 and E2 glycoproteins on each other's folding was evaluated after immunoprecipitation of the expressed proteins from different mammalian cell culture supernatants with the help of a conformation-sensitive monoclonal antibody (for E2) or by analysing intramolecular disulfide bond formation (for E1). The fact that E2 ectodomain could be expressed independently of the polyprotein and that it folds in a structure able to bind to cellular receptors [23, 24] is a further evidence of its independent folding.

One strategy to increase the quantity and even the quality of recombinant proteins using the baculovirus-insect cells system has been to co-express them with a chaperone. Thus, it has been described that co-expression of both protein disulfide isomerase and the chaperone BiP increased the solubility and secretion of immunoglobulin G from insect cell [25, 26]. The postulated chaperone-like function of E1 is not translated into higher quantity but its co-expression gives rise to a slightly more compact E2 ectodomain with a lower tendency to form disulfide-bridge oligomers. In this regard, there is some controversy with respect the functionality of the HCV envelope glycoproteins and the formation of intermolecular covalent complexes. On the one hand, in cells expressing recombinant HCV glycoproteins, E1 and E2 proteins have been shown to form both E1E2 noncovalent heterodimers and large covalent E1E2 complexes, with the latter corresponding to nonfunctional aggregates while the former representing the possible functional unit for HCV entry [27]. On the other hand, more recently it has been described by using the HCVcc system the selective incorporation into secreted virions of large E1E2 disulfide-linked complexes that maintain a native conformation as they were able to bind conformation-sensitive neutralizing antibodies [28]. Furthermore, a recombinant chimeric protein based on the E1 and E2 ectodomains (E1₃₄₁E2₆₆₁) was

produced in the baculovirus system with native-like properties in spite of having a high tendency to self-associate by forming intermolecular disulfide bridges [29]. Then, the disulfide-bridge oligomerization of HCV envelope glycoproteins seems to be compatible with the maintenance of their native properties. The results presented in this work, although related to a processed glycoprotein (E2₆₆₁) instead of the E1E2 heterodimer, confirm this point of view. Moreover, the E2 ectodomain folded in the absence of E1, which can be produced with a much higher yield and with a higher degree of disulfide-linked complexes [11], showed a 30% increased ability to bind to human antibodies, making this recombinant protein potentially suitable for developing future HCV vaccines.

Acknowledgements

This work was supported by Grant BFU 2010-22014 from the Ministerio de Economía y Competitividad (Spain)

References

- [1] Alter, H. Viral hepatitis, *Hepatology* 43 (2006) S230-4.
- [2] Ploss, A. & Dubuisson, J. New advances in the molecular biology of hepatitis C virus infection: towards the identification of new treatment targets, *Gut* 61 Suppl 1 (2012) i25-35.
- [3] Bartosch, B., Dubuisson, J. & Cosset, F. L. Infectious hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes, *J. Exper. Med.* 197 (2003) 633-642.
- [4] Nielsen, S. U., Bassendine, M. F., Burt, A. D., Bevitt, D. J. & Toms, G. L. Characterization of the genome and structural proteins of hepatitis C virus resolved from infected human liver, *J. Gen. Virol.* 85 (2004) 1497-1507.
- [5] Botti, V., Bianchi, A., Fong, S. K. & Merola, M. The hepatitis C virus E1 glycoprotein undergoes productive folding but accelerated degradation when expressed as an individual subunit in CHO cells, *PLoS One* 6 (2011) e23838.
- [6] Michalak, J.-P., Wychowski, C., Choukhi, A., Meunier, J.-C., Ung, S., Rice, C. M. & Dubuisson, J. Characterization of truncated forms of hepatitis C virus glycoproteins, *J. Gen. Virol.* 78 (1997) 2299-2306.
- [7] Patel, J., Patel, A. H. & McLauchlan, J. The transmembrane domain of the hepatitis C virus E2 glycoprotein is required for correct folding of the E1 glycoprotein and native complex formation, *Virology* 279 (2001) 58-68.
- [8] Brazzoli, M., Helenius, A., Fong, S. K. H., Houghton, M., Abrignani, S. & Merola, M. Folding and dimerization of hepatitis C virus E1 and E2 glycoproteins in stably transfected CHO cells, *Virology* 332 (2005) 438-453.

- [9] Cocquerel, L., Quinn, E. R., Flint, M., Hadlock, K. G., Fong, S. K. H. & Levy, S. Recognition of native hepatitis C virus E1E2 heterodimers by a human monoclonal antibody, *J. Virol.* 77 (2003) 1604-1609.
- [10] Whidby, J., Mateu, G., Scarborough, H., Demeler, B., Grakoui, A. & Marcotrigiano, J. Blocking hepatitis C virus infection with recombinant form of envelope protein 2 ectodomain, *J. Virol.* 83 (2009) 11078-89.
- [11] Rodriguez-Rodriguez, M., Tello, D., Yelamos, B., Gomez-Gutierrez, J., Pacheco, B., Ortega, S., Serrano, A. G., Peterson, D. L. & Gavilanes, F. Structural properties of the ectodomain of hepatitis C virus E2 envelope protein, *Virus Res.* 139 (2009) 91-9.
- [12] Krey, T., d'Alayer, J., Kikuti, C. M., Saulnier, A., Damier-Piolle, L., Petitpas, I., Johansson, D. X., Tawar, R. G., Baron, B., Robert, B., England, P., Persson, M. A., Martin, A. & Rey, F. A. The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule, *PLoS Pathog.* 6 (2010) e1000762.
- [13] Yagnik, A. T., Lahm, A., Meola, A., Roccasecca, R. M., Ercole, B. B., Nicosia, A. & Tramontano, A. A model for the hepatitis C virus envelope glycoprotein E2, *Proteins: structure, function, and genetics* 40 (2000) 355-366.
- [14] Khan, A. G., Whidby, J., Miller, M. T., Scarborough, H., Zatorski, A. V., Cygan, A., Price, A. A., Yost, S. A., Bohannon, C. D., Jacob, J., Grakoui, A. & Marcotrigiano, J. Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2, *Nature* 509 (2014) 381-4.
- [15] Kong, L., Giang, E., Nieuwma, T., Kadam, R. U., Cogburn, K. E., Hua, Y., Dai, X., Stanfield, R. L., Burton, D. R., Ward, A. B., Wilson, I. A. & Law, M.

- Hepatitis C virus E2 envelope glycoprotein core structure, *Science* 342 (2013) 1090-4.
- [16] Perczel, A., Hollósi, M., Tusnády, G. & Fasman, G. D. Deconvolution of the circular dichroism spectra of proteins: The circular dichroism spectra of antiparallel β -sheet in proteins., *Protein Eng.* 4 (1991) 669-679.
- [17] Garnier, J., Gibrat, J.-F. & Robson, B. (1996) GOR method for predicting protein secondary structure from amino acid sequence in *Methods in Enzymology* (Doolittle, R. F., ed) pp. 540-553.
- [18] Laemmli, V. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4., *Nature* 227 (1970) 680-685.
- [19] Higginbottom, A., Quinn, E. R., Kuo, C.-C., Flint, M., Wilson, L. H., Bianchi, E., Nicosia, A., Monk, P. N., McKeating, J. A. & Levy, S. Identification of Amino Acid Residues in CD81 Critical for Interaction with Hepatitis C Virus Envelope Glycoprotein E2, *J. Virol.* 74 (2000) 3642-3649.
- [20] Flint, M., Thomas, J. M., Maidens, C. M., Shotton, C., Levy, S., Barclay, W. S. & McKeating, J. A. Functional analysis of cell surface-expressed hepatitis C virus E2 glycoprotein, *J. Virol.* 73 (1999) 6782-6790.
- [21] Petracca, R., Falugi, F., Galli, G., Norais, N., Rosa, D., Campagnoli, S., Burgio, V., Stasio, E., Giardina, B., Houghton, M., Abrignani, S. & Grandi, G. Structure-Function Analysis of Hepatitis C Virus Envelope-CD81 Binding, *J. Virol.* 74 (2000) 4824-4830.
- [22] Wahid, A., Helle, F., Descamps, V., Duverlie, G., Penin, F. & Dubuisson, J. Disulfide bonds in hepatitis C virus glycoprotein E1 control the assembly and entry functions of E2 glycoprotein, *J. Virol.* 87 (2013) 1605-17.

- [23] Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A. J., Houghton, M., Rosa, D., Grandi, G. & Abrignani, S. Binding of hepatitis C virus to CD81, *Science* 282 (1998) 938-941.
- [24] Scarselli, E., Ansuini, H., Cerino, R., Roccasecca, R. M., 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.* 21 (2002) 5017-5025.
- [25] Whiteley, E. M., Hsu, T. A. & Betenbaugh, M. J. Thioredoxin domain non-equivalence and anti-chaperone activity of protein disulfide isomerase mutants in vivo, *J. Biol. Chem.* 272 (1997) 22556-63.
- [26] Hsu, T. A. & Betenbaugh, M. J. Coexpression of molecular chaperone BiP improves immunoglobulin solubility and IgG secretion from *Trichoplusia ni* insect cells, *Biotechnol. Prog.* 13 (1997) 96-104.
- [27] Op De Beeck, A., Voisset, C., Bartosch, B., Ciczora, Y., Cocquerel, L., Keck, Z., Fong, S., Cosset, F. L. & Dubuisson, J. Characterization of functional hepatitis C virus envelope glycoproteins, *J. Virol.* 78 (2004) 2994-3002.
- [28] Vieyres, G., Thomas, X., Descamps, V., Duverlie, G., Patel, A. H. & Dubuisson, J. Characterization of the envelope glycoproteins associated with infectious hepatitis C virus, *J. Virol.* 84 (2010) 10159-68.
- [29] Tello, D., Rodriguez-Rodriguez, M., Yelamos, B., Gomez-Gutierrez, J., Ortega, S., Pacheco, B., Peterson, D. L. & Gavilanes, F. Expression and structural properties of a chimeric protein based on the ectodomains of E1 and E2 hepatitis C virus envelope glycoproteins, *Protein Expr. Purif.* 71 (2010) 123-31.

Table 1. Percentages of secondary structure of E2₆₆₁ and E2_{661p}. Experimental secondary structure percentages were calculated using CCA method [16] from far-UV CD spectra, while theoretical values were obtained using GOR IV method [17] for three components.

% Secondary Structure	Experimental		Theoretical
	E2 ₆₆₁	E2 _{661p}	
α -helix	2	7	2
β -sheet	50	53	34
Non ordered	48	40	64

Figure Legends

Fig. 1. Schematic representation of E1tmbE2 construct. The construct has a His tag in the N-terminal ends of both E1 and E2 proteins. E1 sequence is complete, including its transmembrane domain (tmb), while E2 sequence is truncated at 661 residue (E2 ectodomain). The first four residues of E2 (384-387) (ETHV) were included after E1 sequence and before the His tag of E2, in order to ensure a proper proteolytic cleavage after residue 383 of E1.

Fig. 2. Expression and purification of E2_{661p} protein. (a) Western blot analysis of the extracellular (E) and the intracellular (I) E1 and E2 proteins produced by baculovirus-infected insect cells. The presence of both proteins was detected with a polyclonal anti-E2 antibody, a monoclonal anti-His antibody and a monoclonal anti-E1 antibody. The arrows indicate the position of the corresponding molecular mass marker. A control (C) of uninfected cells is shown. (b) SDS-PAGE (Coomassie brilliant blue R-250 staining) and western blot analysis (with a monoclonal anti-His antibody and a polyclonal anti-E2 antibody) of purified E2_{661p} recombinant protein. The samples were previously reduced with 5% (v/v) β -mercaptoethanol and boiled for 5 min. The blots were developed with ECL reagents as indicated in the Materials and Methods section. The results shown are representative of those obtained for three different experiments.

Fig. 3. Oligomeric state of E2_{661p}. Purified E2_{661p} protein was subjected to SDS-PAGE in the presence (a) or absence (b) of the reducing agent β -

mercaptoethanol. E₂₆₆₁ purified protein was used as reference. The arrows indicate the position of the corresponding molecular mass marker. (1) Monomer, (2) dimer, (3) trimer and (4) tetramer. The gel was stained with Coomassie brilliant blue R-250. The results shown are representative of those obtained for three different experiments.

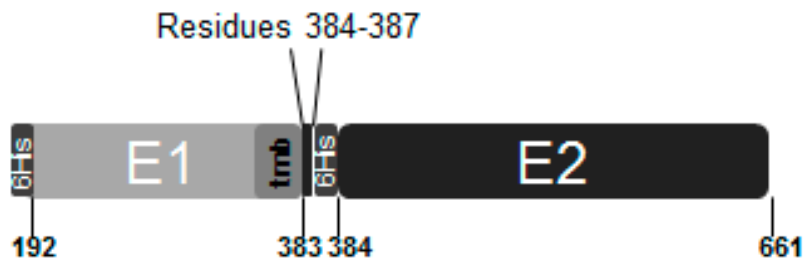
Fig. 4. Spectroscopic characterization of E_{2661p}. (a) Far-UV CD spectra of E₂₆₆₁ (—) and E_{2661p} (- - -) at pH 7.0. The spectra were recorded between 190 and 250 nm with a protein concentration of 0.1 mg/ml in a cylindrical cuvette of 0.1 cm pathlength. Spectra were recorded five times, averaged and corrected for buffer contributions. Data are expressed as mean residue molar ellipticity. (b) Fluorescence emission spectra of E₂₆₆₁ (—) and E_{2661p} (- - -) at pH 7.0 upon excitation at 275 nm. The Tyr contribution, calculated as described in Materials and Methods, is also shown, E₂₆₆₁ (· · ·) and E_{2661p} (-· ·· ·). The protein concentration was 0.05 mg/ml. The buffer employed was 30 mM MOPS pH 7.0, 100 mM NaCl and its contribution was always subtracted. Spectra were collected at room temperature. The spectra shown are representative of those obtained for three different protein preparations. The results shown are representative of those obtained for three different experiments.

Fig. 5. Reactivity of E₂₆₆₁ and E_{2661p} against HCV-positive (1-6 sera) and negative human sera (7 and 8 sera). 96-well plates were coated with 100 ng of purified recombinant protein/well. After washing the wells were incubated with the sera diluted between 1:100-1:500. Absorbance signals were normalized according to the A₄₉₀ provided by E₂₆₆₁ with serum n^o 4 which gave the highest

signal. 10 negative sera were assayed and all of them provided the same reactivity as that obtained with sera 7 and 8. (Inset) Reactivity of E2_{661p} and E2₆₆₁ against H53 conformation-dependent monoclonal antibody. Protein buffer contribution was subtracted. Means and SD of three independent experiments are shown.

Fig. 6. Binding of E2₆₆₁ and E2_{661p} to human CD81-LEL. 96-well plates were coated with 300 ng of CD81-LEL/well. After washing, soluble E2 protein was added at different concentrations and incubated for 2 h at room temperature. H53 and goat anti-mouse-HRP conjugated antibodies were used to detect E2 bound protein. Protein buffer contribution was subtracted. Values were normalized according to the A₄₉₀ provided by E2₆₆₁ at 10 µg/ml. Means and SD of three independent experiments are shown.

Figure 1



ACCEPTED MANUSCRIPT

Figure 2

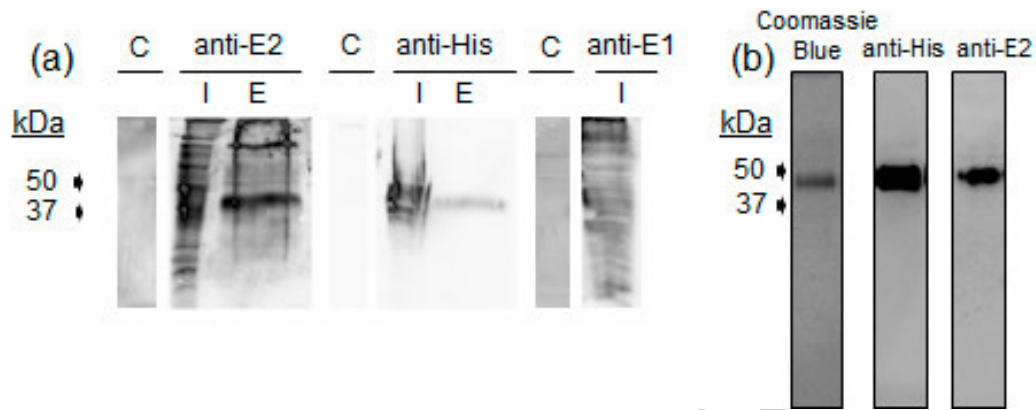


Figure 3

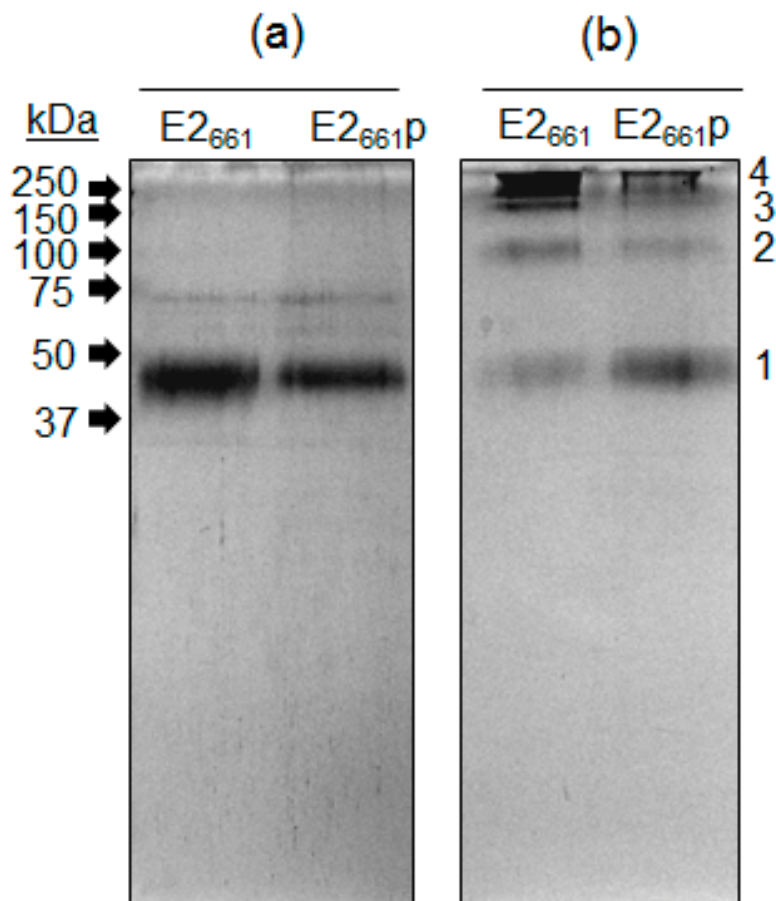


Figure 4

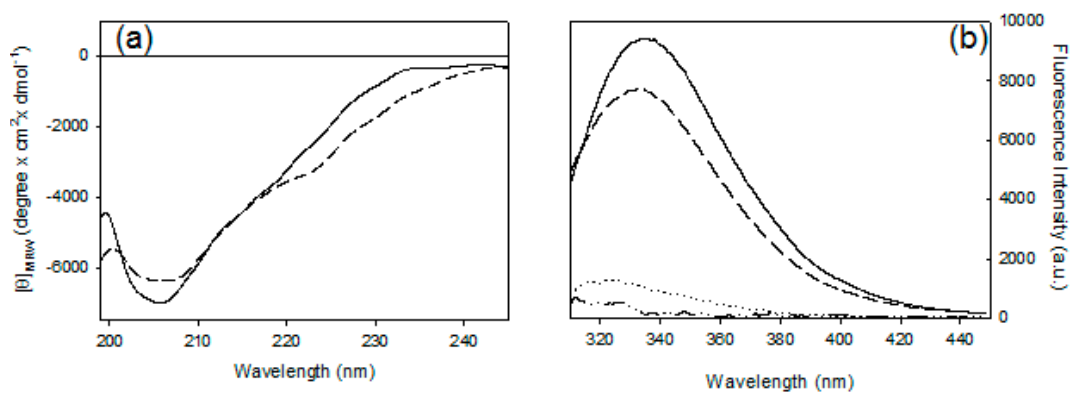


Figure 5

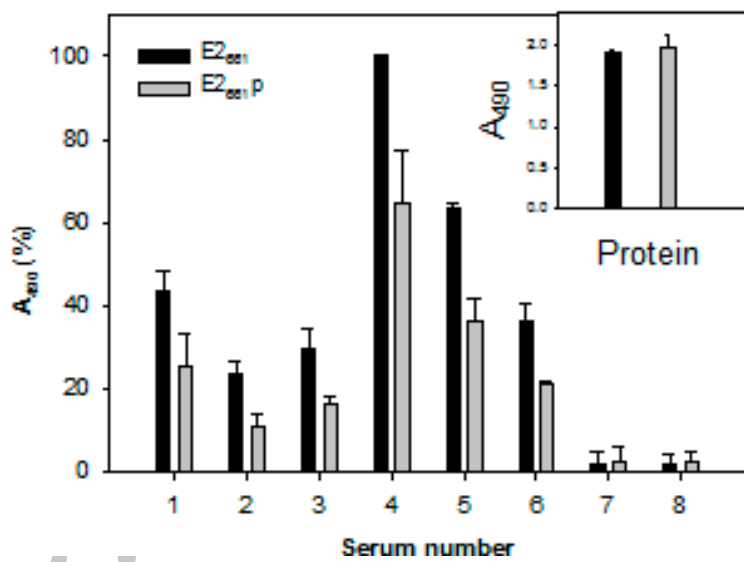
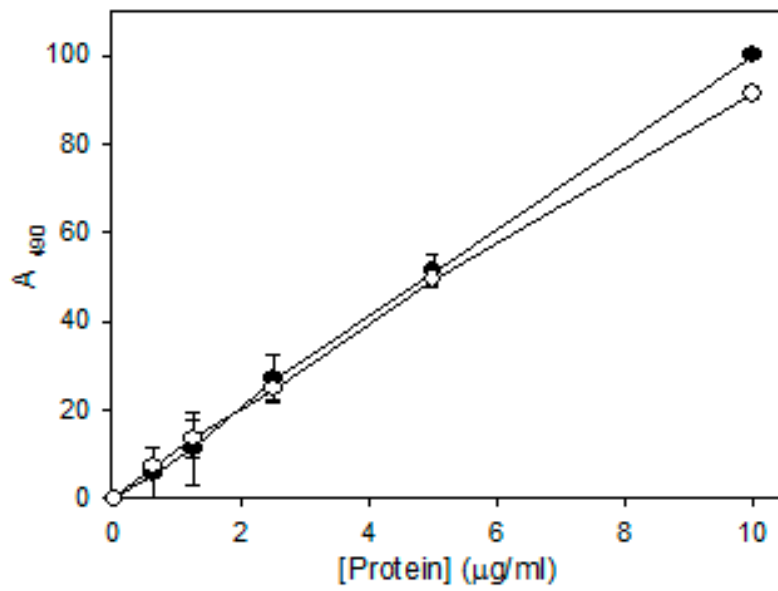


Figure 6



ACCEPTED MANUSCRIPT

SCRIPT