Structural insights into hepatitis C virus receptor binding and entry

https://doi.org/10.1038/s41586-021-03913-5

Received: 5 May 2020

Accepted: 13 August 2021

Published online: 15 September 2021

Check for updates

Ashish Kumar[1], Reafa A. Hossain[1], Samantha A. Yost[2], Wei Bu[3], Yuanyuan Wang[1], Altaira D. Dearborn[1], Arash Grakoui[4,5], Jeffrey I. Cohen[3] & Joseph Marcotrigiano[1]✉

Hepatitis C virus (HCV) infection is a causal agent of chronic liver disease, cirrhosis and hepatocellular carcinoma in humans, and afflicts more than 70 million people worldwide. The HCV envelope glycoproteins E1 and E2 are responsible for the binding of the virus to the host cell, but the exact entry process remains undetermined[1]. The majority of broadly neutralizing antibodies block interaction between HCV E2 and the large extracellular loop (LEL) of the cellular receptor CD81 (CD81-LEL)[2]. Here we show that low pH enhances the binding of CD81-LEL to E2, and we determine the crystal structure of E2 in complex with an antigen-binding fragment (2A12) and CD81-LEL (E2–2A12–CD81-LEL); E2 in complex with 2A12 (E2–2A12); and CD81-LEL alone. After binding CD81, residues 418–422 in E2 are displaced, which allows for the extension of an internal loop consisting of residues 520–539. Docking of the E2–CD81-LEL complex onto a membrane-embedded, full-length CD81 places the residues Tyr529 and Trp531 of E2 proximal to the membrane. Liposome flotation assays show that low pH and CD81-LEL increase the interaction of E2 with membranes, whereas structure-based mutants of Tyr529, Trp531 and Ile422 in the amino terminus of E2 abolish membrane binding. These data support a model in which acidification and receptor binding result in a conformational change in E2 in preparation for membrane fusion.

HCV enters hepatocytes through a multistep process that requires a series of host cellular factors and the viral envelope glycoproteins E1 and E2 (reviewed previously[1]). The HCV glycoproteins mediate cell targeting, endocytosis and membrane fusion, which is ultimately stimulated by endosomal acidification[3]. At least four cellular factors are critical for HCV attachment and entry: CD81, scavenger receptor class B type I (SRBI), claudin-1 (CLDN) and occludin (OCLN); however, blocking the E2–CD81 interaction is the primary means of antibody-mediated neutralization[2]. CD81 is ubiquitously expressed in a variety of cell lines, indicating that it has a role secondary to hepatocyte-specific receptor binding. CD81 translocates with the virion to tight junctions and engages with the late-entry factors CLDN and OCLN in the endosome for acidification and entry. CD81 is an integral membrane protein of the tetraspanin family that contains four transmembrane helices. The CD81-LEL—a globular domain made of five helices (A–E)—binds E2, and residues that are essential for the interaction have previously been identified[4–9]. The molecular mechanisms thereafter, which mediate the cell entry and membrane fusion of HCV, remain to our knowledge undefined.

Initial crystallization trials showed that the presence of low pH is a critical determinant for crystal formation. We therefore measured the affinity of human and tamarin CD81-LEL (hCD81-LEL and tCD81-LEL, respectively) for the ectodomain of E2 (eE2) at neutral and low pH. Differing from human CD81 by only five amino acids (Extended Data Fig. 1), tamarin CD81 supports HCV infection and binds E2 more effectively[10,11]. tCD81-LEL showed a fourfold increase in affinity for eE2 compared to hCD81-LEL (175 nM and 773 nM, respectively) (Extended Data Table 1, Extended Data Fig. 2). Low pH (5.0) increased the affinity of tCD81-LEL for eE2 by 4.7-fold (175 nM at pH 7.5 to 37 nM at pH 5.0), whereas hCD81-LEL showed only a modest 1.1-fold increase in affinity (773 nM at pH 7.5 to 681 nM at pH 5.0) (Extended Data Table 1, Extended Data Fig. 2). A low-pH complex of tCD81-LEL–eE2(ΔHVR1) (in which ΔHVR1 indicates deletion of the hypervariable region 1 of eE2) with 2A12 (an antigen-binding fragment (Fab) of a non-neutralizing antibody)[12] as a crystallization chaperone yielded crystals that diffracted to a resolution of about 3.3 Å. To assist in the identification of changes after complex formation, structures of full-length, fully glycosylated eE2–2A12 as well as tCD81-LEL alone were also determined (Fig. 1, Extended Data Table 2).

Conformational changes in E2 and CD81

The tCD81-LEL–eE2(ΔHVR1)–2A12 structure has two complexes in the asymmetric unit, which allows two independent observations of the interaction (Fig. 1a, Extended Data Fig. 3). The non-crystallographic, two-fold symmetry axis resembles the homodimer that is observed in the tCD81-LEL structure (Fig. 1c) and in previously reported hCD81-LEL structures[13,14], and is likely to be a biochemical artifact as the interface clashes with the transmembrane helices of the full-length CD81 structures[15,16]. Each tCD81-LEL is bound to a copy of eE2(ΔHVR1), and each eE2(ΔHVR1) is, in turn, bound to a 2A12 Fab. There are no contacts between the E2 molecules in the asymmetric unit (Extended Data Fig. 3)

[1]Structural Virology Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [2]Center for Advanced Biotechnology and Medicine, Rutgers, The State University of New Jersey, Piscataway, NJ, USA. [3]Medical Virology Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA. [5]Emory Vaccine Center, Division of Microbiology and Immunology, Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA. ✉e-mail: joseph.marcotrigiano@nih.gov

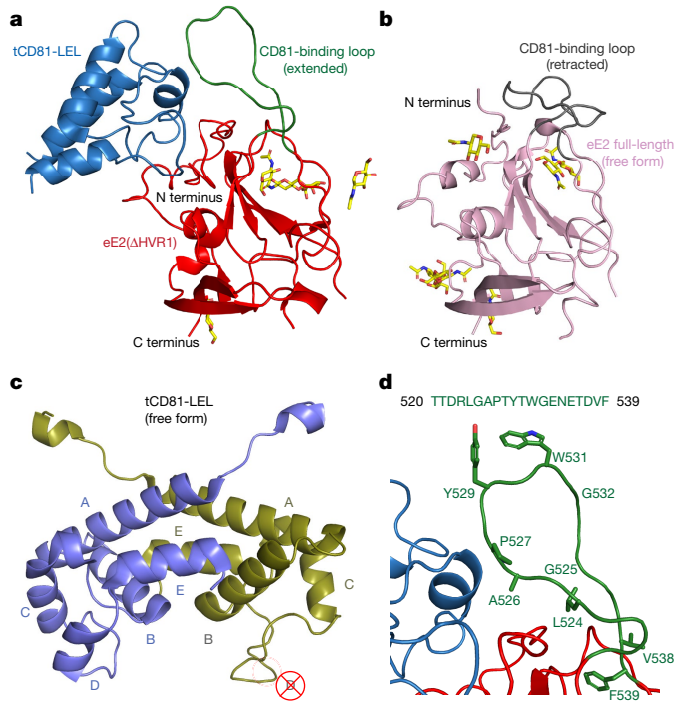


Fig. 1 | Ribbon diagrams of the tCD81-LEL–eE2(ΔHVR1)–2A12, eE2–2A12 and tCD81-LEL X-ray crystal structures. a, **b**, The tCD81-LEL (blue) and eE2(ΔHVR1) (red and green) complex (**a**) and eE2 alone (pink and grey) (**b**), highlighting the location of the CD81-binding loop (green in **a** and grey in **b**). The 2A12 Fab is not shown. Carbohydrate molecules are coloured by heteroatom. **c**, The homodimeric, asymmetric unit of tCD81-LEL, free form, with helices A–E labelled on the intermediate (violet) and open (olive) conformations. The unwound helix D in the open conformation is labelled with a crossed-out D. **d**, The tCD81-LEL(blue) and eE2(ΔHVR1) (red and green) interface in detail with the side chains of the CD81-binding loop residues (green and heteroatom) represented in sticks. The CD81-binding loop sequence is shown at the top.

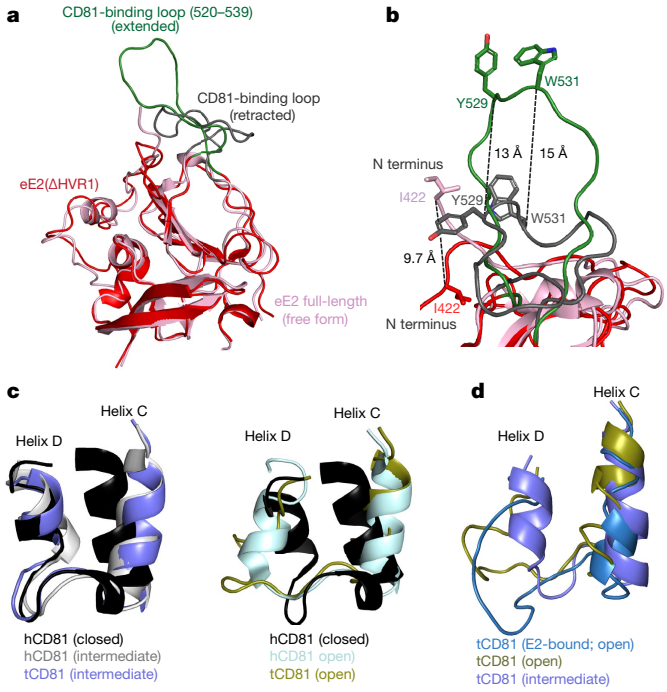


Fig. 2 | Ribbon diagrams of the conformational variation in E2 and CD81-LEL. a, **b**, Superposition of eE2 in the presence (red and green) and absence (pink and grey) of tCD81-LEL. The CD81-binding loops (residues 520–539) are shown in the presence (green) and absence (grey) of tCD81-LEL. In **b**, the relative alpha carbon movements (dotted lines) of Ile422, Tyr529 and Trp531 side chains (sticks) are shown and N termini are labelled. **c**, Superposition of the unbound C and D helices of tCD81-LEL (violet and olive) and hCD81-LEL (PDB entries 5M2C and 5M33 show the black, light grey and aqua in the closed, intermediate and open conformations, respectively). **d**, Superposition of tCD81-LEL unbound intermediate (violet) and open (olive) conformations with the open E2-bound conformation (blue).

but the complexes are highly similar, with an alpha carbon root mean squared deviation (r.m.s.d.) of 1.0 Å.

The overall structure of E2 in the eE2–2A12 and tCD81-LEL–eE2(ΔHVR1)–2A12 complexes (Fig. 1a, b) is similar to that reported previously[12,17–20]. We observed two noteworthy conformational changes in residues 418–422 and 520–539 (the CD81-binding loop) of E2 (Fig. 1a, b, d, Extended Data Fig. 4). In the absence of CD81, residues 384–421 (HVR1 and antigenic site (AS) 412) of eE2 are disordered, and the CD81-binding loop is packed against residues 422–427 (Figs. 1b, 2a, b, Extended Data Fig. 4). The tip of the CD81-binding loop (represented by Tyr529) interacts with Ile422 (Fig. 2b). In the presence of tCD81-LEL, residues 418–421 become ordered and wrap around CD81, with Ile422 moving more than 9 Å (Fig. 2b). The CD81-binding loop extends by 13–15 Å to pack partly against tCD81-LEL (Figs. 1a, 2a, b). The CD81-binding loop adopts a similar conformation in each complex in the asymmetric unit (Extended Data Fig. 5). CD81 binding appears to stabilize an alternate conformation of residues 418–427, including Ile422, which allows for loop extension (Fig. 2b). This mechanism is supported by a previous structure of the E2 core, in which the 520–539 loop was disordered owing to the deletion of residues 384–455, including Ile422 (ref. [12]).

Previous hCD81-LEL structures were classified into 'open', 'intermediate', and 'closed' conformations on the basis of the relative orientations of the C and D helices[14] (residues 160–188) (Fig. 2c). The vast majority of the described hCD81-LEL structures are in the closed conformation, including the LEL from the full-length structure of hCD81[14,21]. The D helix in each tCD81-LEL molecule is unwound in the tCD81-LEL–eE2(ΔHVR1)–2A12 complex structure, adopting instead an extended, open conformation, consistent with the flexibility that was observed in this region using nuclear magnetic resonance[22] (Fig. 2d). Furthermore, in the structure of tCD81-LEL alone the two molecules in the asymmetric unit adopt an intermediate conformation with a D helix, and an extended, open conformation, respectively (Figs. 1c, 2d). This open conformation is distinct from the open conformation of human CD81 when bound to CD19, in which helices D and E join the fourth transmembrane helix as a contiguous structure and half of helix C is unwound[16]. The five differences in amino acids between human CD81 and tamarin CD81 were mapped in their relation to the E2-binding site (Extended Data Fig. 1b). Of the five differences, only T163S and N180S contribute to the CD81–E2 interface, whereas D155N, V169M and D196E are directed away from E2. The enhanced affinity of tamarin CD81 for E2 may be due to tCD81-LEL adopting an extended open conformation that is needed for E2 binding and/or these contact differences.

Neutralizing antibody and CD81 competition

The CD81-binding site on E2 is discontinuous, comprising the AS412 epitope and front layer (residues 412–459), the central CD81-binding loop and residues 616–617 of the back layer (Fig. 3, Extended Data Figs. 4, 6, Extended Data Table 3). The two tCD81-LEL–eE2(ΔHVR1) complexes in the asymmetric unit have 872 Å^2 and 959 Å^2 of buried surface with an r.m.s.d. of 1.1 Å for similar alpha carbon positions, suggesting binding tolerance in the interface. In both tCD81-LEL–eE2(ΔHVR1) complexes, eE2 residues Leu441, Phe442, Tyr443 and His445 of the front layer and Tyr617 of the back layer make contact with CD81 (as defined by a distance of at least 4 Å) (Extended Data Table 3).

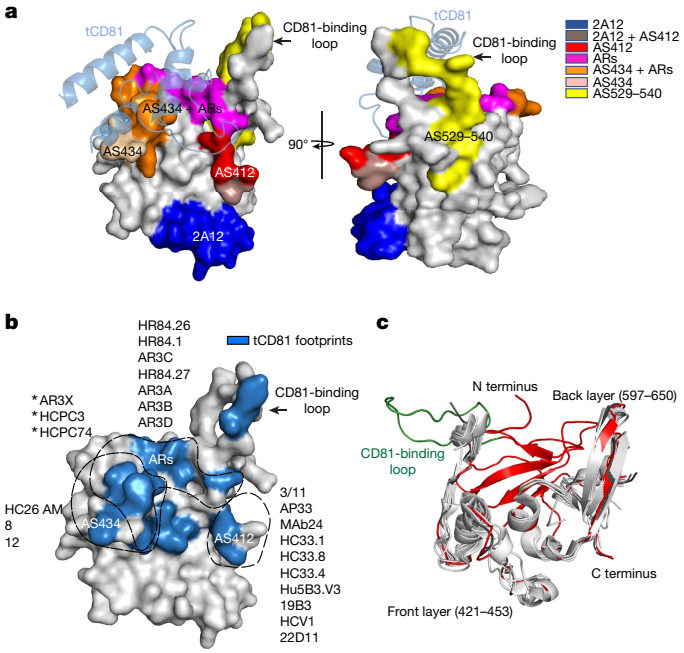


Fig. 3 | Neutralizing antibodies compete directly with CD81 for E2 binding. **a**, Antibody footprints, coloured as in the key, mapped onto surface rendering (light grey) of eE2(ΔHVR1) relative to tamarin CD81 (transparent blue ribbon diagram). The right panel is rotated 90° about a vertical axis relative to the left panel. ARs, antigenic regions. **b**, Antibody-binding regions (dashed line, labelled) that compete for the CD81-LEL footprint (blue) on surface-rendered (light grey) eE2(ΔHVR1). Asterisks denote antibodies that share common regions on Fabs antigenic regions and AS434. **c**, Superimposed ribbon diagrams of the E2 front and back layers from chain C (red) of the tCD81-LEL–eE2(ΔHVR1)–2A12 complex with the CD81-binding loop in green and eE2 free form (dark grey) with various E2–Fab complexes (PDB entries 6MEI, 6MEH, 6BKC, 6URH, 4WEB, 4MWF, 6BKD and 6BKB are coloured in light grey).

These residues are highly conserved across HCV genotypes, with Leu441, Tyr443 and Tyr617 being invariant across genotypes. HK6a (genotype 6a) has a lysine residue at position 445 and SA13 (genotype 5a) has a leucine residue at position 442 (Extended Data Fig. 4).

DAO5, a non-neutralizing antibody, recognizes the AS529–540 epitope on the CD81-binding loop distal from CD81, consistent with its inability to neutralize infection[23] (Fig. 3a). A crystal structure of DAO5 bound to a synthetic peptide shows a portion of the CD81-binding loop in an alpha-helical conformation, which is not observed in either the retracted or the extended form in the eE2 structures. DAO5 is also capable of capturing HCV pseudoparticles, cell-culture-derived virions and eE2 in solution[23], suggesting that the CD81-binding loop may be dynamic.

The E2–CD81 interaction is targeted by a series of well-characterized broadly neutralizing antibodies, and several structures of E2 in complex with broadly neutralizing antibodies have been determined[2]. Potent broadly neutralizing antibodies isolated from individuals infected with HCV mainly target the overlapping AS412, AS434 and antigenic regions (Fig. 3a, b). These antibodies sterically clash with linear and conformational epitopes that are involved in the binding of CD81 to E2, indicating that neutralization is accomplished by directly blocking the CD81-binding surface on E2 (Fig. 3b, Extended Data Fig. 7, Extended Data Table 3). Superposition of the front and back layers of the published structures of E2–antibody structures onto the E2–CD81 complex yielded an r.m.s.d. that ranged from 0.45 Å to 1.44 Å, indicating that receptor and broadly neutralizing antibody binding may result in slight changes in the conformation of these layers (Fig. 3c). All of the determined broadly neutralizing antibody–eE2 structures have the CD81-binding loop in the retracted position[12,17–20].

Superposition onto full-length CD81

To provide a better understanding of the effect of CD81 binding on HCV entry, each tCD81-LEL–eE2(ΔHVR1) structure was superimposed onto the full-length, human CD81 structure (Protein Data Bank (PDB) entry 5TCX[21]) and docked into an idealized membrane bilayer. The two planes of spheres represent the carbonyl moieties of the phospholipids and approximate the hydrophobic core of the bilayer[24] (Fig. 4a). E2 binds to the edge of CD81 distal from the four transmembrane helices and proximal to the bilayer.

To investigate the effect of charge and pH-sensitive regions in the context of CD81 binding, the electrostatic-potential surfaces of E2 and CD81 were calculated for both the bound and the free forms at pH 7.5 and pH 5.0 (Extended Data Fig. 7). The electrostatic surface of the eE2–tCD81-LEL interface does not show a marked change when the pH is decreased (Extended Data Fig. 7a, b, e, f). Protonation of histidine side chains after endosomal acidification can actuate protein conformational changes in other viral fusion proteins[25]. E2 His421 and His445 are within 4 Å of CD81 (Extended Data Table 3), but far from the extended loop of E2. These observations are consistent with a previous proposal that the protonation state of His445 of E2 is a key regulator of the low-pH-dependent fusion mechanism that is used by HCV[26]. In fact, the presence of a lysine residue in the HCV HK6a strain retains a positive charge at this position (Extended Data Figs. 4, 6). The electrostatic surface of the eE2–CD81-LEL complex is basic and becomes more so when the pH is lowered (Extended Data Fig. 7c, d). The superposition model places half of the basic surface of eE2 proximal to the negatively charged membrane while the other half wraps around the side of the complex (Fig. 4b). More studies are needed to clarify the function of the extended form of the E2 loop and pH sensing during HCV infection.

In each complex of the asymmetric unit, the CD81-binding loop extends away from E2 towards the bilayer. Tyr529, Trp531 and Gly532, located at the tip of the loop, are invariant among the major genotypes (Extended Data Fig. 4) and are critical for CD81 binding and virus entry[4,7]. The side chains of Tyr529 and Trp531 are oriented towards the outer leaflet of the membrane at distances of less than 5 Å and less than 10 Å from the idealized phosphatidyl carbonyl layer, respectively (Fig. 4a). The hydrophilic head groups of the phospholipids would extend around 3–5 Å from the hydrophobic core, placing Tyr529 in contact with the membrane. Given the inherent flexibility of both the CD81-binding loop and the membrane bilayer, it would be energetically favourable to insert the side chains of Tyr529 and Trp531 into the outer leaflet of the membrane. The intervening residue at position 530 is not conserved within the different HCV genotypes but is generally polar, which may serve to orient the loop relative to the lipid head groups.

To evaluate the effects of pH, tCD81-LEL and select E2 mutations on membrane binding, liposome flotation was used as described previously[27]. eE2 Y529A and W531A have greatly reduced human CD81 binding and I422A retains about 50% binding, but each is recognized by conformational antibodies, suggesting that they are correctly folded[4,7]. Wild-type eE2 was incubated with liposomes in the presence or absence of tCD81-LEL at either pH 7.5 or pH 5.0, separated in a sucrose gradient and detected by E2-specific western blotting (Fig. 4c). Liposome-bound proteins migrate to the top of the gradient whereas free proteins remain at the bottom. Wild-type eE2 showed weak membrane binding at pH 7.5, which improved slightly in the presence of tCD81-LEL. Lowering the pH to 5.0 increased liposome binding, which was enhanced further in the presence of tCD81-LEL. The partial flotation of eE2 is consistent with previous results[27] and is probably due to the small hydrophobic exposed surface formed by Tyr529 and Trp531 (Fig. 4b). During HCV entry, insertion of the loop would be assisted by the membrane-embedded, full-length CD81. Mutant versions of eE2 (I422A, Y529A, W531A and Y529A/W531A) exhibited liposome binding of less than 10% relative to wild-type eE2 at pH 5.0 in the presence of

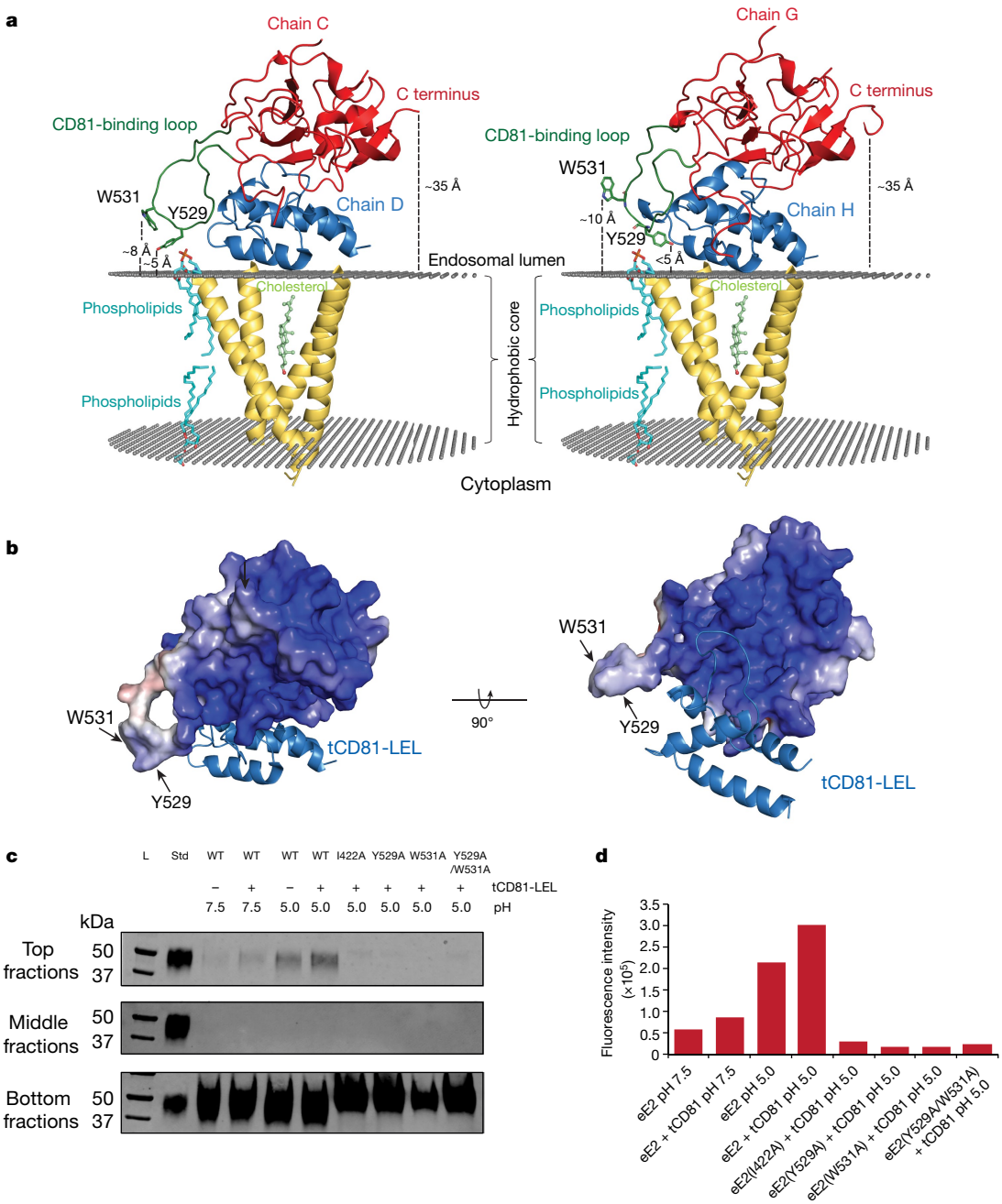


Fig. 4 | Proximity and interaction of eE2 with membranes. a, The full-length human CD81 structure (ribbon diagram and transmembrane regions in yellow) and coordinated cholesterol molecule (light-green heteroatom sticks) (PDB entry 5TCX) were docked into an idealized palmitoyloleoylphosphatidylcholine (POPC) membrane bilayer. Parallel planes of grey spheres represent the carbonyl moieties that define the hydrophobic core, and representative phospholipids (cyan heteroatom sticks) are shown. Ribbon diagrams of each complex of tCD81-LEL (blue) and eE2(ΔHVR1) (red) (the CD81-binding loop is shown in green) in the asymmetric unit are superimposed on the LEL and are shown in the same orientation. Relative distances (dotted lines) of Tyr529 and Trp531 (green heteroatom sticks) and of the carboxyl terminus of eE2(ΔHVR1) from the hydrophobic membrane core are labelled. **b**, Electrostatic surface of eE2(ΔHVR1) bound to tCD81-LEL (blue ribbon diagram) calculated at pH 5.0 in the same orientation as A (left) and rotated 90° about a horizontal axis (right) to show membrane-proximal surface. The surface is coloured by electrostatic potential corresponding to +5 kcal/(mol·e) (blue) and −5 kcal/(mol·e) (red), where e is a unit of charge. **c**, E2-specific western blots of the top, middle and bottom fractions from liposome flotation in sucrose gradients with a protein molecular weight marker (L) and eE2 loading control marker (std). One representative western blot is shown from three independent experiments. Each experiment showed the same trend; that is, enhanced flotation of eE2 in the presence of tCD81-LEL and low pH. For gel source data, see Supplementary Fig. 1. **d**, Quantification of the top-fraction western blot in **c** with arbitrary units. In **c**, **d** the pH, inclusion of tCD81-LEL, and eE2 wild-type (WT) and mutants are labelled.

tCD81-LEL (Fig. 4c, d, Extended Data Fig. 8). Doubling the amount of mutant eE2 proteins per flotation assay did not enhance membrane flotation (Extended Data Fig. 8). Liposome binding by these mutants was markedly lower than CD81-independent liposome binding by the wild-type eE2 at pH 7.5 (Fig. 4c). Thus, although these E2 mutations are important for CD81 binding and do not noticeably misfold the protein, their effect on liposome binding cannot be attributed solely to reduced CD81 binding.

Discussion

Viral membrane fusion involves a two-step mechanism: priming (for example, proteolysis) and triggering (for example, acidification and/or receptor binding)[28]. Once triggered, the trimeric viral glycoprotein introduces a fusion loop or peptide into the cellular membrane, and this is followed by a conformational rearrangement that draws the two membranes together. HCV entry involves cell-type recognition and binding, translocation to the tight junctions and membrane fusion to the endosome. HCV fusion requires both E1 and E2 glycoproteins as well as low pH and is primed by CD81-LEL[27]. Thus far, there is little evidence of an eE2 trimer, although the E2 stem, transmembrane helix or E1 may influence oligomerization. The structural data we provide here demonstrate a mechanism by which HCV E2 binds to the cellular receptor CD81 at low pH, which results in the extension of an internal E2 loop towards the endosomal membrane. A fusion loop for HCV has yet to be identified, but the CD81-binding loop exhibits many necessary characteristics (that is, membrane binding, low-pH trigger and CD81-dependent extension). Furthermore, there is some additional evidence that a fusion loop may exist in E1 (residues 264–294)[29]. For fusion to occur, a conformational change must bring the viral, membrane-embedded transmembrane helix at the carboxyl terminus of E2 into contact with the host membrane—a distance of 35 Å according to our docking model (Fig. 4a). The intervening stem region of E2, which was omitted in this study, could span this distance and warrants further investigation. Together, our results show that during entry to the host cell, E2 uses a hybrid triggering mechanism in which both acidification and CD81 interaction are necessary for optimal membrane binding.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03913-5.

1. Gerold, G., Moeller, R. & Pietschmann, T. Hepatitis C virus entry: protein interactions and fusion determinants governing productive hepatocyte invasion. *Cold Spring Harb. Perspect. Med.* **10**, a036830 (2020).
2. Tzarum, N., Wilson, I. A. & Law, M. The neutralizing face of hepatitis C virus E2 envelope glycoprotein. *Front. Immunol.* **9**, 1315 (2018).
3. Tscherne, D. M. et al. Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. *J. Virol.* **80**, 1734–1741 (2006).
4. Rothwangl, K. B., Manicassamy, B., Uprichard, S. L. & Rong, L. Dissecting the role of putative CD81 binding regions of E2 in mediating HCV entry: putative CD81 binding region 1 is not involved in CD81 binding. *Virol. J.* **5**, 46 (2008).
5. Drummer, H. E., Wilson, K. A. & Poumbourios, P. Identification of the hepatitis C virus E2 glycoprotein binding site on the large extracellular loop of CD81. *J. Virol.* **76**, 11143–11147 (2002).
6. Drummer, H. E., Boo, I., Maerz, A. L. & Poumbourios, P. A conserved Gly4 36-Trp-Leu-Ala-Gly-Leu-Phe-Tyr motif in hepatitis C virus glycoprotein E2 is a determinant of CD81 binding and viral entry. *J. Virol.* **80**, 7844–7853 (2006).
7. Owsianka, A. M. et al. Identification of conserved residues in the E2 envelope glycoprotein of the hepatitis C virus that are critical for CD81 binding. *J. Virol.* **80**, 8695–8704 (2006).
8. Zhao, Z. et al. A neutralization epitope in the hepatitis C virus E2 glycoprotein interacts with host entry factor CD81. *PLoS One* **9**, e84346 (2014).
9. Higginbottom, A. et al. Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J. Virol.* **74**, 3642–3649 (2000).
10. Flint, M. et al. Diverse CD81 proteins support hepatitis C virus infection. *J. Virol.* **80**, 11331–11342 (2006).
11. Allander, T., Forns, X., Emerson, S. U., Purcell, R. H. & Bukh, J. Hepatitis C virus envelope protein E2 binds to CD81 of tamarins. *Virology* **277**, 358–367 (2000).
12. Khan, A. G. et al. Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. *Nature* **509**, 381–384 (2014).
13. Kitadokoro, K. et al. CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs. *EMBO J.* **20**, 12–18 (2001).
14. Cunha, E. S. et al. Mechanism of structural tuning of the hepatitis C virus human cellular receptor CD81 large extracellular loop. *Structure* **25**, 53–65 (2017).
15. Dearborn, A. D. & Marcotrigiano, J. Hepatitis C virus structure: defined by what it is not. *Cold Spring Harb. Perspect. Med.* **10**, a036822 (2020).
16. Susa, K. J., Rawson, S., Kruse, A. C. & Blacklow, S. C. Cryo-EM structure of the B cell co-receptor CD19 bound to the tetraspanin CD81. *Science* **371**, 300–305 (2021).
17. Flyak, A. I. et al. HCV broadly neutralizing antibodies use a CDRH3 disulfide motif to recognize an E2 glycoprotein site that can be targeted for vaccine design. *Cell Host Microbe* **24**, 703–716 (2018).
18. Kong, L. et al. Hepatitis C virus E2 envelope glycoprotein core structure. *Science* **342**, 1090–1094 (2013).
19. Flyak, A. I. et al. An ultralong CDRH2 in HCV neutralizing antibody demonstrates structural plasticity of antibodies against E2 glycoprotein. *eLife* **9**, e53169 (2020).
20. Tzarum, N. et al. Genetic and structural insights into broad neutralization of hepatitis C virus by human VH1-69 antibodies. *Sci. Adv.* **5**, eaav1882 (2019).
21. Zimmerman, B. et al. Crystal structure of a full-length human tetraspanin reveals a cholesterol-binding pocket. *Cell* **167**, 1041–1051 (2016).
22. Rajesh, S. et al. Structural basis of ligand interactions of the large extracellular domain of tetraspanin CD81. *J. Virol.* **86**, 9606–9616 (2012).
23. Vasiliauskaite, I. et al. Conformational flexibility in the immunoglobulin-like domain of the hepatitis C virus glycoprotein E2. *mBio* **8**, e00382-17 (2017).
24. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
25. Kielian, M. Mechanisms of virus membrane fusion proteins. *Annu. Rev. Virol.* **1**, 171–189 (2014).
26. Boo, I. et al. Distinct roles in folding, CD81 receptor binding and viral entry for conserved histidine residues of hepatitis C virus glycoprotein E1 and E2. *Biochem. J.* **443**, 85–94 (2012).
27. Sharma, N. R. et al. Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *J. Biol. Chem.* **286**, 30361–30376 (2011).
28. White, J. M. & Whittaker, G. R. Fusion of enveloped viruses in endosomes. *Traffic* **17**, 593–614 (2016).
29. Li, H. F., Huang, C. H., Ai, L. S., Chuang, C. K. & Chen, S. S. Mutagenesis of the fusion peptide-like domain of hepatitis C virus E1 glycoprotein: involvement in cell fusion and virus entry. *J. Biomed. Sci.* **16**, 89 (2009).

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Construct design, expression and purification of eE2(ΔHVR1), eE2, tCD81-LEL and eE2 mutants

E2(ΔHVR1) (residues 406–656) and eE2 (full length, residues 384–656) from HCV genotype J6, and CD81 (residues 112–202 from human and tamarin) were expressed in HEK293T GnTI$^-$ cells (D. Comoletti) and purified as described previously[30]. In brief, the proteins of interest were cloned into a lentiviral vector containing a CMV promoter, a prolactin signal sequence, the desired gene fragment and an HRV3C cleavage site followed by a C-terminal protein-A and Flag tags. Stable expressing GnTI$^-$ HEK293T cells were produced by lentiviral transduction. Cells were grown in an adherent cell bioreactor (Cesco Bioengineering) for long-term growth and protein production. Supernatants were collected every two days and purified by IgG affinity chromatography and eluted by GST-HRV3C protease digestion. The elution was purified by subtractive chromatography over GST and Q columns followed by size-exclusion chromatography. Final yields for all constructs were 5–10 mg per litre of supernatant. eE2 mutants I422A, Y529A, W531A and Y529A/W531A (double mutant) of eE2 (full length, residues 384–656) from HCV genotype J6 were cloned into the same plasmid and expressed in suspension Expi239 GnTI$^-$ cell culture (Thermo Fisher Scientific) by transient transfection method. An ExpiFectamine 293 transfection kit (Thermo Fisher Scientific) was used per the manufacturer's protocol for high protein yield. Supernatants were collected on the sixth day after transfection. eE2 mutants were purified by IgG, GST and Q column as detailed above.

Production, purification and production of the 2A12 Fab

The protocol is adapted from a previous study[12] with slight modification. The large-scale growth of the mouse 2A12 hybridoma (provided by A.G.) was achieved through a CELLine Classic bioreactor flask (Sigma-Aldrich). A total of 6×10^6 cells in 6 ml of Iscove's modified Dulbecco's medium (IMDM) with 15% low-IgG FBS, and 10 mM HEPES pH 7.5 (culture medium) were inoculated in the inner layer of the cell, and the upper membrane was covered with 350 ml of IMDM with 1% low-IgG FBS, and 10 mM HEPES pH 7.5 (nutrient medium). The culture medium was collected after 4–6 days when hybridoma cell confluency reaches 6×10^8 cells. The medium was centrifuged at 1,000 rpm for 20 min, and the supernatant was further purified. 2A12 was purified through protein G column and dialysed in 10 mM EDTA and 20 mM sodium phosphate pH 7.0. Just before digestion with papain, cysteine-HCl was added to a final concentration of 20 mM and the pH was adjusted to 7.0. Approximately 100 µl of immobilized agar bead papain (Thermo Fisher Scientific) was used for 20 mg of 2A12 antibody and incubated for 3 h at 37 °C by gentle inversion. Reducing and non-reducing gels confirmed complete digestion of 2A12 antibody. Immobilized papain was separated by centrifuge at 4,200*g* at 4 °C for 20 min. Supernatant was loaded on protein A column and flow-through was collected and further loaded onto a Protein G column that was pre-equilibrated with 20 mM HEPES pH 7.5, 250 mM NaCl and 5% glycerol. 2A12 Fabs were eluted from the Protein G column by 0.05% TFA, and immediately neutralized by 1 M Tris pH 8.0 and desalted into 20 mM Tris pH 8.0.

Crystallization

eE2–2A12 complex. The 2A12 Fab was incubated with eE2 at a ratio of 1.1:1 (w/w) for 1–2 h at 4 °C before purification over a Superdex200 size-exclusion column (Cytiva Life Sciences) equilibrated with 20 mM HEPES pH 7.5 and 100 mM NaCl. The complex was concentrated to 10 mg ml^{-1} and crystals were grown at 4 °C using the hanging drop vapour diffusion method. Exactly 2 µl of protein complex was mixed with 2 µl of well solution containing 4% v/v tacsimate pH 5.0, 14% w/v PEG3350, and 4% D-(+)-trehalose dihydrate. Crystals were first seen after 7–8 days and continued to grow until day 10. Crystals were cryoprotected using well solution supplemented with 30% glycerol and flash-cooled in liquid nitrogen. Data were collected at a wavelength of 0.979 Å using the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) 31-ID beamline at the Advanced Photon Source (APS), Argonne National Laboratory.

Tamarin CD81-LEL. Tamarin CD81-LEL was purified over a Superdex200 column equilibrated with 20 mM HEPES pH 7.5 and 100 mM NaCl. The protein was then concentrated to approximately 14 mg ml^{-1} before being grown into crystals using the hanging drop vapour diffusion method at 20 °C. Exactly 2 µl of protein was mixed with an equal amount of reservoir solution containing 0.1 M HEPES pH 7.5, and 55% w/v PEG200. Single cubic crystals were obtained one day after setting up. Crystals were flash-cooled in liquid nitrogen without additional cryoprotectant and data were collected at a wavelength of 0.979 Å at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the APS, Argonne National Laboratory.

eE2–(ΔHVR1)–2A12–tCD81-LEL. Initial crystallization trials with several different E2 and CD81-LEL constructs produced crystals that diffracted poorly (4.5 Å resolution or worse). In an attempt to improve the resolution, plasmid-encoded amino acids and HVR1 were removed from tCD81-LEL and eE2, respectively. eE2(ΔHVR1) and 2A12 Fab complex was mixed in a 1:1.2 w/w ratio and then mixed with tCD81-LEL in a 1:1.2 molar ratio in a buffer containing 20 mM sodium acetate pH 4.5 and 100 mM NaCl, and incubated overnight on ice at 4 °C. The complex was purified over a Superdex200 column equilibrated with 20 mM sodium acetate pH 4.5 and 100 mM NaCl. The complex was concentrated to 10 mg ml^{-1} and crystals were grown at 4 °C using the hanging drop vapour diffusion method. Diffraction quality crystals were grown by microseeding in 0.2 M sodium acetate pH 4.5, 4% v/v tacsimate pH 4.0 and 12% w/v PEG Smear Medium 1 to 2 days after setting up the 24-well plate. Crystals were cryoprotected in a stepwise manner with mother liquor supplemented with 30% w/v PEG Smear Medium, and flash-cooled in liquid nitrogen. Data were collected at 0.979 Å at the SER-CAT 22-ID beamline at the APS, Argonne National Laboratory.

Structure determination and refinement

eE2 (genotype J6)–2A12 crystals belong to space group $C222_1$ with cell parameters a = 95.84 Å, b = 155.56 Å and c = 129.42 Å. Phases were determined by the molecular replacement method using PHENIX_phaser[31] and the coordinates were obtained from PDB entry 4WEB. Unambiguous placement of the Fab heavy and light chains provided the necessary phases to extend the map to cover eE2 coordinates, using iterative rounds of model building and density modification by Coot[32], and the model was evaluated during refinement for various parameters. The final model was built to a resolution of 2.71 Å, comprising residues 422–523, 490–571 and 596–650 of eE2 from the J6 genome with 6 *N*-linked *N*-acetylglucosamine, 1 beta-D-mannopyranose (ref. [33]) and 102 solvent molecules. The 2A12 Fab light chain and heavy chain consist of 1–218 and 1–219 amino acids, respectively. The model coordinates were refined to R_{work} 0.218 and R_{free} 0.271 with 92.11% Ramachandran favoured, allowed 7.05, and 0.84% outliers calculated from MolProbity[34]. The overall $CC_{1/2}$ (Pearson correlation coefficient) of the processed data is 0.989, and the $CC_{1/2}$ in the outer shell is 0.487.

Tamarin CD81-LEL crystals belong to space group $I23$ with cell parameters a = 113.26 Å, b = 113.26 Å and c = 113.26 Å, and data were collected at 1.80 Å resolution. PDB entry 5TCX was truncated to LEL and the chain was mutated to polyalanine. The edited model was used to determine the initial phases by using PHENIX_phaser[31]. Iterative rounds of model building in Coot[32] and refinement in PHENIX_refine led to R_{work} 0.207

and R_{free} 0.231 with 96.83% Ramachandran favoured, 3.17% allowed, and 0% outliers. Chain A and chain B start from 106–201 and 106–202, respectively. Residues 106–111 were from the backbone of the pJG construct, and were also built in the electron density with 66 solvent, 2 polyethylene glycol and 1 glycerol molecule. The overall $CC_{1/2}$ of the processed data is 0.997 and the outer-shell $CC_{1/2}$ is 0.612.

eE2(ΔHVR1)–2A12–tCD81-LEL complex crystals belong to space group $P2_12_12_1$ with cell parameters a = 76.95 Å, b = 127.77 Å and c = 212.37 Å. Phases were determined by the molecular replacement method using PHENIX_phaser[31] and the Fab coordinates were taken from the eE2 J6–2A12 structure. We found two molecules per asymmetric unit, and further eE2(ΔHVR1) and tamarin CD81 coordinates were placed by PHENIX_phaser, which further improved the phases. The model was built and refined in Coot[32] and refined in PHENIX_refine at 3.32 Å resolution. The final structure comprised chain C 418–453, 490–571 and 597–650 residues, and chain G 415–453, 490–571, and 596–654 of eE2(ΔHVR1). Both heavy chains (A, E) and light chains (B, F) of the 2A12 Fab consist of 1–218 amino acids. Chains D and H of tamarin CD81 consist of 119–199 and 116–199 residues, respectively. Seven *N*-linked *N*-acetylglucosamine and one BMA were also modelled in the structure[33] with refined R_{work} 0.241 and R_{free} 0.286, and 91.87% Ramachandran favoured, 7.4% allowed and 0.73% outliers. The overall $CC_{1/2}$ of the processed data is 0.994 and the $CC_{1/2}$ in the outer shell is 0.398. The statistics for the data processing and structure refinement of all three structures are summarized in Extended Data Table 2.

Sequence alignment and surface electrostatic potential

Multiple sequence alignment was performed by Clustal Omega[35], and surface electrostatic potential was visualized using PyMOL with the assistance of the APBS plug-in[36]. The input files for the APBS plug-in at different pH values (5.0 and 7.5) were generated from the online APBS server (https://server.poissonboltzmann.org) using PDB2PQR tools[37] using the default setting for forcefield and output naming scheme. All structural figures were prepared using PyMOL[38].

Isothermal titration calorimetry and data analysis

The protein samples were dialysed against buffers containing 20 mM HEPES pH 7.5 and 100 mM NaCl, or 20 mM sodium citrate pH 5.0 and 100 mM NaCl. A MicroCal ITC-200 isothermal calorimeter was used to carry out calorimetric experiments at 20 °C with stirring at 750 rpm (Malvern Panalytical). The interface between the cell, containing 15–30 μM eE2, and the syringe was equilibrated with 0.4 μl (0.14–0.19 μM tCD81-LEL and 0.30–0.58 μM hCD81-LEL) of ligand with a spacing of 0.8 s, followed by 16 subsequent injections of 2.45 μl (0.87–1.17 μM tCD81-LEL and 1.82–3.51 μM hCD81-LEL) with a spacing of 180 s.

Isothermal titration calorimetry (ITC) thermograms were integrated using the NITPIC program[39] and normalized peak area plots were fitted using SEDPHAT[40]. The A + B ⇄ AB heterodimer model was used to determine eE2–CD81 interaction. Enthalpogram fitting parameters included K_D and ΔH, and eE2 was considered the incompetent fraction. The binding parameter confidence level of 95% and error surface of the fit were estimated with SEDPHAT. ITC measurements were performed at the NHLBI Biophysics Core Facility.

Liposome flotation assay

Three micrograms of wild-type or mutant eE2 was mixed with 3.6 μg tCD81-LEL (a six-fold excess of tCD81-LEL) or buffer. The volume was adjusted to 50 μl with either 20 mM sodium citrate pH 5.0 and 100 mM NaCl, or 20 mM HEPES pH 7.5 and 100 mM NaCl buffer. The sample was incubated on ice overnight at 4 °C. Fifty microlitres of 200-nm Soy PC:Cholesterol liposomes (70:30 molar ratio) from Encapsula NanoSciences (stock 10 mM or 8 μg μl^{-1}) (CPC-610) were added and incubated at 37 °C for 1 h. After incubation, 67 μl of 3 M KCl was added to a final concentration of 1 M KCl and incubated at 37 °C for 15 min to minimize the nonspecific electrostatic association between proteins and lipids. Then 67% (w/v) sucrose in matching buffer was added to a final concentration of 40% in a final volume of 500 μl, mixed thoroughly, and underlaid in a buffer-matched, step gradient of 0.5 ml 5% and 10 ml of 25% (w/v) sucrose in an Open-Top Thinwall Ultra-Clear Centrifuge Tube (Beckman Coulter, 344060). Gradients were centrifuged at 281,000*g* for 75 min at 4 °C in an SW 40 Ti swinging bucket rotor (Beckman Coulter Optima XL-100K Ultracentrifuge). After centrifugation, each gradient was fractionated, from the top down, into 16 fractions of 700 μl. Samples were analysed by western blot.

Western blot analysis

All the fractions were diluted with 10× SDS–PAGE reducing sample buffer to a final concentration of 1× and denatured at 95 °C for 5 min. Samples were run along with either a eE2FL marker (std) or an Odyssey Protein Molecular Weight Marker (Li-Cor) (L) on 4–20% Bis-Tris precast gels (Bio-Rad). Proteins were then transferred to PVDF membranes using a Trans-Blot Turbo Transfer System (Bio-Rad). The membrane was blocked by Intercept (PBS) Blocking Buffer (LI-COR) for 1 h at 37 °C followed by incubating the blot with a 1:1,000 dilution of purified 2C1 mouse antibody (against HVR1 of HCV J6 E2 produced in the laboratory of A.G.) and incubated overnight at 4 °C. Primary antibody dilution was prepared in Odyssey Blocking Buffer in PBS with 0.1% Tween 20 (Sigma-Aldrich). The secondary antibody, IRDye 800CW Goat anti-Mouse IgG (Li-Cor), was used at a 1:10,000 dilution. The western blot was scanned using Li-Cor Odyssey software (v.3.0). Fluorescence signals of top fractions were background-corrected and measured by the Image Studio Lite software (v.5.2.5). The fluorescence intensities were exported in Excel format and histograms were prepared.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The coordinates and structure factors for eE2–2A12, tCD81-LEL and tCD81-LEL–eE2(ΔHVR1)–2A12 have been deposited into the RCSB PDB (https://www.rcsb.org) under accession numbers 7MWW, 7MWS and 7MWX, respectively.

30. Yost, S. A., Whidby, J., Khan, A. G., Wang, Y. & Marcotrigiano, J. Overcoming challenges of hepatitis C virus envelope glycoprotein production in mammalian cells. *Methods Mol. Biol.* **1911**, 305–316 (2019).
31. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
32. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
33. Emsley, P. & Crispin, M. Structural analysis of glycoproteins: building *N*-linked glycans with Coot. *Acta Crystallogr. D* **74**, 256–263 (2018).
34. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
35. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
36. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10037–10041 (2001).
37. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667 (2004).
38. DeLano, W. L. *The PyMOL molecular graphics system*. http://www.pymol.org (Schrödinger, 2002).
39. Keller, S. et al. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal. Chem.* **84**, 5066–5073 (2012).
40. Zhao, H., Piszczek, G. & Schuck, P. SEDPHAT—a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **76**, 137–148 (2015).

Acknowledgements We acknowledge A. Khan, M. Miller,M. Paskel and L. Tuberty for technical assistance; D. Wu and G. Piszczek for help with the ITC measurements; C. Rice for reagents and advice; and F. Jiang for his dedication to science and friendship. This work was supported by the Intramural Research Programs of the National Institute of Allergy and Infectious Diseases (J.I.C. and J.M.) and NIH grants R01AI136533, R01AI124680, R01AI126890 and U19AI159819 to A.G. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract numbers W-31-109-Eng-38 (SER-CAT) and DE-AC02-06CH11357 (LRL-CAT). SER-CAT is supported by its member institutions, and equipment grants (S10_RR25528 and S10_RR028976) from the National Institutes of Health.

Author contributions A.K., R.A.H., S.A.Y. and Y.W. purified the proteins and determined crystallization conditions. A.K., R.A.H., S.A.Y., W.B., A.D.D., J.I.C. and J.M. collected, processed and analysed the results. A.G. provided the antibody hybridoma. All authors helped to write and edit the manuscript.

Competing interests A.K., W.B., A.D., J.I.C, and J.M., are named as inventors on a patent application describing the data presented in this paper, which has been filed by the National Institutes of Health.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03913-5.

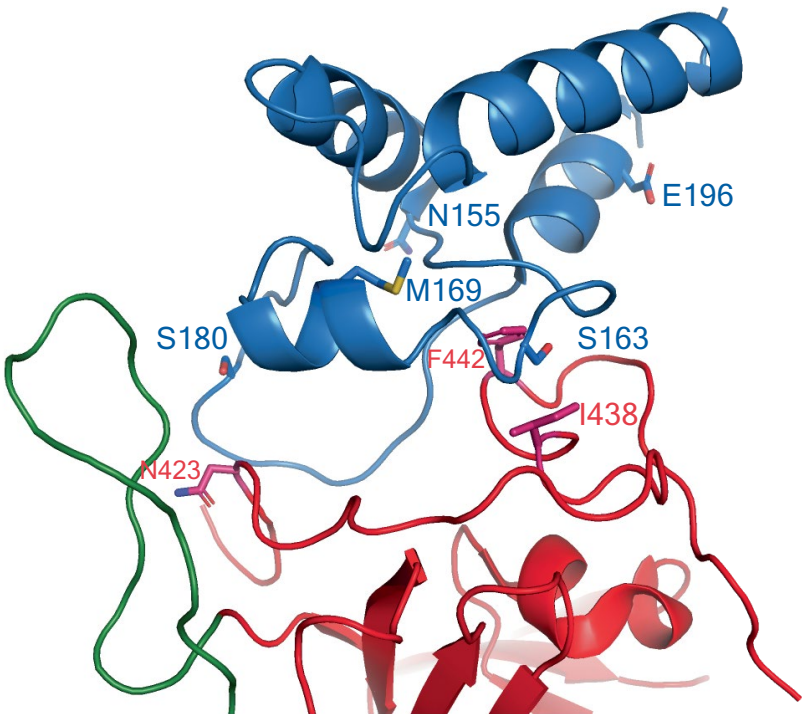
Correspondence and requests for materials should be addressed to Joseph Marcotrigiano.

Peer review information *Nature* thanks Yorgo Modis and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

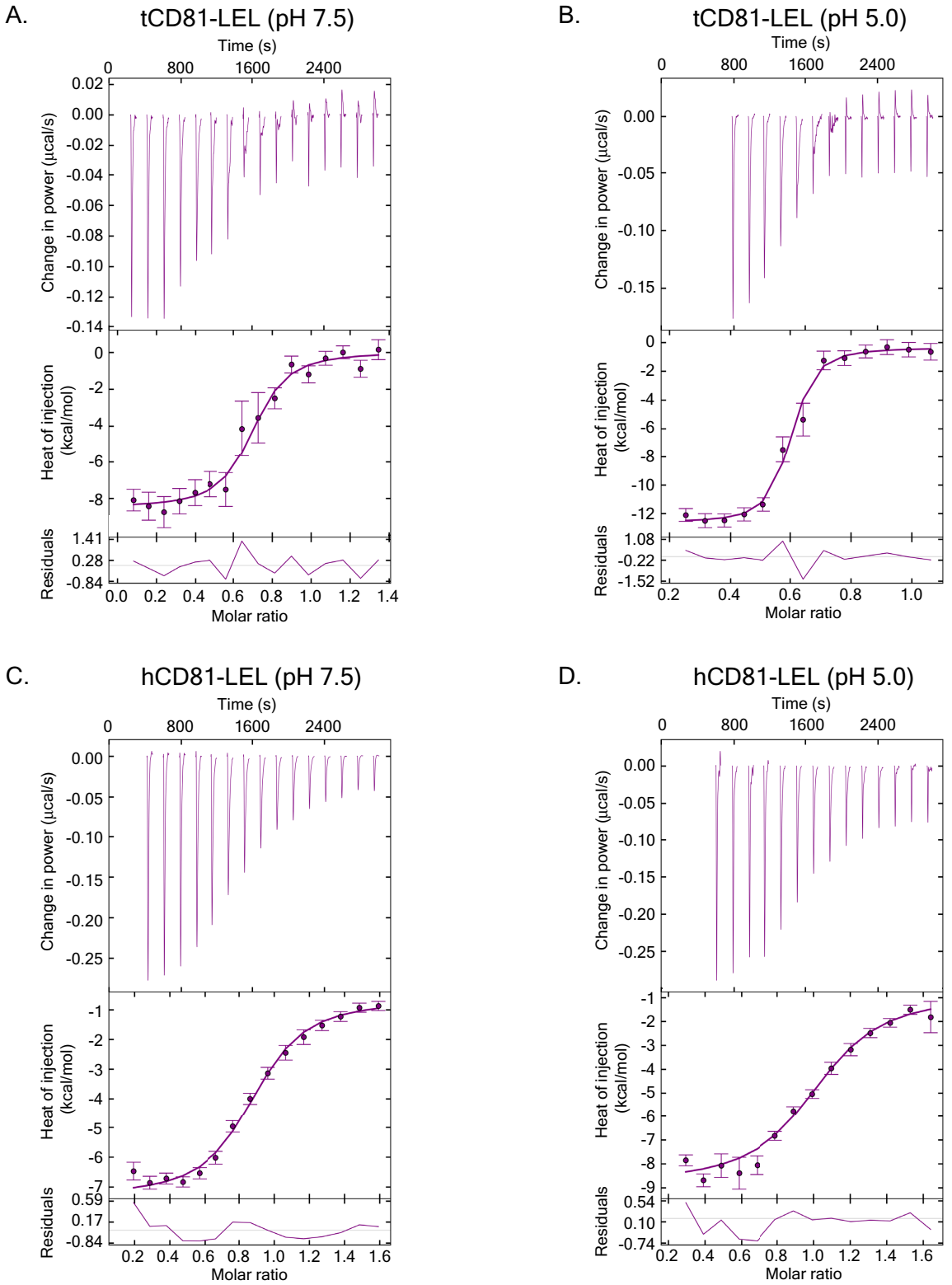
A

```
Human      MGVEGCTKCIKYLLFVFNFVFWLAGGVILGVALWLRHDPQTTNLLYLELGDKPAPNTFYVGIYILIAVGAVMMF
Tamarin    MGVEGCTKCIKYLLFVFNFVFWLAGGVILGVALWLRHDPQTTNLLYLELGDKPAPNTFYVGIYILIAVGAVMMF

Human      VGFLGCYGAIQESQCLLGTFFTCLVILFACEVAAGIWGFVNKDQIAKDVKQFYDQALQQAVVDDDANNAKAVVK
Tamarin    VGFLGCYGAIQESQCLLGTFFTCLVILFACEVAAGIWGFVNKDQIAKDVKQFYDQALQQAVVDDDANNAKAVVK

Human      TFHETLDCCGSSTLTALTTSVLKNNLCPSGSNIISNLFKEDCHQKIDDLFSGKLYLIGIAAIVVAVIMIFEMI
Tamarin    TFHETLNCCGSSTLSALTTSMLKNNLCPSGSSIISNLFKEDCHQKIDELFSGKLYLIGIAAIVVAVIMIFEMI

Human      LSMVLCCGIRNSSVYVP
Tamarin    LSMVLCCGIRNSSVY--
```

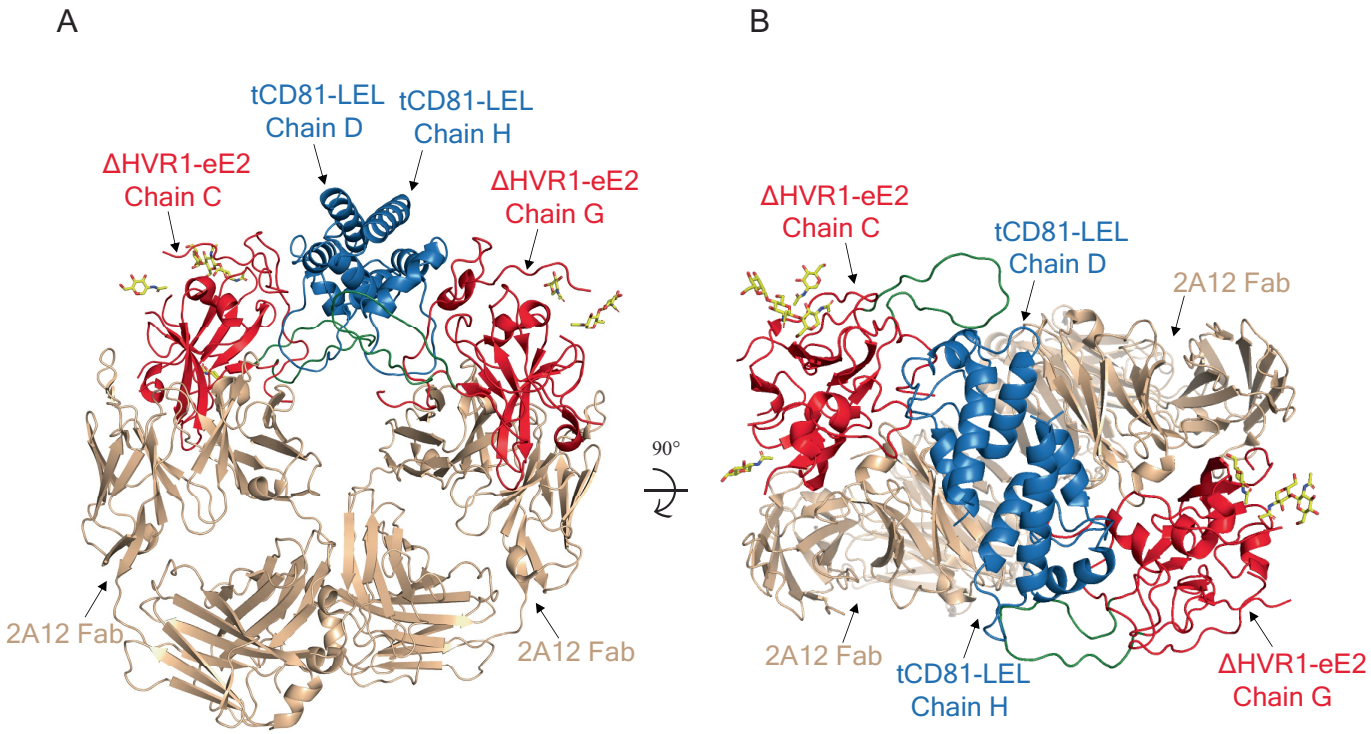
B



Extended Data Fig. 1 | Sequence divergence between human and tamarin CD81. a, Sequence alignment (light blue and black font) of full-length human and tamarin CD81 (Accession numbers: Human NM_004356, Tamarin CAB89875.1). The CD81-LEL (black font) has five divergent residues (green and yellow highlights represent nonidentical and similar amino acids, respectively). **b**, Ribbon diagram of tamarin CD81-LEL (blue) bound to eE2(ΔHVR1) (red and CD81-binding loop green) with side chains of the five, diverging CD81 residues (blue sticks) and proximal residues in eE2 (red sticks).

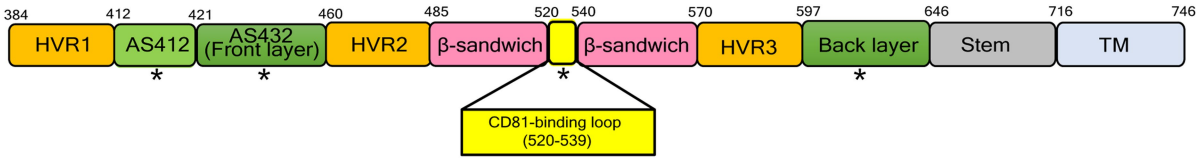


Extended Data Fig. 2 | Thermodynamic characterization of tamarin and human CD81-LEL interaction with eE2. a–**d**, ITC for the titration of tCD81-LEL (**a**, **b**) or hCD81-LEL (**c**, **d**) into eE2 at pH 7.5 (**a**, **c**) and pH 5.0 (**b**, **d**). Thermogram (upper panel), integrated heats and error bars (middle panel), and fit residuals (lower panel) are shown for each. The measurements were performed at 20 °C and analysed with an A + B ⇄ AB heterodimer model. Error bars indicate the error of peak integration over an interpolated baseline with a 68% (1 sigma) confidence interval. Residuals are the *y*-axis difference between the data point and the fitted curve in kcal mol^{-1}.

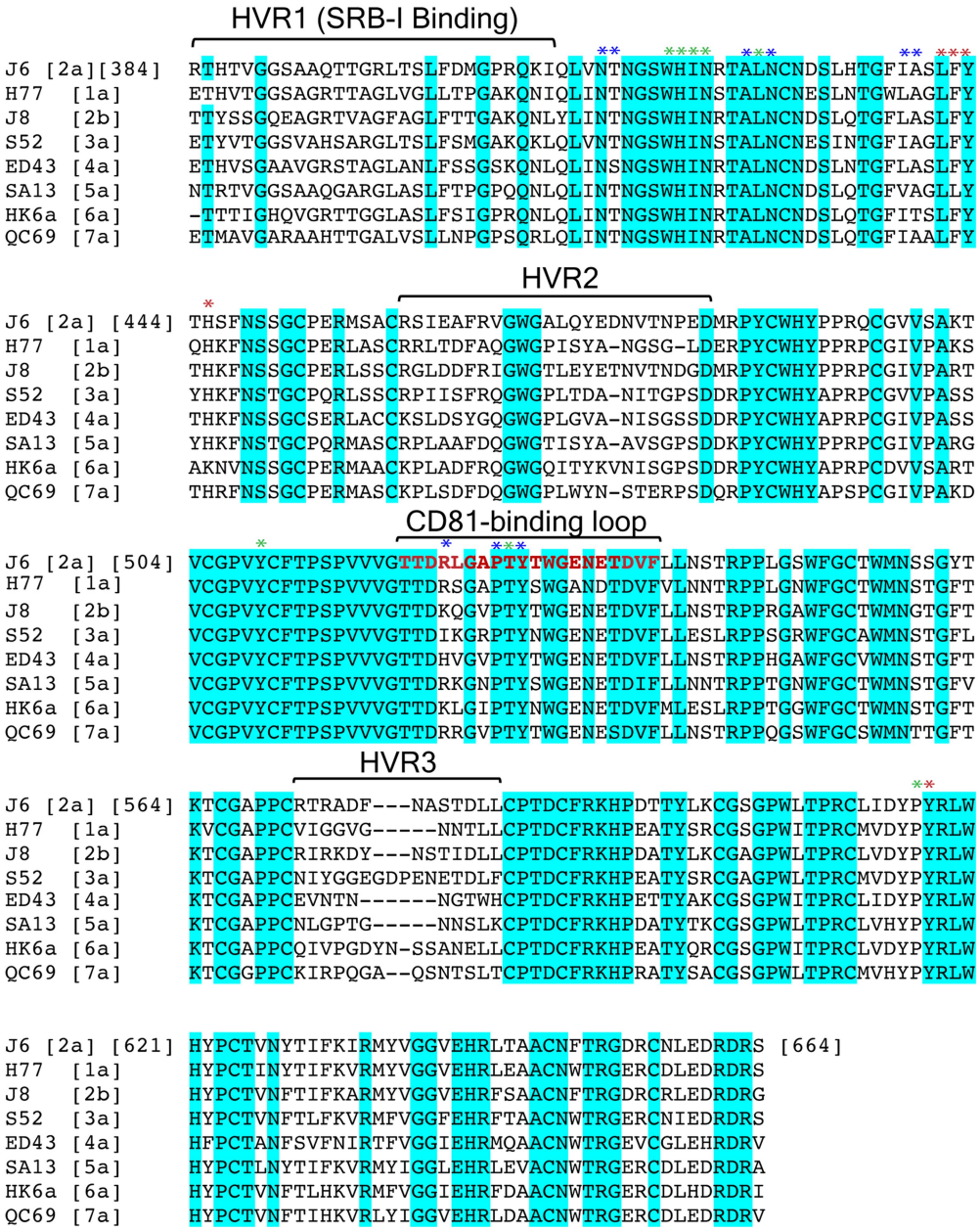


Extended Data Fig. 3 | The asymmetric unit for the tCD81-LEL–eE2(ΔHVR1)–2A12 complex. a, **b**, eE2(ΔHVR1) chains C and G (red with extended CD81-binding loop in green), tCD81-LEL chains D and H (blue), 2A12 (wheat) ribbon diagrams in the asymmetric unit of the tCD81-LEL–eE2(ΔHVR1)–2A12 complex from side (**a**) and top (**b**) views. The 90° axis of rotation is indicated. Carbohydrate moieties (yellow heteroatom sticks) are also shown.

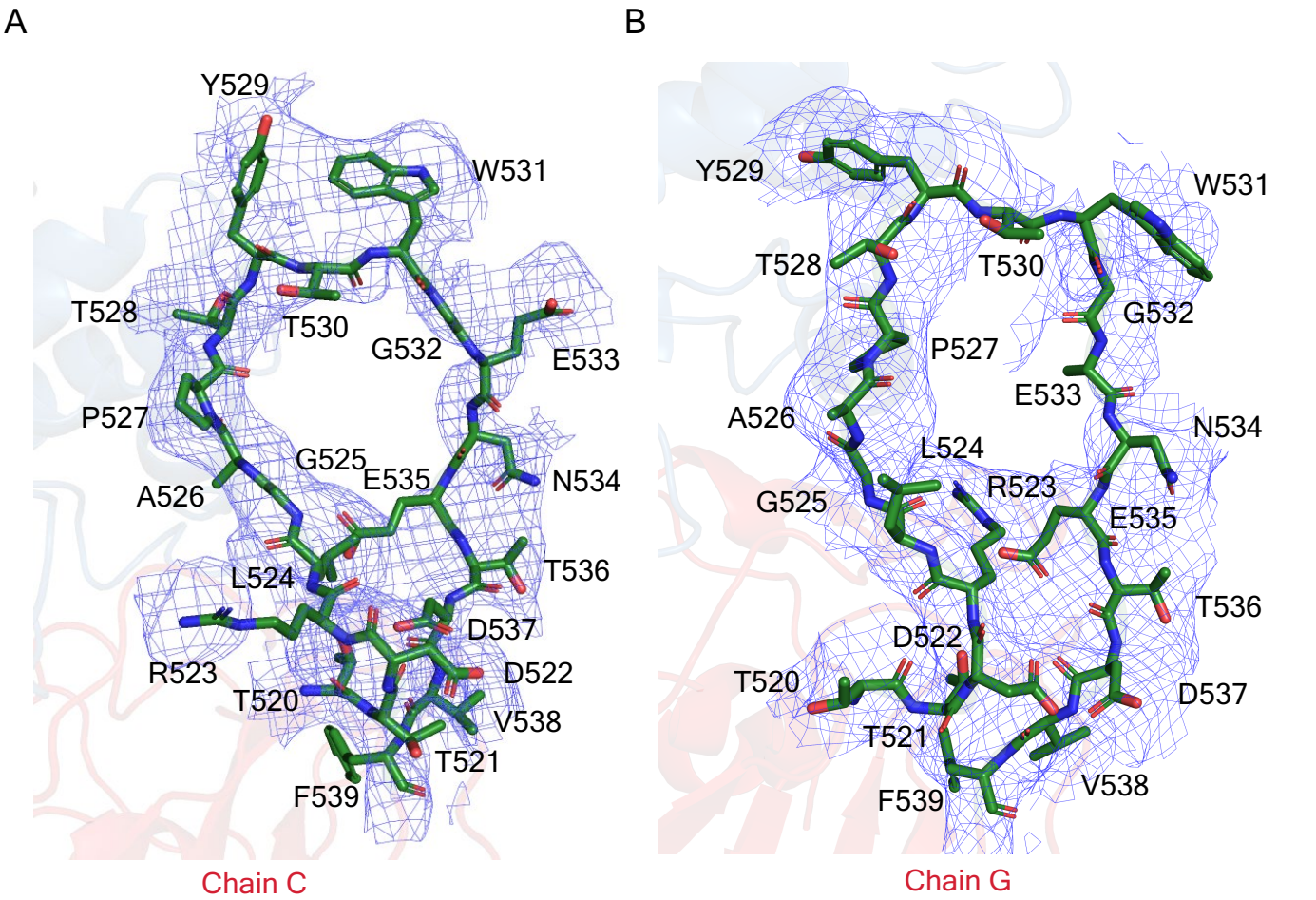
A



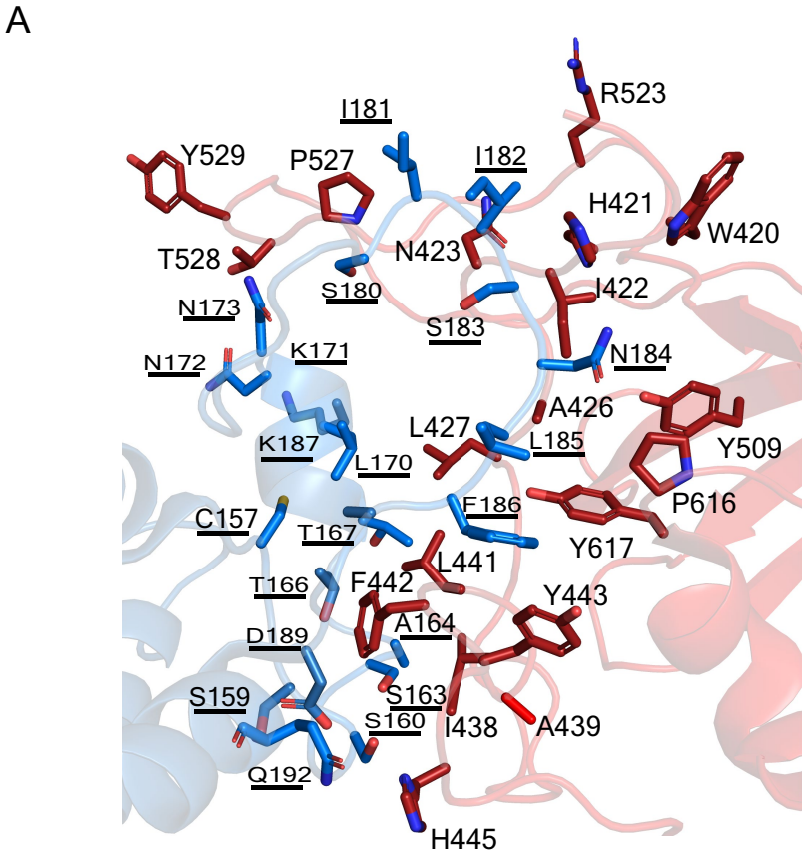
B



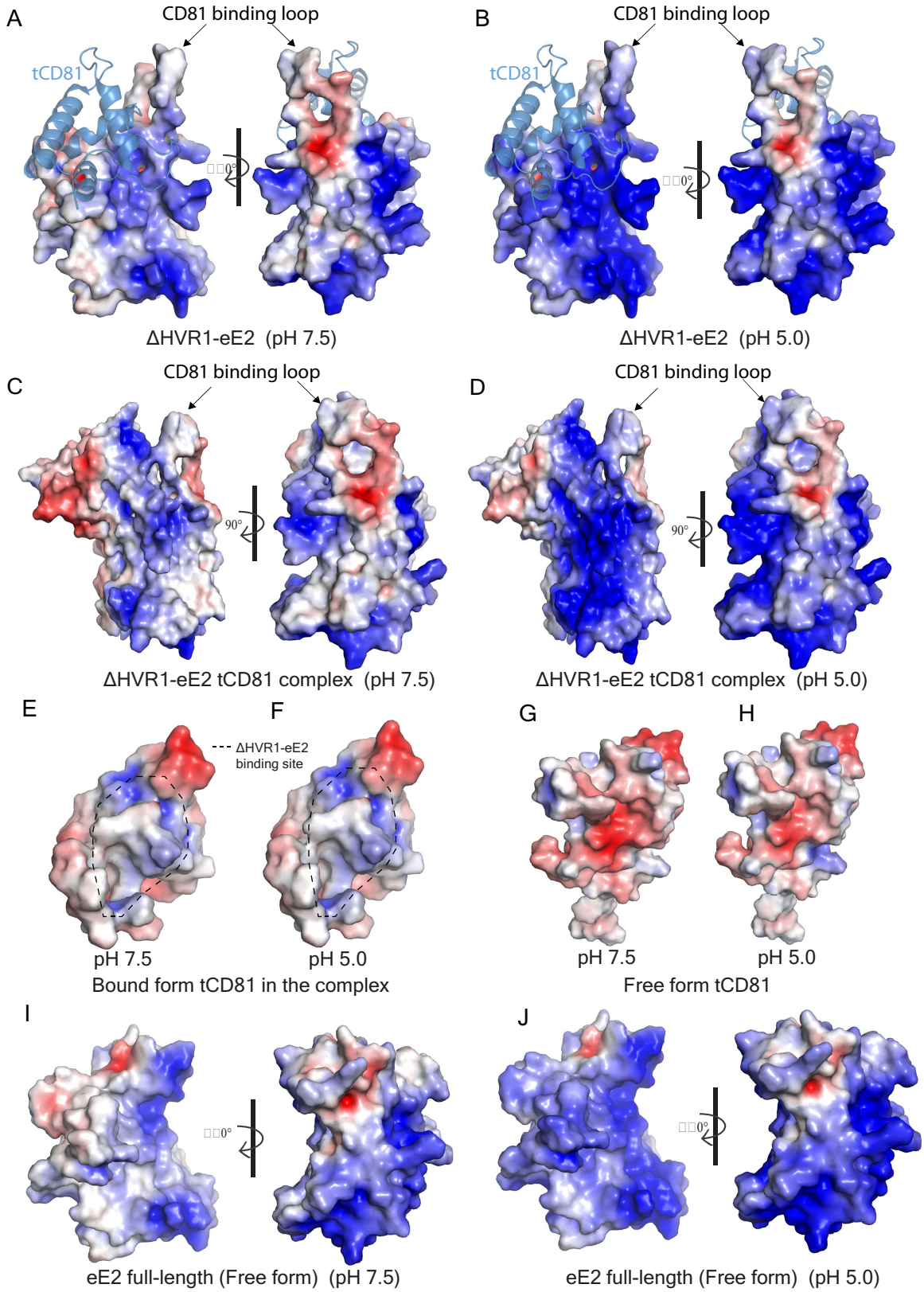
Extended Data Fig. 4 | Diagram and conservation of HCV E2. a, Schematic representation of the E2 protein with the CD81-binding loop highlighted in yellow and the asterisks highlighting regions associated with CD81 binding. **b**, Multiple sequence alignment of eE2 from representative strains (as labelled) of the seven genotypes. Conserved residues (cyan highlights) and CD81-binding loop residues (red font) are noted. Asterisks indicate residues ≤4Å from tamarin CD81 common to both chains C and G (red), chain G only (blue), and chain A only (green). Hypervariable region, antigenic site, and transmembrane are labelled HVR, AS and TM, respectively.



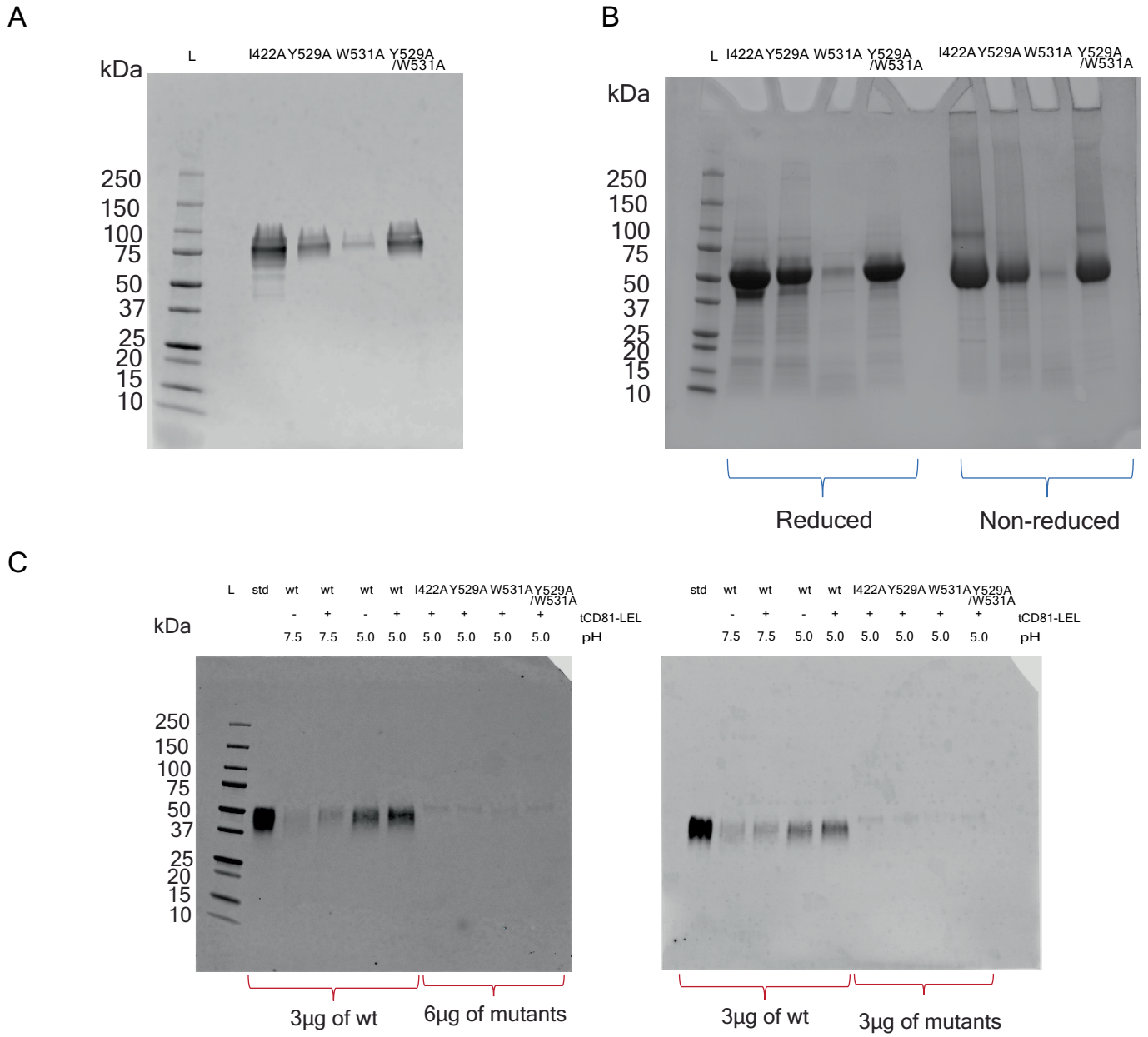
Extended Data Fig. 5 | A simulated-annealing $2F_o - F_c$ composite omit map for the eE2(ΔHVR1) CD81-binding loop in the X-ray crystal structure of the complex. a, **b**, CD81-binding loop in (**a**) Chain C and (**b**) Chain G (green heteroatom sticks), residues as labelled, in a 0.8σ contour level $2F_o - F_c$ composite omit map (blue mesh) calculated from the omission of residues 415–426 and 520–539, and packed against the tCD81-LEL (blue) and eE2(ΔHVR1) (red) ribbon diagrams.



Extended Data Fig. 6 | Interface between tCD81-LEL and eE2(ΔHVR1). Ribbon diagram of tCD81-LEL (blue) and eE2(ΔHVR1) (red) interface, chains C and D, with side chains (blue and red heteroatom sticks, respectively). The labels for tCD81-LEL residues are underlined.



Extended Data Fig. 7 | Electrostatic-potential surface maps of E2 and tCD81-LEL. a–j, Electrostatic-potential surface maps of eE2(ΔHVR1) in complex (**a**, **b**), tCD81-LEL–eE2(ΔHVR1) complex (**c**, **d**), tCD81-LEL in complex (**e**, **f**) and free form (**g**, **h**), and full-length eE2 free form (**i**, **j**). The surfaces are coloured by electrostatic potential corresponding to +5 kcal/(mol·e) (blue) and −5 kcal/(mol·e) (red) at 298 K calculated at pH 7.5 (**a**, **c**, **e**, **g**, **i**) and 5.0 (**b**, **d**, **f**, **h**, **j**) as labelled. Panels **a**, **b**, **i**, and **j** are depicted in the same orientation; panels **e**–**h** are depicted in the same orientation. **a**, **b**, tCD81-LEL is shown as a transparent, blue ribbon diagram. **e**, **f**, The eE2(ΔHVR1)-binding surface is outlined with a dotted line.



Extended Data Fig. 8 | Expression, purification and liposome flotation of eE2 mutants. a, E2-specific western blot of cell culture supernatants showing secreted protein levels of eE2 mutants I422A, Y529A, W531A, and double mutant Y529A/W531A. Expi293 GnTI$^-$ cells were transfected and supernatants (uncleaved eE2 protein) were mixed with reduced 2x sample buffer. 15 ul of supernatant was loaded in each well. E2 2C1 primary antibody was used for western blotting. **b**, Coomassie-stained 4-20% Bis-Tris SDS-PAGE gels of purified eE2 mutant proteins in the presence (Reduced) and absence (Non-reduced) of β-mercaptoethanol. **c**, E2-specific western blot of top fractions from liposome flotation assays, comparing increased loading (as labelled under each blot) of mutants. Protein molecular weight maker (L) and wild-type eE2 is provided as a marker (std). Sample pH, inclusion of tCD81-LEL, and eE2 mutant proteins are labelled.

Extended Data Table 1 | Affinity measurements of eE2 to human and tamarin CD81-LEL at neutral and low pH

Substrate	Ligand	pH	ΔH (kcal/mol)	K_d (nM)	N
eE2	tCD81	7.5	-8.52 (-11.21, -6.98)	175 (14, 870)	0.67 (0.57, 0.76)
	tCD81	5.0	-12.27 (-13.67, -11.12)	37 (8, 100)	0.57 (0.54, 0.61)
	hCD81	7.5	-6.65 (-8.09, -5.76)	773 (316, 1905)	0.86 (0.81, 0.94)
	hCD81	5.0	-7.89 (-10.10, -6.09)	681 (208,1698)	0.99 (0.89, 1.31)

Numbers in parentheses represent 95% confidence intervals.

Extended Data Table 2 | Data collection and refinement statistics for eE2–2A12, tCD81-LEL and tCD81-LEL–eE2(ΔHVR1)–2A12 complexes

	eE2/2A12	tCD81-LEL	tCD81-LEL/ΔHVR1-eE2/2A12
Data collection			
Space group	$C222_1$	I23	$P2_12_12_1$
Cell dimensions			
***a*, *b*, *c* (Å)**	95.84, 155.56, 129.42	113.26, 113.26, 113.26	76.95, 127.77, 212.37
(°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	33.34-2.71(2.81-2.71)	46.24-1.80 (1.83-1.80)	54.74-3.32 (3.5-3.32)
R_{sym}	0.11 (0.49)	0.18 (3.74)	0.28 (6.4)
$I / \sigma I$	6.3 (2.0)	14.2 (1.7)	6.8 (1.2)
Completeness (%)	99.01 (98.48)	100 (100)	99.5 (100)
Redundancy	1.98 (1.95)	40.1 (41.4)	13.2 (13.2)
Refinement			
Resolution (Å)	33.34-2.71 (2.81-2.71)	40.04-1.80 (1.87-1.79)	52.10-3.32 (3.4-3.32)
No. reflections	26376 (2583)	22508 (2237)	31547 (3090)
R_{work} / R_{free}	0.2183/0.2705	0.2072/0.2309	0.2413/0.2856
No. atoms			
Protein	606	193	1389
Ligand/ion	95	32	109
Water	102	66	0
***B*-factors**			
Protein	56.48	57.59	167.40
Ligand/ion	127.91	82.34	252.70
Water	34.0	60.86	0
R.m.s. deviations			
Bond lengths (Å)	0.005	0.014	0.002
Bond angles (°)	0.847	1.379	0.668
Ramachandran plot (%)			
Favored	92.11	96.83	91.87
Allowed	7.05	3.17	7.4
Outliers	0.84	0	0.73

A single crystal was used for each data collection. Values in parentheses are for the highest-resolution shell.

Extended Data Table 3 | Residues making interactions within ≤4 Å

	ΔHVR1-eE2 chain C	tCD81-LEL chain D
1	TRP 420	ASN 184
2	TRP 420	PHE 186
3	HIS 421	SER 183
4	ILE 422	SER 183
5	ILE 422	ILE 182
6	ASN 423	ILE 181
7	LEU 427	THR 167
8	LEU 441	THR 166
9	PHE 442	SER 159
10	PHE 442	SER 163
11	PHE 442	THR 166
12	PHE 442	CYS 157
13	PHE 442	PHE 186
14	PHE 442	LYS 187
15	TYR 443	PHE 186
16	HIS 445	SER 160
17	HIS 445	GLN 192
18	TYR 509	ASN 184
19	**THR 528**	**ASN 172**
20	**THR 528**	**ASN 173**
21	PRO 616	ASN 184
22	TYR 617	LEU 185
23	TYR 617	PHE 186

	ΔHVR1-eE2 chain G	tCD81-LEL chain H
1	ASN 415	SER 180
2	ASN 415	ILE 181
3	THR 416	SER 180
4	ALA 426	LYS 171
5	ASN 428	THR 167
6	ILE 438	SER 163
7	ILE 438	ALA 164
8	ALA 439	SER 163
9	LEU 441	PHE 186
10	PHE 442	THR 166
11	PHE 442	LEU 170
12	TYR 443	SER 160
13	TYR 443	ASP 189
14	HIS 445	SER 160
15	**ARG 523**	**LYS 171**
16	**PRO 527**	**LYS 171**
17	**PRO 527**	**ASN 172**
18	**TYR 529**	**ASN 172**
19	TYR 617	LEU 185
20	TYR 617	PHE 186
21	TYR 617	ASN 184

Amino acids in bold letters represent the CD81-binding loop, and those in red represent shared interactions in both chains of eE2(ΔHVR1).

nature research

Corresponding author(s): Joseph Marcotrigiano

Last updated by author(s): Jul 17, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The X-ray crystallographic data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID and Lilly Research Laboratories Collaborative Access Team (LRL-CAT) 31-ID beamlines at the Advanced Photon Source (APS), Argonne National Laboratory, using software developed for those beamlines at APS. The ITC data was collected on a MicroCal ITC-200 isothermal calorimeter, using commercial software (Malvern Panalytical, UK)
Data analysis	The X-ray crystallographic data were analyzed using publicly available software: CCP4 (www.ccp4.ac.uk) and Phenix (http://www.phenix-online.org). The ITC thermograms were integrated by NITPIC (Keller, et al., Anal. Chem. 2012) and fitted using SEDPHAT (Zhao, et al., Methods, 2015)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The coordinates and structure factors for eE2/2A12, tCD81-LEL, and tCD81-LEL/DeltaHVR1-eE2/2A12 have been deposited into the Protein Data Bank under accession numbers 7MWW, 7MWS, and 7MWX, respectively.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The eE2/2A12, tCD81-LEL, and tCD81-LEL/DHVR1-eE2/2A12 crystallographic data sets are 99, 100, and 99.5% complete with 2, 40, and 13 fold redundancy, respectively.
Data exclusions	No data was excluded
Replication	The structures were determined multiple times from different crystals. The manuscript reports the highest resolution example for each. The isothermal calorimetry data was collected multiple times, yielding similar thermodynamic values. The membrane flotation experiments were done multiple times and the results from two independent experiments are provided in Extended Data Figure 8.
Randomization	The experiments were not randomized as this is not relevant to structural biology studies
Blinding	Blinding is not relevant to this study as knowledge of the protein sequence is needed in order to accurately determine its structure

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	2A12 and 2C1 are anti HCV E2 from genotype J6 and were generated in Dr. Arash Grakoui's lab Emory University School of Medicine against purified, recombinant protein. Secondary antibody IRDye 800CW Goat anti-Mouse IgG was purchased from a commercial source (cat. no. 926-32210 from Li-Cor)
Validation	2C1 neutralizes HCV infection and binds to the HVR1 region of E2. 2A12 was shown to bind E2 by ELISA, isothermal calorimetry, size exclusion chromatography, and X-ray crystallography (Khan et al. Nature, 2014).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T GNTI- and 293expi GNTI- cells were purchased from ATCC (cat. no. CRL-3022) and ThermoFisher (Cat. No. A39240), respectively
Authentication	The cells line were not authenticated
Mycoplasma contamination	The cells were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	no commonly misidentified cell lines were used