ARTICLE IN PRESS

npj Vaccines

https://doi.org/10.1038/s41541-026-01449-1

Article in Press

Structural and genetic signatures of two classes of HCV E2 neutralizing face antibodies from non-human primates immunized with a recombinant E1E2

Received: 8 September 2025

Accepted: 8 April 2026

Cite this article as: Nguyen, Y.T., Chen, F., Giang, E. *et al.* Structural and genetic signatures of two classes of HCV E2 neutralizing face antibodies from non-human primates immunized with a recombinant E1E2. *npj Vaccines* (2026). https://doi.org/10.1038/s41541-026-01449-1

Yen T. K. Nguyen, Fang Chen, Erick Giang, Swati Saha, Lynn A. Ueno, Christopher Chen, Corey T. Watson, Netanel Tzarum, Ian A. Wilson, Mansun Law & Robyn L. Stanfield

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

Structural and genetic signatures of two classes of HCV E2 neutralizing face antibodies from non-human primates immunized with a recombinant E1E2

Yen T.K. Nguyen[1], Fang Chen[2], Erick Giang[2], Swati Saha[3], Lynn A. Ueno[1], Christopher Chen[4,†], Corey T. Watson[3], Netanel Tzarum[1,‡], Ian A. Wilson[1,5,*], Mansun Law[2,*], and Robyn L. Stanfield[1,*]

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
[2]Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037, USA.
[3]Department of Biochemistry and Molecular Genetics, University of Louisville School of Medicine, Louisville, KY 40202, USA.
[4]Southwest National Primate Research Centre (SNPRC) at Texas Biomedical Research Institute, San Antonio, TX 78245, USA.
[5]Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.
[†]Current Address: PKD Foundation, Kansas City MO 64197, USA.
[‡]Current address: Department of Biological Chemistry, Alexander Silberman Institute of Life Sciences, Faculty of Mathematics & Science, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel.
[*]To whom correspondence may be addressed. Email: wilson@scripps.edu (I.A.W.); mlaw@scripps.edu (M.L.); robyn@scripps.edu (R.L.S.)

ARTICLE IN PRESS

ARTICLE IN PRESS

Abstract
Hepatitis C continues to be a significant public health problem despite advancements in antiviral therapeutics. To eliminate this disease, an effective vaccine against new infections and re-infections is needed. However, to date only one Hepatitis C virus (HCV) envelope protein (E1E2) immunogen, developed by Chiron Inc., has been tested in a Phase I clinical trial (ClinicalTrials.gov identifier NCT00500747). To establish a benchmark for elicitation of broadly neutralizing antibodies (bnAbs) by E1E2, we previously immunized non-human primates (NHPs) with this immunogen and isolated monoclonal nAbs that exhibit neutralization potency comparable to human nAbs. Here we show that NHP nAbs, encoded by germline genes IGHV1-138*01 and IGHV4-NL_5*01 (homologs of human IGHV1-69*10 and IGHV4-59*12, respectively), recognize a relatively conserved E2 region (neutralizing face) proximal to antigenic region 3 (AR3). These NHP AR3-targeting nAbs share highly similar binding modes to human AR3-targeting nAbs, suggesting a similarity in human and NHP immune responses to the same HCV immunogen.

Introduction
Although anti-viral drugs to eliminate HCV infection are extremely effective [1], the number of HCV-infected individuals worldwide is still estimated to be over 50 million (World Health Organization, Hepatitis C: Fact sheet (2025), https://www.who.int/news-room/fact-sheets/detail/hepatitis-c). Despite treatment with direct-acting antiviral (DAA) drugs that can completely eradicate viral infection, some patients are developing liver cancer within 10 years of treatment [2]. Thus, elimination of HCV may not always be sufficient to protect against cancer and other chronic health problems. Additionally, viral clearance via DAA does not prevent future re-infection [3]. To prevent future infections and help eradicate HCV, an effective vaccine is needed [4].

A vaccine to prevent infection by the induction of bnAbs should target the viral envelope proteins E1 and/or E2 that have been shown by cryoEM to form a heterodimeric complex [5-7]. To gain cell entry, these proteins interact with multiple receptors on the cell surface, including heparan sulfate proteoglycan, low-density lipoprotein receptor, CD81, claudin, scavenger receptor-B1, and occludin [8-12]. The CD81 receptor binding site on E2 serves as a key target for numerous nAbs [13-18] and is part of the E2 neutralizing face that features highly conserved and hydrophobic sites, including antigenic site 412 (AS412, amino acid (aa) 412-420), AS434 (aa 434-446), and the discontinuous epitope antigenic region 3 (AR3) [17,18]. The AR3 comprises multiple structural domains, including the E2 front layer (FL, aa 421-459), CD81 binding loop (CD81BL, aa 518-535), and back layer (BL, aa 614-622) [17,18]. Of particular interest, AR3-targeting Abs exhibit broad neutralization activity against multiple HCV genotypes, suggesting that a broadly protective vaccine is possible [13,14,19,20]. Furthermore, crystal structures of E2 cores in complex with Fabs [21] or the CD81 long extracellular loop [22,23] have revealed that structural flexibility exists within AR3 regions, such as the FL and CD81BL. Whether this flexibility will aid, or hinder design of potential immunogens is not clear.

A large number of human bnAbs have been isolated from HCV-infected participants. Thus far, the majority of structurally characterized bnAbs utilize heavy chains (HCs) derived from the IGHV1-69 germline family and block binding to CD81 by recognition of E2 AR3 [16,20,24-26]; however, neutralizing antibodies (nAbs) that recognize other epitopes and are encoded by other germline families are also being isolated and studied [27]. The human IGHV1-69 germline gene is unique in

ARTICLE IN PRESS

that its codes for hydrophobic residues at the tip of CDRH2 that have been often observed to bind to hydrophobic surfaces on the HCV E2 antigen. Many IGHV1-69 encoded antibodies also target human viruses including influenza and HIV [25,28,29].

Until controlled human infection models are approved and available [30,31], the rhesus macaque (RM) is of interest as a model for antibody elicitation. While macaques cannot be infected with HCV, their immune system is highly similar to that of humans [32-37], and macaque Abs elicited by immunization can be compared to their human counterparts. However, despite the growing availability of NHP genomic databases [38,39], the RM IG loci remain poorly understood because of their genomic complexity and incomplete assembly [40-42]. The identification of individual germline genes for each tested NHP animal can ensure the accurate assignments of identified Abs to specific germline genes/alleles and estimation of SHM [38,42,43]. This process is critical for tracing Ab affinity maturation pathways in vaccinated NHP as a means to identify parallels in human-NHP genetics [36,37,44,45].

ARTICLE IN PRESS

Surprisingly, the only human vaccine trial carried out to date using a hepatitis envelope protein as immunogen was initiated in 2003 by Chiron (ClinicalTrials.gov identifier NCT00500747) [46], where a recombinant, full-length, genotype 1a E1E2 envelope protein was used as immunogen with MF59 as adjuvant. This vaccine advanced to a Phase I trial, but cross-neutralizing responses were produced in only a few vaccine recipients [47]. We previously obtained and analyzed serum samples from that trial [44]; however, no samples were available from which Abs could be isolated. Thus, to better understand this study, we previously immunized four RMs with the Chiron E1E2 immunogen (Fig. 1a) and found that both animals and humans mounted similar Ab responses, which are largely strain-specific at the serum level [44,45]. Despite this, cross-neutralizing Abs were isolated from RM plasmablasts and memory B cells, although their circulating levels were insufficient to confer broad serum neutralization [44,45]. Many of these nAbs target AR3 and are mostly derived from germline IGHV1-138*01_S6073 (KIMDB nomenclature [38]) that is homologous to the human IGHV1-69*10 gene (94.6% DNA sequence identity) [44,45] (Supplementary Fig. 1).

To gain deeper insight into the conserved properties of AR3-targeting nAbs, we have chosen five AR3-targeting nAbs isolated from RMs previously immunized with the Chiron E1E2 vaccine [44,45] for an in-depth analysis including crystal structure determinations of their complexes with E2. In this work we have also confirmed their verified germline sequences directly through targeted long-read IG loci genomic sequencing of the two study animals (#30734 and #31881) to verify the genetic features of these RM AR3-targeting class Abs. For four of these nAbs (RM1-73, RM11-48, RM10-30, and RM1-36) that utilize the same IGHV gene segment, IGHV1-138*01_S6073, with limited SHM we find that they exhibit a very similar E2 binding mode as Abs isolated from human elite neutralizers [20]. NAb RM5-16 is derived from a different germline gene, IGHV4-NL_5*01_S5158, and we show that it recognizes a similar epitope footprint but with a different binding approach angle to E2. These new results confirm that the NHP immune system can mimic that of humans in response to HCV immunogens but can also recognize the conserved AR3 epitope (within the neutralizing face) in different ways using different germline genes.

Results

Genetic features of vaccine-induced AR3-targeting Abs from RM

ARTICLE IN PRESS

Previously, we reported the isolation of 100 HCV-specific mAbs from two Chiron E1E2-immunized RMs: 56 from RM#3074 and 44 from RM#31881 (Fig. 1a) [44,45]. Of these, 24 mAbs exhibited neutralizing activity (15 from RM#3074 and 9 from RM#31884) [45]. Twelve of these nAbs (including RM1-73, RM11-48, and RM10-30) demonstrated potent cross-neutralizing activity (classified as bnAbs), 10 (including RM1-36 and RM5-16) showed limited breadth, and the remaining two were strain-specific [45]. In this work we present detailed characterization and structural analysis for five of the mAbs: RM1-73, RM11-48, RM10-30, RM1-36, and RM5-16.

Germline gene usage of five RM AR3-targeting nAbs (RM1-73, RM11-48, RM10-30, RM1-36, and RM5-16) studied was initially assigned using KIMDB [38] and IMGT [39] (Fig. 1b). More recently, we searched a data set derived from sequencing of 106 macaques (Macaque United Set of Alleles, MUSA) [48], which was compiled from combined genomic and repertoire sequencing. These data are searchable at: https://vdjbase.org/ reference_book/Rhesus_Macaque. Using the MUSA data and naming system, RM AR3-targeting nAb germline genes were assigned as IGHV1-BSGF*02 and IGHV4-5ARF*17 (Fig. 1c). The IGHV1-BSGF*02 and IGHV4-5ARF*17 germline alleles were detected in 14% (15/106) and 48% (51/106) of macaques in the MUSA group, respectively (Fig. 1d). None of the databases searched contained an IGHD germline allele aligning with RM10-30 while no corresponding IGHD germline gene was found for RM1-36 or IGHV germline gene for RM5-16 in the IMGT database (Fig. 1b,c).

To confirm the authentic germline alleles of these five RM nAbs, we next employed targeted single molecule real-time (SMRT; Pacific Biosciences) long-read sequencing from PBMC-derived genomic DNA of the same vaccinated RMs (#30734 and #31881) (Fig. 1a; see Methods for details). These data allowed us to generate haplotype-resolved assemblies spanning the IG loci for both animals (#30734 and #31881), from which we conducted comprehensive curation of their germline IG alleles using Digger [49]. Next, we performed multiple sequence alignment (MSA) of the identified Ab sequences and their corresponding best-matched germline alleles from previous databases, as well as the curated germline annotation of each animal. As a result, the direct targeted IG long-read sequencing and individualized germline set (IGS) confirmed the expected germlines allele for these five RM AR3-targeting nAbs. Specifically, RM1-73, RM10-30, RM11-48, and RM1-36 arise from IGHV4-2HU4*01 (identical to IGHV1-138*01_S6073), while RM5-16 is derived from IGHV6-2BVG*16 (identical to IGHV4-NL_5*01_S5158) (Fig. 1e, Supplementary Fig. 1a). RM #30734 appears homozygous to IGHV4-2HU4*01, whereas RM #31881 is also homozygous to IGHV6-2BVG*16.

IGS identified the IGHD germline alleles for RM1-73, RM11-48, RM1-36, and RM5-16, each showing 100% DNA identity to counterparts in the other databases (Fig. 1e, Supplementary Fig. 1b). Notably, the IG long-read sequencing data allowed us to find the closest putative IGHD gene for RM10-30 as IGHD1-C7RJ*01_i0gga_c1g_t2g_a3g_c5g_t6a_g8a_12g12, which showed a 6 nucleotide base pair (bp) match to the mature Ab sequence (Fig. 1f). Comparison with the putative IGHD germline allele thus shows conservation as well as variation (Fig. 1f); thus, whether it best represents the underlying germline gene for this Ab may require additional data analyses. In addition, IGS confirmed that the IGHJ germline alleles for these five RM AR3-targeting nAbs are also completely identical at the nucleotide level to corresponding germline alleles found in other databases (Supplementary Fig. 1c).

ARTICLE IN PRESS

Our RM germline sequences revealed consistently low SHM rates for the IGHV of these five RM AR3-targeting nAbs of 1.4-4% at the nucleotide level (Fig. 1e). The RM1-73 Ab has the lowest SHM (only 1.4%) and the longest CDRH3 of 25 aa (Fig. 1e,g). RM1-73, RM11-48, RM10-30, and RM1-36, which were all isolated from the same animal, use the same IGHV (Fig. 1e,g, Supplementary Fig. 1a). RM1-73 and RM11-48 also used the same IGHD gene, while RM11-48, RM10-30, and RM1-36 used the same IGHJ (Fig. 1e, Supplementary Fig. 1c). In contrast, these RM Abs used different IGHD and IGHJ genes from those of previously described RM2-01 and RM11-43 Abs [44,45]. RM5-16, isolated from a different animal, used distinct IGHV, D, J germline genes (Fig. 1e,h, Supplementary Fig. 1).

To compare rhesus and human homologs, we also performed MSAs of these IGHV alleles at both nucleotide and amino acid levels, incorporating their five closest human alleles to illustrate evolutionary similarity (Supplementary Fig. 1). From these alignments, we identified IGHV1-69*10 (94.6 % DNA identity and 90.8 % protein identity, L at codon 54) as the closest human counterpart to IGHV1-138*01_S6073, although differences to various IGHV1-69 alleles were small (0.7% at the DNA and 0 % at protein level). IGHV4-NL_5*01_S5158 aligned most closely with IGHV4-59*12 (93.6% DNA and 92.9 % protein identity), but again with small differences to other IGHV4-59 alleles (Supplementary Fig. 1e). To assess whether RM antibodies undergo maturation similar to human bnAbs, we aligned the human AR3-targeting IGHV1-69-encoded mAb1382 [20] with the IGHV1-69*10 and compared its SHM profiles to that of RM IGHV1-138*01 class Abs (Fig. 1g). This alignment revealed that both human and RM Abs retained germline framework regions across species and accumulated SHMs in CDRH1-2 and HFR3 regions (Fig. 1g). Henceforth, we will use the KIMDB germline identifiers for clarity and abbreviate IGHV1-138*01_S6073 to IGHV1-138*01, and IGHV4-NL_5*01_S5158 to IGHV4-NL_5*01.

Overall structure of RM IGHV1-138*01 class Abs and complexes with HK6a E2c3

We determined structures for the four RM IGHV1-138*01 class Fabs RM11-48, RM10-30, RM1-73, and RM1-36 in complex with the E2 core domain (E2c3) from HCV isolate HK6a, as well as an unliganded structure for Fab RM1-73 (Fig. 2, Supplementary Fig. 2, Supplementary Tables 1, 2). The complex structures, ranging from 2.6-2.9 Å resolution, revealed that all four RM Fabs recognize the canonical AR3 region of HK6a E2c3 (Fig. 2b), consistent with previous epitope mapping data [45]. The Fab binding to E2c3 arises mainly from the HC, which accounts for 85% (RM1-73), 65% (RM11-48), 79% (RM10-30), and 90% (RM1-36) of the total buried surface area (BSA) on the Fab (Fig. 2c, left panel). This HC domination is also observed in human anti-HCV IGHV1-69 class Abs and two previously reported RM IGHV1-138*01 class Abs RM2-01 and RM11-43 (Fig. 2c, left panel) [45]. Most of the interactions are from CDRH3, which contributes 45-59% of the total BSA of the HC (Fig. 2c, right panel).

Epitope analysis revealed all four RM Fabs recognize largely overlapping epitopes at the central hydrophobic groove within the AR3 region, contacting a cluster of hydrophobic residues including the E2 FL N-terminus (I422, T425, and L427), α1 helix (I438, T439, L441, and F442), C-terminus (Y443 and A444), β-sandwich (P505), CD81BL (W529), and BL (P612 and Y613) (Fig. 3, left panel). Among these residues, T425, L427, L441, Y443, P505, W529, P612, and Y613 are highly conserved across 6 HCV genotypes (Supplementary Fig. 3), suggesting all RM Abs recognize a conserved hydrophobic core of E2. RM11-48 engages additional hydrophobic residues (V447,

ARTICLE IN PRESS

V622, and L626) adjacent to the core hydrophobic groove, whereas RM1-73 uniquely contacts the AS412 region (N417-W420) (Fig. 3, left panel).

RM1-73: Fab RM1-73 undergoes a large conformational change upon binding E2, where CDRH3 changes from a loop conformation in the unliganded Fab to a helical conformation when bound to E2 (Supplementary Fig. 2a), with a Cα root-mean-square deviation (RMSD) of 5.3 Å for CDRH3 residues C92-W103 in bound and unbound Fab structures. In the E2-RM1-73 complex structure, the RM1-73 HC and LC both interact with E2 (Figs. 2c, 3a), burying 972 $Å^2$ of surface on the E2 (821 $Å^2$ for HC and 151 $Å^2$ for LC). CDRH1 and CDRH2 form many hydrophobic interactions with E2 core hydrophobic groove residues I422, L427, I438, T439, F442, Y443, and W529 (Fig. 3a). Whereas CDRH2 of previously characterized Abs RM2-01 interacts only with the C-terminus of the FL α1 helix (aa 439-443) within the core hydrophobic groove [45], CDRH2 of RM1-73 retains this footprint and makes additional contact towards the FL N-terminus L427 and C-terminus K445 (Fig. 3a). Two hydrogen bonds (H-bonds) are formed between CDRH2 E50 and N58 with E2 Y443 and K445, respectively (Fig. 3a, Supplementary Table 3). Compared to CDRH1-2, the CDRH3 loop binds lower in the core hydrophobic groove and extends toward AS412 (G418, S419, and W420), FL C-terminal tail A444, β-sandwich P505, and the BL (P612 and Y613) (Fig. 3a, Supplementary Table 4). CDRH3 K98 H-bonds with E2 F442 and P612, while CDRH3 L99 H-bonds with Y613 (Fig. 3a, Supplementary Table 3). CDRH3 A100c backbone H-bonds with conserved AS412 S419 and W420 (Fig. 3a, Supplementary Fig. 3, Supplementary Table 3). The only contacts from the LC arise from CDRs L1 and L3 (Fig. 3a, Supplementary Fig. 4, Supplementary Tables 3,4), where CDRL1 Y32 contacts E2 FL A444 and CDRL3 D93/T94 interact with the FL helix α1 Y443, K445, and N446.

RM11-48: RM11-48 binds E2 with a BSA of 968 $Å^2$ (632 $Å^2$ for HC and 336 $Å^2$ for LC) (Fig. 2c). The CDRH1-2 loops of RM11-48 bind E2, recognizing a similar epitope as those of RM1-73 (Fig. 3b). In addition, HFR1 residue T28 is within vdW contact distance of the N423 glycan (Supplementary Fig. 5, Supplementary Table 3). Like RM1-73, RM11-48 CDRH3 binds into the AR3 hydrophobic groove (Fig. 3b, Supplementary Table 4). CDRH2 E50 H-bonds with E2 Y443 and CDRH3 R98 engages in a H-bond network with E2 FL S440 and BL P612; I99 H-bonds to Y613, and N100f H-bonds to F442 (Fig. 3b, Supplementary Table 3). Among the four RM Fabs, the LC of RM11-48 forms the most extensive interactions with E2 (Fig. 3b, Supplementary Tables 3, 4). CDRL1 F27d and F28, as well as CDRL3 N93, P95, and Y96 create a hydrophobic cluster that interacts with another hydrophobic patch formed by the E2 FL C-terminus (Y443 and A444) and BL C-terminus (L615 and V622) (Fig. 3b, Supplementary Fig. 4, Supplementary Table 4). CDRL3 K92 and N93 form H-bonds with E2 A444 and N446 (Fig. 3b, Supplementary Fig. 4, Supplementary Table 3).

RM10-30: RM10-30 recognizes E2 with 889 $Å^2$ of BSA (706 $Å^2$ for HC and 182 $Å^2$ for LC) (Fig. 2c). RM10-30 CDRH1-2 bind into a slightly wider E2 AR3 hydrophobic patch as compared to that bound by RM11-48 (Fig. 3c, Supplementary Table 4). HFR1 T28 is within H-bonding distance of the N423 glycan (Supplementary Fig. 5, Supplementary Table 3). CDRH2 E50 also H-bonds with E2 helix α1 Y443, while R53 forms a salt bridge with CD81BL E531 and a π-cation interaction with W529 (Fig. 3c, Supplementary Table 3). CDRH3 L100a binds in the lower hydrophobic pocket bordered by E2 I422, L441, F442, and Y613 (Fig. 3c, Supplementary Table 4) while CDRH3 R100b and R100d form several H-bonds with E2 FL helix α1 S440, L441, F442, and

ARTICLE IN PRESS

Y443 (Fig. 3c, Supplementary Tables 3, 4). CDRH3 R100d forms a π-cation interaction with E2 Y443 (Fig. 3c, Supplementary Table 4). CDRL3 is involved in hydrophobic interactions with E2 FL helix α1 residues Y443 and A444 while CDRL1 E27 forms a charged interaction with E2 C-terminal residue K445 (Fig. 3c, Supplementary Fig. 4, Supplementary Tables 3, 4). The LC framework LFR1 I2 makes a van der Waals (vdW) interaction with E2 K445 (Fig. 3c, Supplementary Fig. 4, Supplementary Table 4).

RM1-36: RM1-36 buries 839 $Å^2$ of surface area (758 $Å^2$ for HC and 81 $Å^2$ for LC) in its complex with E2 (Fig. 2c). CDRH1 makes hydrophobic interactions with E2 FL helix F442 and CD81BL W529 (Fig. 3d, Supplementary Table 4). CDRH2 binds to the same hydrophobic patch as CDRH2 from RM10-30 (Fig. 3d, Supplementary Table 4). As for the other three RM IGHV1-138*01 Fabs, CDRH2 E50 forms a H-bond with E2 α1 helix Y443 (Fig. 3d, Supplementary Table 3). CDRH2 R53 engages in polar and charged interactions with E2 FL N428 and CD81BL E531, H-bonds with FL L427/N428 and CD81BL W529 (Fig. 3d, Supplementary Tables 3,4). CDRH3 also occupies the AR3 core hydrophobic groove (Fig. 3d, Supplementary Table 4). CDRH3 S99 forms several H-bonds with E2 FL S440, L441, Y443, and A444, while Y98 H-bonds with FL F442 and BL Y613, and E100 also H-bonds to BL Y613 (Fig. 3d, Supplementary Table 3). Additionally, CDRH3 Y100a H-bonds with E2 FL N-terminus H421 (Fig. 3d, Supplementary Table 3). The RM1-36 LC has little contact with E2, with only CDRL3 Y96 forming a hydrophobic interaction with FL Y443 (Fig. 3d, Supplementary Fig. 4).

Characterization of RM IGHV4-NL_5*01 class Ab RM5-16 and structure with HCV-1 E2ecto

To investigate the structural basis of RM IGHV4-NL_5*01 class Ab recognition, crystal structures of RM5-16 Fab, unliganded and with the HCV-1 strain E2 ectodomain (E2ecto, aa 384-645), were determined at 1.46 and 3.39 Å resolutions, respectively (Figs. 2d, Supplementary Fig. 2b, Supplementary Tables 1, 2). RM5-16 also recognizes AR3 of E2ecto (Fig. 2d), consistent with previous epitope mapping [45]. The overall RM5-16 structures for unliganded and ligand-bound RM5-16 are highly similar, with a Cα RMSD of 0.2 Å calculated for variable domain (residues variable VH 1-113 and VL 1-110). Unlike RM1-73, there is essentially no difference in the CDR regions of regions of bound versus unbound RM5-16, with RMSD of only 0.1 Å for CDRH3 (residues C92-W103) (Supplementary Fig. 2b). Unlike the four RM IGHV1-138*01 class Abs, the HC and LC of RM5-16 contribute more equally to E2 binding, accounting for 54% (368 $Å^2$) and 46% (317 $Å^2$) of the total BSA on E2ecto, respectively (Fig. 2e). CDRH3 of RM5-16 remains the largest contributor to E2ecto binding (57% of total HC BSA) (Fig. 2e). The epitope of RM5-16 overlaps to some extent with those of RM IGHV1-138*01 class Abs but also includes contacts to the top of the E2ecto FL loop (aa 431-439) and CD81BL (aa 528-529) (Fig. 3e). Although RM5-16 has an N-linked glycosylation sequon at HC N81 we observe no electron density for a glycan at that position in either unliganded or ligand-bound RM5-16 structures.

The resolution of this complex structure is relatively low (3.39 Å), limiting precise H-bond assignment; however, the high-resolution Fab structure (1.46 Å) and previously determined higher resolution E2ecto structures aided as templates during our model building. H-bonds described here are within H-bonding distance in our model. Within the HC, the RM5-16 CDRH1-2 loops contact only E2 CD81BL (Fig. 3e). CDRH1 Y33 is within H-bonding distance of E2 E531 (Fig. 3e, Supplementary Table 3). CDRH2 is buried in a hydrophobic region of CD81BL formed by E2

ARTICLE IN PRESS

conserved S528, W529, E531, and N532 (Fig. 3e, Supplementary Fig. 3, Supplementary Table 4). The side chain of T54 H-bonds with E2 E531 (Fig. 3e, Supplementary Table 3). R56 makes H-bond and charged interactions with E2 E531 and the main-chain carbonyl of G53 is within H-bonding distance of the N532 glycan (Fig. 3e, Supplementary Fig. 5 and Supplementary Table 3). CDRH3 W100 is positioned adjacent to a hydrophobic cluster in E2 FL L427, C429, L438, and F442, and forms an H-bond with the main-chain carbonyl of W420 (Fig. 3e, Supplementary Tables 3, 4). CDRH3 G97, S98, and W100 are involved in a network of vdW contacts with E2 FL L438 and CD81BL W529 (Fig. 3e, Supplementary Table 4). The RM5-16 LC contacts E2 via all three CDRLs (Fig. 3e, Supplementary Fig. 4). CDRL1 hydrophobic residues also bind into the hydrophobic pocket formed by E2 FL L438, A439, and F442 (Fig. 3e, Supplementary Table 4). In LFR3, K53 forms a H-bond with E2 H421 while CDRL3 N93 and S94 form H-bonds to E2 FL N434 (Fig. 3e, Supplementary Table 3).

Neutralization and binding breadth of RM AR3-targeting Fabs for HCV E2

We evaluated the neutralization breadth of all RM antibodies against five selected HCVpp strains spanning antibody resistance tiers 1–4, together with the autologous immunizing strain HCV-1 (Fig. 2f). All five RM antibodies effectively neutralized HCV-1. RM1-73 displayed the broadest activity, cross-neutralizing the four heterologous tier 1–3 strains, with breadth exceeding that of human reference bnAb AR3A. RM11-48 also neutralized three of the four tier 1–3 viruses. In contrast, RM10-30, RM1-36 and RM5-16 exhibited limited activity, neutralizing only one of the tier 1-3 viruses. None of the RM antibodies neutralized the tier 4 virus.

Previous studies using BLI to characterize the binding of RM1-73 and RM11-48 IgGs to HCV E2 showed the two Abs bound strongly to all tested isolates with nM K_d values [45]. To avoid avidity issues of the IgG format, in this study we tested binding of E2 with five RM Fabs. Our BLI data show that Fabs RM1-73, RM11-48, and RM10-30 bind to E2 proteins from different isolates and genotypes including H77 (1a), HCV-1 (1a), J6 (2a), S52 (3a), and HK6a (6a) in the nM range (Fig. 2g, Supplementary Table 5). Fabs RM1-73 and RM10-30 generally showed stronger binding than RM11-48, RM1-36, and RM5-16 (Fig. 2g, Supplementary Table 5). Consistent with our previous study using IgGs, Fab RM1-73 also exhibited enhanced binding to E2 compared to Fab RM11-48 [45]. Fab RM5-16 showed nM binding to all tested strains except H77 (1a), aligning with previous neutralization data (Fig. 2f,g) [45].

The E2 epitopes recognized by the RM antibodies include 22-28 residues and/or glycans, and as epitope sequence conservation increases, it is more likely that an antibody to that epitope will be broadly neutralizing, although neutralization can also be affected by conformational changes (such as in the FL[21]) and by residues far from the actual epitope [50]. In order to help rationalize the neutralization profiles (Fig. 2f) for the RM antibodies and control human bnAb AR3A, we analyzed each epitope position for sequence variability. The epitope residues (defined by the Molecular Surface package MS [51]) are listed in Supplementary Table 7, followed by the amino acid at that position in the crystal structure and then by the amino acid types found at the position in the 6 viral strains used for the neutralization assay (in the order: HCV-1, 5.2.1, H77, 1a123, 4.2.2, and 3.1.2). Also listed for each epitope position are the buried molecular surface area for that residue, % of the total buried surface area, and Wu-Kabat variability [52]. Wu-Kabat variability is 1.0 for a completely conserved residue, and 400 for a position where all 20 amino acids appear at equal frequencies and was calculated for each position using 5609 E2 sequences extracted from

ARTICLE IN PRESS

the BV-BRC [53]. While most epitope residues are well conserved among the 6 isolates tested (Fig. 2f), there are several positions with high variability that may be involved in neutralization resistance or enhancement. Residue 531 (E,S,A,E,E,G in the HCV-1, 5.2.1, H77, 1a123, 4.2.2, and 3.1.2 isolates, respectively) is found in the epitopes of all Abs excepting RM1-73 (Fig. 3, Supplementary Table 3), and hydrogen bonds or charged interactions involving this residue are found in all Abs excepting RM1-73 and RM11-48. Glu is found at position 531 in about 65% of all HCV E2 sequences extracted from the BV-BRC and is the residue at this position in the immunizing HCV-1 strain. None of the Abs tested here neutralized the Tier 4, 3.1.2 isolate that has a Gly at 531, and only two Abs (RM1-73 and AR3A) neutralized the H77 isolate that has Ala at 531, while 3 Abs (RM1-73, RM11-40 and AR3A) neutralized the 5.2.1 strain with Ser at 531. Consistent with neutralization profiles, we observed strong binding for RM1-73 to the H77 E2, but reduced/no binding of RM11-48, RM10-30, RM1-36, and RM5-16 to H77 E2 (Fig. 2g, Supplementary Table 5). Another interesting residue is the highly variable FL residue 444 (H,F,Q,S,H,Y in the HCV-1, 5.2.1, H77, 1a123, 4.2.2, and 3.1.2 isolates, respectively) that is part of all epitopes excepting those recognized by AR3A and RM5-16. Residue 444 is the most variable residue in the FL region and can be occupied by residues of different sizes and properties. There are no hydrogen bonds to the 444 side chain (Supplementary Table 3) in any of our structures but the high variability in size and charge at this position may influence neutralization by antibodies when it is part of their epitopes. For example, the immunizing HCV-1 strain has a His at this position, and the Tyr found in the 3.1.2 strain (not neutralized by any antibodies, Fig. 2f) may be difficult to accommodate due to its larger size. Another FL residue, 431 is Asp or Glu in all our tested isolates except the 3.1.2 isolate where it is a smaller and uncharged Ala. This residue is found in the epitopes for AR3A, RM10-30 and RM5-16 and may contribute to the resistance of 3.1.2 to these antibodies, although it makes hydrogen bonds to E2 only in the AR3A structure. One other notable position is 430 that is a highly conserved glycosylation site. In our sequence data from the BV-BRC, almost 97% of the isolates have a glycan at this position. However, the 5.2.1 isolate has a 430Q that eliminates that glycan. It is possible that removal of this glycan will allow easier antibody access to surrounding epitopes, and thus enhance neutralization in most cases. Although not making close contacts to any antibodies, residue 430 and its glycan are buried in the epitopes for Abs AR3A, RM1-36, and RM5-16. While AR3A neutralizes 5.2.1, RM5-16 and RM1-36 show no neutralization for this Tier 1 isolate. Isolate 5.2.1 also lacks a moderately conserved glycan at position 476 (~40-75% conservation; estimating the exact level of conservation for this position is difficult due to insertions and deletions in the VR2 region confusing automated sequence alignments) and thus has only 9 N-linked E2 glycosylation sequons rather than the more commonly found 11 sequons [54].

RM AR3-targeting Abs mimic neutralizing Abs isolated from human elite neutralizers

Previously described human elite neutralizer mAbs 1198 and 1382 share similar binding modes to RM Fabs RM2-01 and RM11-43 [20]. To further understand differences/similarities in how these five RM Abs bind to E2, we compared RM Ab-E2 structures to previously reported AR3-targeting RM and human IGHV1-69 class Abs. All five RM Abs target the same AR3 region as the previously reported RM2-01, RM11-43, as well as human bnAbs AR3C, HEPC74, mAbs 1198, and 1382 and they share a heavy-chain–dominant mode of approach toward AR3 (Fig. 4a and Supplementary Fig. 6) [20,26,45,55]. The RM Abs and human elite neutralizer mAb1198 share similar overall binding poses with mAb1382 but differ in their precise approach angles relative to mAb1382 (Supplementary Fig. 6). These approach angles are defined here between the center of

mass of the respective Fabs and E2, ranging from counterclockwise (up to 26°) to clockwise (up to 15°) (Supplementary Fig. 6). In contrast, all of these Fabs adopts markedly different approach modes compared to human bnAbs AR3C and HEPC74 (Supplementary Fig. 6). HEPC74 shows an inverted HC–LC positioning relative to the RM Abs and mAb1382, while AR3C engages the epitope with an orthogonal orientation (~90° relative to the RM Abs), positioning its HC toward the AR3 epitope face (Supplementary Fig. 6).

RM1-36 closely matches the approach angle of human mAb1382 (Fig. 4a and Supplementary Fig. 7a). Similar to RM2-01 and RM11-43, RM1-73, RM11-48, and RM10-30 rotate away from mAb1382 toward the BL at counterclockwise angles of 22°, 26°, and 18°, respectively (Supplementary Fig. 6a). By contrast, RM5-16 rotates away from mAb1382 towards the E2 CD81BL at a clockwise angle of 15° (Supplementary Fig. 6a). As the CDRH3 dominates the binding interactions of AR3-targeting Abs (Fig. 2c), we also performed structural alignments between the CDRH3 regions of RM Abs compared to human mAb1382 (Supplementary Fig. 6b). Similarly, the alignments revealed CDRH3 of RM1-36 has the same approach angle as that of mAb1382, while CDRH3s from other Abs have different binding approach angles in comparison with mAb1382 (Supplementary Fig. 6b).

When we compared the position of CDRH1-3 loops of these RM IGHV1-138*01 class Abs upon interacting with E2 protein, we found their CDRH3 loops contact the BL and central β-sandwich and lower FL region, extending toward the BL, similar to the CDRH3 loops of mAbs 1198 and 1382 (Fig. 4b) [20,45]. In contrast, CDRH3 loops in human AR3C and HEPC74 contact the FL and CD81BL regions of E2 (Fig. 4b) [26,55]. For IGHV4-NL_5*01 RM5-16, the CDRH1-2 loops shift towards the E2 CD81BL center and lower FL region (Fig. 4b) with the CDRH3 tip of RM5-16 exhibiting a binding mode more like those of HEPC74 and AR3C, as it is centrally located in a cleft formed by the E2 FL, CD81BL, and BL, rather than projecting toward the BL (Fig. 4b). The CDRL1-3 loops of RM5-16 lie in the middle of the FL region of E2, unlike other RM IGHV1-138*01 class Abs and human IGHV1-69 class Abs (Fig. 4c).

When we compared the CDRH1 loops of all RM AR3-targeting Abs, we observed that while RM1-73, RM2-01, and RM1-36 only contact E2 residues from the E2 FL N-terminus and CD81BL, the those of RM11-48, RM10-30, and RM11-43 extend their contacts to residues from the helix α1 C-terminus of E2 (Fig. 4d). As previously mentioned, the CDRH1 loop of RM5-16 only contacts E2 CD81BL (Fig. 4d). Meanwhile, the CDRH1 loops of RM11-48 and mAb1198 share a similar binding mode as they both contact only the helix α1 C-terminus (Fig. 4d).

Next, we compared the CDRH2 loops and found that all RM Abs have slightly different approach angles for CDRH2 (Fig. 4e). The CDRH2 loops of RM1-36 and RM10-30 bind at a similar location as those of mAbs 1382 and 1198 (Fig. 4e). In contrast, CDRH2 of RM2-01 interacts only with the E2 FL helix α1 (aa 438-443), whereas CDRH2 of RM1-73, RM11-48, RM11-43, and RM1-36 extend their footprint to E2 FL N-terminus (aa 427-429), similar in footprint but not in angle to mAbs 1198 and 1382 (Fig. 4e, Supplementary Table 4). Notably, the CDRH2 tip of RM10-30 and RM1-36 also contacts E2 CD81BL W529 whereas the RM5-16 CDRH2 tip only interacts with E2 CD81BL (Fig. 4e, Supplementary Table 4).

ARTICLE IN PRESS

While binding orientation of these AR3-targeting Abs reflects alternative structural solutions to reach the AR3 epitope, neutralization efficacy likely depends more on epitope residues, CDR engagement, and preservation of CD81-blocking interactions, rather than just the binding angle itself.

AR3-targeting Abs compared to CD81 receptor binding

Clearly, the structures and competition data show that RM AR3-targeting Abs compete with host CD81 receptor binding to E2 [44,45]. To further understand how these RM Abs can block CD81 receptor binding and to rationalize their relative neutralization activity, we superimposed the E2 structures from complexes with all RM Abs (RM1-73, RM11-48, RM10-30, RM1-36, RM2-01, RM11-43, and RM5-16)**,** human AR3-targeting Abs (mAbs 1198 and 1382), and CD81 receptor long extracellular loop (PDB ID 7MWX) (Fig. 5a). As expected, all RM AR3-targeting Abs target a similar E2 region as the CD81 receptor (Fig. 5a) [22]. These and other AR3-targeting Abs bind to E2 with CD81BL in a retracted conformation, unlike its structure in complex with the CD81 receptor where it displays an extended conformation (Fig. 5a) [22].

Loop D of CD81 binds into the hydrophobic groove of E2, which is comprised of the FL N-terminus (aa 421-422), helix α1 C-terminus (aa 442-445), β-sandwich Y507, and BL (aa 612-613) [22]. We found that the CDRH3 loops of RM IGHV1-138*01 Abs (RM1-73, RM11-48, RM10-30, RM1-36, RM2-01, and RM11-43) and human IGHV1-69 mAbs 1198 and 1382 also bind into this groove (Fig. 5b,c). As a result, residues from CDRH3 of RM IGHV1-138*01 Abs, human IGHV1-69 Abs (mAbs 1198 and 1382), and CD81 loop D form somewhat similar interactions with E2 (Fig. 5c) [22]. The CD81 helix C interacts with E2 at the hydrophobic pocket formed by FL N-terminus (aa 426-429), FL helix α1 (aa 438-443), and CD81BL (aa 523-531) [22] and the CDRH2 regions of IGHV1-138*01 Abs (RM10-30 and RM1-36) make a similar interaction (Figs. 4e, 5b).

Similarities and differences within AR3-targeting Abs

Structural analyses of the RM Ab-E2 complexes show similar Fab approach angles and epitope conservation. Although IGHV4-NL_5*01 encoded RM5-16 uses a different HC germline gene, it recognizes a similar but not identical epitope as RM IGHV1-138*01 Abs (Fig. 3). For example, RM5-16 shares some common epitope residues such as FL L427, I438, A439 (A439 in HCV-1 strain), F442, Y443, and CD81BL W529 (Fig. 3, Supplementary Table 4). RM5-16 also shares numerous common epitopes within the FL loop (aa 429-438) with human AR3-targeting IGHV1-69 Abs such as AR3C, HEPC74, mAbs 1198, and 1382 (Fig. 3e) [16,20,26]. Interestingly, the CDRH3 tip of RM5-16 inserts into E2ecto at a location similar to that bound by AR3C and HEPC74 CDRH3s (Fig. 3e, Supplementary Table 4) [16,26]. While LCs possibly contribute auxiliary interactions especially in IGHV4-NL_5*01 RM5-16, the RM IGHV1-138*01 class Abs primarily depend on their HCs for E2 recognition (Figs. 2c,e, 3).

Six RM IGHV1-138*01 Abs (RM11-48, RM1-73, RM1-36, RM10-30, RM2-01, RM11-43) target overlapping conserved epitopes including E2 FL (I422, T425, L427, I438, T439, L441, F442, Y443, and A444); E2 β-sandwich (P505), CD81BL (W529), BL(Y613) (Fig. 3, Supplementary Fig. 3, Supplementary Table 4) [20,45]. The conserved epitope residues E2 FL A/S440 (A440 in E2c3 HK6a; S440 in E2ecto 1b09 strain) and BL (P612) are present for RM1-73, RM11-48, RM1-36, RM11-43, mAbs 1198, and 1382, but not RM2-01 (Fig. 3, Supplementary Fig. 3, Supplementary Table 4) [20,45].

RM1-73 has the longest CDRH3 (25 aa) among RM AR3-targeting nAbs studied here (Fig. 1f). In the complex of RM1-73 with HK6a E2c3, we see ordered electron density for E2 AS412 (aa 417-420), β-sandwich loop (aa 542-550), and post variable region 3 (pVR3) (aa 566-576), that are disordered in other RM IGHV1-138*01 Fab-HK6a E2c3 complex structures (Supplementary Fig. 7a). In previous crystal structures, the AS412 region has displayed many conformations (e.g. beta-hairpin, semi-open, and open) [56]. In the low pH structure of CD81-E2, one complex in the asymmetric unit shows AS412 residues 415-424 folding around the CD81 loop D, facilitating interaction of E2 with the CD81 receptor (Supplementary Fig. 7b) [23]. The second CD81-E2 complex in the asymmetric unit has density for E2 starting at residue 418, with residues 418-424 forming the same interaction with CD81 [23]. Upon RM1-73 binding, AS412 residues 417-424 adopt an extended conformation, following the general path of AS412 in the CD81 complex, binding close to the tip of CDRH3, bringing E2 N417, G418, S419, and W420 into close contact with RM1-73 (Supplementary Fig. 7b, Supplementary Table 4). The CD81 loop D and the long CDRH3 of RM1-73 may serve to stabilize the flexible AS412 region (Supplementary Fig. 6b, 7a), at least in the crystal structures.

Although not in the Fab epitope, the HCV-1 E2ecto AS412 region, as well as the VR2 and VR3 loops have ordered electron density in the RM5-16 complex structure (Supplementary Fig. 7c). These regions are disordered in structures in complex with human mAbs 1198 and 1382 (Supplementary Fig. 7c).

While the HCV E2 FL has been seen to adopt an alternate ‘B’ conformation in complex with human bnAbs 212.1.1 and HC1AM [21], it is found in the more common ‘A’ conformation in our seven RM Ab structures targeting AR3 as well as in the CD81 receptor complex structure (Supplementary Fig. 8).

ARTICLE IN PRESS

The motifs and somatic mutations in HCs of AR3-targeting RM Abs critical for E2 binding and virus neutralization

Both germline genes of AR3-targeting RM Abs have hydrophobic residues in their CDRH2: IGHV1-138*01 encodes $I^{51}I^{52}P^{52a}L/R^{53}V^{54}G^{55}I^{56}T^{57}$ and IGHV4-NL_5*01 encodes $I^{51}Y^{52}G^{52a}G^{53}S^{54}G^{55}S^{56}T^{57}$ (Figs. 1f,g, 4e). It has been noted that the hydrophobic CDRH2 tip of human IGHV1-69 AR3 and AS434-targeting Abs plays an important role for recognizing conserved hydrophobic epitopes on E2 [25]. Similar to human IGHV1-69 bnAbs, most RM IGHV1-138*01 AR3-targeting Abs have hydrophobic CDRH2 loops that pack into a hydrophobic pocket surrounded by the FL and CD81BL (Fig. 4e and Supplementary Fig. 9a). In contrast, the IGHV4-NL_501-encoded RM5-16 antibody does not rely on a hydrophobic CDRH2-driven recognition mechanism (Fig. 3e, 4e). Instead, RM5-16 engages E2 predominantly through polar and charged interactions at the FL-CD81BL interface, reflecting a distinct mode of antigen engagement (Fig. 3e, 4e). In RM10-30 and RM1-36 Abs, CDRH2 L53 undergoes somatic mutation to R53 (Fig. 4e and Supplementary Fig. 9a) enabling RM10-30 and RM1-36 to form a salt bridge with E531 and hydrophobic interaction with W529 of E2 CD81BL (Figs. 3, 4e and Supplementary Fig. 9a). A V54G somatic mutation occurs in RM2-01 and RM11-43 removing the hydrophobic interaction of V54 with E2 FL (Fig. 4e and Supplementary Fig. 9a).

CDRH1 S30 (in RM1-36, RM1-73, RM10-30, and RM11-48) is somatically mutated to R30 in RM2-10 and RM11-43 (Fig. 4d and Supplementary Fig. 9a). This R30 side chain occupies a similar spatial position as CDRH2 L/R53 of RM1-73, RM11-48, RM10-30, and RM1-36 (Supplementary Fig. 9a). These Arg residues make similar interactions with the FL and CD81BL regions (Supplementary Fig. 9a) [45].

In CDRH3 of AR3-targeting RM Abs, the basic, long side chains of K98 (RM1-73), R100b (RM10-30), R98 (RM11-48), and K98 (RM11-43) bind to similar regions in E2 (Fig. 5c and Supplementary Fig. 9b). In contrast, the β and γ atoms of the small S99 (RM1-36) and A98 (RM2-01) are found in similar positions as the longer side chains K98/R98/R100b (Fig. 5c and Supplementary Fig. 9b). Notably, these residues engage in extensive interactions with the E2 hydrophobic pocket formed by FL (L441, F442, Y443, and A444), which is also the binding site for the aromatic F83 of CD81 (Fig. 5c). Mutating CDRH3 residues R98A (in RM11-48) and K98A (in RM11-43) substantially reduced both the stability (or melting temperature (Tm)) and binding affinity to E2 (Supplementary Fig. 9d,e). While the RM10-30 R100bA mutation did not affect the Tm, its binding affinity to E2 decreased nine-fold compared to the RM10-30 wild-type (WT) (Supplementary Fig. 9d,e). The RM1-73 K98A mutant retained Tm and binding affinity to E2 compared to RM1-73 WT (Supplementary Fig. 9d,e), likely due to main chain H-bond interactions from RM1-73 K98 to E2 (Fig. 3a, Supplementary Table 3). In contrast, RM1-36 S99R and RM2-01 A98R mutants did not have much effect on either Tm or binding affinity to E2 when compared to the WT Fabs (Supplementary Fig. 9d,e). Previous studies revealed that mutation of K98R in RM11-43 substantially improved its ability to bind to and neutralize heterologous viruses [45], suggesting that the positive long side-chain CDRH3 K/R of RM Abs play a key role in E2 binding and virus-neutralization.

The RM5-16 CDRH2 SHM S54T and S56R play an important role in E2 binding (Fig. 3e, Supplementary Tables 3, 4). For example, CDRH2 T54 forms H-bonds and vdW contacts with CD81BL E531 and N532 while R56 is involved in extensive vdW and salt bridge interactions with E531 (Fig. 3e, Supplementary Tables 3, 4). R56 of RM5-16 binds to CD81BL E531 in a similar manner as CDRH2 R53 of RM10-30, RM1-36, and CDRH1 R30 in RM2-01 and RM11-43 (Fig. 3, Supplementary Table 4). We also observed that RM5-16 CDRH3 W100 is positioned in a similar location as the key CDRH3 disulfide motif in human IGHV1-69 class Abs AR3C and HEPC74 (Supplementary Fig. 9c) [16,26], close to a pocket formed by E2 aromatic residues (L427, L438, L441, and F442) and the E2 C429-C503 disulfide bond. The CDRH3 W100A mutation had no affinity for E2, but its Tm relative to RM5-16 WT was not affected (Supplementary Fig. 9d, e). This result supports the idea that RM5-16 CDRH3 W100 is critical for its role in E2 binding, potentially similar to the disulfide motif of human IGHV1-69 AR3C and HEPC74 (Supplementary Fig. 9c,10) [16,26].

In contrast to the dominant roles of CDR-derived mutations, heavy-chain framework regions (HFR) showed little involvement in antigen recognition, with only HFR1 residue T28 contacting antigen in RM1-73, RM11-48, RM10-30, and RM1-36 (Supplementary Table 4). Although SHMs are found in HFR3 of several RM AR3-targeting Abs (RM1-73, RM11-48, RM10-30, RM1-36, or RM5-16) (Fig. 1g,h), structural analyses revealed no contacts between HFR3 residues and the E2 antigen (Fig. 3, Supplementary Table 4). Consistently, HFR2 contained no SHMs across all Abs (Fig. 1g,h), and the HFR1 mutations (K19R in RM1-73 and V2L/V12M in RM5-16) did not

ARTICLE IN PRESS

participate in antigen contact. These findings indicate that HFR-associated SHMs are unlikely to play a functional role in antigen recognition in these RM AR3-targeting Abs.

Collectively, our sequence and mutation analyses indicate that specific somatically derived amino acids in RM AR3-targeting nAbs that are positioned at the Ab-E2 binding interface play important roles in mediating E2 binding as well as in neutralization activity.

Discussion

One of the challenges for HCV vaccine research is the absence of an immunocompetent animal model for preclinical immunization and protection studies [4,57,58]. The extensive genetic diversity of HCV, along with the flexibility of the E2 glycoprotein, may necessitate rational design of a cross-genotype vaccine that targets conserved epitopes [4,21,58,59]. The rational design and evaluation of such HCV vaccine immunogens can be guided by structural information of nAb epitopes [18,25,55,58,60-63].

Compiling information about shared germline gene segments, structural features, and antigen-interaction patterns of nAbs derived from humans or animals is important for guiding effective HCV immunogen design. These data are increasing as more human bnAbs against HCV are identified from both chronically and spontaneously infected HCV patients [14,25-27]. Anti-HCV human or human-like IGHV1-69 class Abs, from both phage display and HCV-infected patients or NHPs, seem to preferentially target the conserved AR3 region of E2 [25,44,45]. Human IGHV1-69 AR3-targeting Abs also bind to a similar epitope with most of the contacts from their HCs [25]. Human IGHV1-69 AR3-targeting nAbs possess signature hydrophobic sequences consisting of CDRH2 $I/V/T^{52}$-P^{52a}-X^{53}-F/S^{54} (X = hydrophobic residue) and an aromatic residue Tyrosine (Y) in CDRH3 that insert into the hydrophobic AR3 of E2 [25].

Our RM Ab structural analyses reveal that, like human IGHV1-69 class Abs, the RM IGHV1-138*01 class Abs primarily use their HC to interact with E2 (Fig. 2c). CDRH3 contributes substantial BSA in most RM IGHV1-138*01 (or human IGHV1-69) class Abs (Fig. 2c). Importantly, the RM IGHV1-138*01 is highly similar to the human IGHV1-69 germline gene, and their mature Abs have very low SMH (~1-4 %) (Supplementary Fig. 1d). Basic K/R and hydrophobic residues in CDRH3 and a hydrophobic motif $I/V^{52}P^{52a}L/R^{53}V/G^{54}$ at the CDRH2 tip interacts with the conserved and hydrophobic AR3 region of E2. Their long K/R CDRH3 residues can reach into and extensively interact with E2 in the AR3 hydrophobic groove through multiple types of interaction such as H-bonds, vdW, and hydrophobic interactions (Supplementary Fig. 9b), facilitating the specificity and affinity of the antigen-Ab interactions. Some key CDRH1 and CDRH2 residues have exchangeable roles on binding E2 (Supplementary Fig. 9a).

Notably, these RM IGHV1-138*01 class Abs share very similar binding modes and epitopes with human IGHV1-69 Abs from human elite neutralizers [20]. Although NHP-derived AR3-targeting Abs approach the E2 neutralizing face with slightly different orientations and angles, they recognize a shared group of conserved AR3 residues with human anti-HCV IGHV1-69 bnAbs. As NHPs share many similarities with humans in their immune systems [32], the functional and structural similarities of the antibodies described here suggest that the NHP model can to some extent recapitulate the human immune response to the same HCV antigen and is a valuable preclinical study model to study how Abs are elicited by HCV vaccine candidates.

ARTICLE IN PRESS

Anti-HCV AR3-targeting Abs were previously grouped into two classes based on their CDRH3 loop shape and secondary structure: either straight β-hairpin (HEPC3, HEPC74, hcab55, hcab64, and AT1209) [5,26,27] or bent β-hairpin (AR3A, AR3C, and HC11) [16,21,24]. CDRH3s for RM1-73, RM11-48, and RM11-43 incorporate helical segments when bound to E2 (Fig. 4a, Supplementary Fig. 2a) [45], in contrast to loop conformation of the unliganded RM1-73 Fab (Supplementary Fig. 2a). Like CDRH3 loops of mAbs 1198 and 1382, those of RM10-30, RM2-01, RM1-36, and RM5-16 have loop conformations when bound to E2 (Fig. 5a). These observations suggest that the shape and length of the CDRH3 loop may influence the approach angles of Abs to the antigen.

RM5-16 from a different immunized animal arose from the IGHV4-NL_5*01 germline (a homolog of human IGHV4-59*12) but targets the same AR3 region using interactions from both HC and LC (Figs. 2e, 3e). As noted, IGHV4-NL_5*01 germline alleles appear 3.4-fold more frequently than IGHV1-138*01 germline alleles in the animals contributing to the MUSA dataset (Fig. 1d). Moreover, the human IGHV4-59 germline gene was also found to be the most frequently used among B cell receptors in chronic HCV patients, both prior to and during DAA treatment [64], indicating that RM IGHV4-NL_5*01 or human IGHV4-59 germline genes may be prevalent in effective Ab-mediated responses against HCV. To date, structural information for anti-HCV human IGHV4-59 as well as NHP IGHV4-NL_5*01 class Abs remains sparse.

Recently, structures for FL-binding human Abs hcab55 and hcab64 encoded by the human IGHV1-46 germline gene have been reported [27]. Although encoded by a different IGHV gene, IGHV1-46 class nAbs utilize similar approach angles to recognize overlapping epitopes in the E2 FL as human IGHV1-69 class HEPC74 (Fig. 4, Supplementary Fig. 10) [26,27]. Structural comparison revealed that RM5-16, HEPC74, AR3C, hcab55, and hcab64 all target the AR3 neutralizing face of E2, with substantial contact to the FL and CD81BL region (Fig. 4, Supplementary Fig. 10a-c). Notably, none of these Abs contact the BL with their CDRH3, unlike Abs RM1-73, RM11-48, RM10-30, RM1-36, RM2-01, RM11-43, mAb1198, and mAb1382 that do contact the BL with CDR H3 (Fig. 4, Supplementary Fig. 10). Interestingly, the CDRH3 W100 of RM5-16 occupies a position analogous to the conserved CDRH3 disulfide motif observed in HEPC74, AR3C, hcab55, and hcab64 (Supplementary Fig. 10d). Consistent with this structural convergence, mutation of RM5-16 W100 (Supplementary Fig. 9e) or disruption of the CDRH3 disulfide in HEPC74 or AR3C significantly reduced binding to HCV E2 [26], indicating that residues engaging the cleft between the FL and CD81BL (Supplementary Fig. 9c, 10d) play a critical role in E2 recognition.

Unlike IGHV1-138*01 Abs, RM5-16 does not interact with the E2 BL C-terminal region (aa 612-616), but instead makes extensive contact with the CD81BL (Fig. 3e). E2 CD81BL Y529 and W531 are thought to be proximal to the membrane and involved in membrane binding [22,23], suggesting that RM5-16 might have an alternative neutralization mechanism beyond most characterized AR3-targeting class Abs. Of note, despite different angles of approach, we found that the IGHV4-NL_5*01 (a homolog of human IGHV4-59*12) and IGHV1-138 *01 (a homolog of human IGHV1-69*10) class Abs have similar binding footprints and occupy the same set of conserved hydrophobic pockets in the AR3 region of E2, confirming convergent recognition in NHPs as well as humans. Notwithstanding, the binding mode of the NHP IGHV4-NL_5*01 class Abs might be specific to NHPs or not seen yet in humans. Thus, RM5-16 offers another starting point for uncovering analogous human Ab responses to HCV infection.

ARTICLE IN PRESS

ARTICLE IN PRESS

Importantly, all RM AR3-targeting Abs studied here block CD81 binding to E2 [44] and exhibit a binding mode analogous to that of the host CD81 receptor (Fig. 5). Collectively, host receptor functional binding mimicry of AR3-targeting Abs highlights parallels between receptor engagement and Ab neutralization.

While HCV does not naturally infect RMs, limiting their use for studying the full viral life cycle, our structural analysis of RM Ab-antigen complexes, alongside comparisons with human Abs and host receptor, provide valuable insights into the immune response to HCV immunogens. Our findings offer critical preclinical data and provide a baseline for evaluating the ability of vaccine candidates to elicit nAbs and protect against diverse HCV strains.

Methods

IG-targeted capture genome sequencing, assembly, and analysis protocol

From the RM PBMCs, 2.5 ug of high molecular weight DNA was sheared using a g-tube (Covaris, Woburn, MA, United States) to ~14 Kbp at 4000 RPM and size selected via bead-based size selection to remove fragments less than 3 Kbp. Sheared gDNA underwent end-repair and A-tailing using the standard KAPA library protocol (KAPA Hyper Prep Kit; Roche, Indianapolis, IN, United States). Barcodes were added to samples sequenced on the Sequel IIe platform and universal primers ligated to all samples. PCR amplification was performed for 8 cycles using PrimeSTAR GXL Polymerase (Takara, San Jose, CA, United States) at an annealing temperature of 60 °C. Small fragments and excess reagents were then removed using 0.7X vol: vol KAPA Pure beads (Roche). Genomic DNA target enrichment was carried out using oligo probes designed directly from the reference sequence for IGH/K/L of rhemac10 [65]. (IGH (chr7:167,854,585-169,917,761), IGK (chr13:16,744,193-18,180,859), IGL (chr10:29,581,424-30,956,134), and assembly following Cirelli et al. [66]. Constructed capture libraries were washed using the KAPA Hyper-Cap protocol (KAPA HyperCapture Reagent kit and KAPA HyperCapture Bead kit, Roche), and post-capture PCR amplification performed for 15 cycles using PrimeSTAR GXL Polymerase (Takara) at an annealing temperature of 62 °C. Sequencing SMRTbell libraries were prepared using the SMRTbell Template Prep Kit 2.0 and SMRTbell Enzyme Cleanup Kit 1.0 (Pacific Biosciences, Menlo Park, CA, United States). Each sample was treated with a DNA Damage Repair and End Repair mix to repair nicked DNA, followed by A-tailing and ligation with SMRTbell hairpin adapters. These libraries were treated with an exonuclease cocktail to remove unligated gDNA and cleaned with 0.6X AMPure PB beads (Pacific Biosciences). The resulting SMRTbell libraries were prepared for sequencing according to the manufacturer’s protocol and sequenced on the Sequel IIe system using 2.0 chemistry and 30 hr movies. HiFi data, consisting of circular consensus sequences filtered at a quality threshold of QV20 (99%), were generated on the instrument and used for all downstream analysis.

HiFi sequencing reads were assembled into haplotype-resolved assemblies using Hifiasm [67] with default parameters. The IG loci were extracted and annotated using Digger [49]. Sequences of Abs for each animal were aligned to their respective IG assemblies using BLAT [68]. The locations of the highest BLAT matches were linked to the corresponding germline allele Digger annotations. Allele assignments were based on names used by KIMDB [38] and MUSA [48] databases, both of which included exact matches to the closest germline identified in the assemblies of each animal. The SHM for the RM Abs were assigned by IgBLAST with the input of *Macaca mulatta* (IGHV,

ARTICLE IN PRESS

D, and J germline gene) sequence files from the sequenced germline alleles. All HC and LC sequences of RM Ab in this study are aligned by Clustal Omega [69]. E2 isolate sequence alignment formatted with ESPript[70]. The Fab residues were renumbered according to Kabat nomenclature [71].

Protein expression and purification

RM Fab-derived V_L and V_H DNA fragments were inserted into the pHCMV mammalian cell expression vector containing the corresponding human λ or κ CL region and the human IgG C_H1 region, terminating with a stop codon following the CH1 sequence [44,72]. To enhance crystallization, we introduced the 'CK' mutation into κ CL regions of RM1-73, RM10-30, RM11-48, RM1-36, and HEPC46 Fabs [73]. The HC and LC encoding plasmids for each Fab were co-transfected into Expi293F cells using ExpiFectamine (Thermo Fisher Scientific) according to the manufacturer's instructions. The cells were grown and harvested after 7 days. Affinity chromatography using CaptureSelect CH1-XL affinity matrix (Thermo Fisher Scientific) was used to purify recombinant Fabs from culture supernatant, followed by size-exclusion chromatography (SEC) using a Superdex 200 column (GE Healthcare).

The HK6a (genotype 6a) E2c3 proteins were purified as previously described [24]. E2 proteins from isolates H77 (1a), HCV-1 (1a), J6 (2a), and S52 (3a) were prepared for biophysical assays as reported for the HK6a E2c3 protein. For HCV-1 E2ecto and HEPC46 Fab, the complex was expressed by co-transfecting expression vectors encoding HEPC46 Fab and untagged E2ecto (residues 384-643) from strain HCV-1. HEPC46-E2ecto complexes were then purified from Expi293F cell supernatants using CaptureSelect CH1-XL affinity matrix followed by SEC. All purified proteins were quantified by optical absorbance at 280 nm and stored in 20 mM Tris pH 8.0 and 150 mM NaCl (TBS) at -80 ºC for further experiments. The purities of recombinant proteins were analyzed by reducing and nonreducing SDS-PAGE.

Biolayer interferometry

E2 binding with Fabs was evaluated by BLI using an Octet Red instrument (ForteBio) at 10 ºC [45]. The Fabs at 20 μg/mL in 1× kinetics buffer (1× PBS, pH 7.4, 0.01% BSA, and 0.002% Tween 20) were immobilized onto Fab-2G biosensors and interacted with a two-fold gradient dilution of E2 analyte proteins starting at 540 nM. The kinetic assay consisted of the sequential steps: baseline equilibration (60 s, buffer), Fab protein immobilization (180 s), washing unbound Fab proteins (30 s, buffer), second baseline stabilization (60 s, buffer), E2 binding association (600 s), and E2 dissociation measurement (600 s, buffer). For estimating the dissociation constant (K_d), a 1:1 binding model was used.

HCV neutralization assays

HCV pseudoparticles (HCVpp) neutralization assays were performed as previously described [74,75]. Briefly, HCVpp were generated by co-transfection of 293T cells with pNL4-3.lucR-E- and E1E2 expression plasmids from selected isolates with tier 1-4 antibody resistance [76-79] at a 4:1 ratio by polyethylenimine (PEI, Polysciences). Neutralization was assessed on Huh-7 cells with mAbs (25 μg/ml) as recently described [74]. Virus infectivity was quantified using the Bright-Glo luciferase assay system (Promega), and percent neutralization was calculated relative to no-antibody controls after background subtraction. Background signal was defined using pseudoparticles generated

ARTICLE IN PRESS

ARTICLE IN PRESS

with pNL4-3. lucR-E− alone. Pseudoparticles displaying the vesicular stomatitis virus envelope glycoprotein G (VSVpp) were used as control for nonspecific neutralizing activity.

Differential scanning fluorimetry (nanoDSF)

A Promethius Panta (nanoTemper Technologies) was used to measure the melting temperatures of the wild-type (WT) and mutant RM Fabs. The WT and mutant Fabs at 1 mg/ml in TBS were loaded into capillaries and inserted into the sample holder. We applied a temperature gradient of 1°C/min from 25 to 95 ºC for each sample. The intrinsic protein fluorescence at 330 and 350 nm was recorded. Three independent measurements were performed for each sample. The data was analyzed by Panta Analysis v1.8. The melting point (Tm) was determined as the maximum of the first derivative of the melting curve.

Crystal structure determination

For complexes of RM1-36, RM10-30, RM11-48, and RM1-73 with HK6a E2c3, E2 proteins were combined with Fab in a 1:1.2 molar ratio. For RM5-16, the HCV-1 E2ecto complex with HEPC46 was mixed with RM5-16 Fab in a 1:1.2 molar ratio. The mixture was incubated overnight at 4ºC before further purification by SEC (Superdex 200 GL column) in TBS buffer to remove uncomplexed Fab. Crystallization experiments were set up using the sitting drop vapor diffusion method at 20°C. Initial crystallization conditions were obtained from trials using our robotic, high-throughput Rigaku CrystalMation system. Crystals were manually optimized using the sitting-drop vapor diffusion method at 20 °C using Cryschem plates (Hampton Research Corp). Crystallization precipitant conditions are listed in Supplementary Table 6. The protein crystals were flash-cooled at 100 K with 20% (w/v) ethylene glycol added as cryoprotectant.

Diffraction data were collected at synchrotron beamlines and integrated and scaled with HKL-2000 [80], or for the unliganded RM1-73 Fab were collected in-house with a Rigaku MicroMax-007 generator at 1.5418 Å wavelength and a Mar345dtb area detector. Data collection statistics are summarized in Supplementary Table 1, 2. Crystal structures were phased by the molecular replacement (MR) method using the program Phaser [81]. The initial input models for RM Fabs were generated by Repertoire Builder (https://sysimm.ifrec.osaka-u.ac.jp/rep_builder/) [82]. The MR template for E2 was derived from HK6a E2c3 (PDB ID: 7JTG) or HCV-1 E2ecto (PDB ID: 6MEJ). Refinement was carried out in Phenix [83]. Structure models were examined and modified with the program Coot to assess model fit and any potential clashes [84]. Final refinement statistics are summarized in Supplementary Table 1, 2 and their quality analyzed with MolProbity [85].

Structural analysis

Epitope and paratope residues, molecular interactions, BSA of E2, Fabs, and CD81, were identified and calculated with the Protein Interfaces, Surfaces and Assemblies (PISA) web server at the European Bioinformatics Institute (www.ebi.ac.uk/pdbe/pisa/) [86] followed by manual inspection. Structure figures were created by MacPyMol (DeLano Scientific LLC). RMSD calculations were done in PyMOL following pairwise Cα alignments without excluding outliers. Epitope variability analyses were carried out using MS [51] with a 1.4Å probe radius and Wu-Kabat variability calculated as described in [52].

Data availability

ARTICLE IN PRESS

The X-ray coordinates and structure factors have been deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank under accession codes 9MS9, 9MRZ, 9MNS, 9MNU, 9MNT, 9MNQ, and 9MSC.

Acknowledgments
We thank Henry Tien for automated robotic crystal screening at The Scripps Research Institute and Xiaoping Dai for help with in-house data collection. X-ray diffraction data were collected at the Advanced Photon Source (APS) beamline 23ID-B (GM/CA CAT) and the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2. GM/CA CAT is funded in whole or in part with federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (NIGMS) (Y1-GM-1104). Use of the APS was supported by the US Department of Energy (DOE), Basic Energy Sciences, Office of Science, under Contract DE-AC02-06CH11357. The SSRL is a Directorate of SLAC National Accelerator Laboratory, and an Office of Science User Facility operated for the US DOE of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the NIH, NIGMS (including P41GM103393), and the National Center for Research Resources (NCRR) (P41RR001209). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIAID, NIGMS, NCRR, or NIH.

This investigation used resources that were supported by the Southwest National Primate Research Center grant P51 OD011133 from the Office of Research Infrastructure Programs, National Institutes of Health. Research reported in this publication was supported by the Office of the Director, National Institutes of Health under Award Numbers S10OD028732 and S10OD032443. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding: this work was funded in part by NIH AI168251 (ML, IAW, and RLS) and AI168917 (ML).

ARTICLE IN PRESS

Author contributions
Conceptualization: YTKN, FC, RLS, ML, IAW
Methodology: YTKN, FC, RLS, ML, IAW
Investigation: YTKN, FC, EG, SS, NT, LAU, RLS, CC
Visualization: YTKN, FC, RLS, CW
Supervision: RLS, CTW, ML, IAW
Writing—original draft: YTKN, FC, RLS, SS, CTW, ML, IAW
Writing—review & editing: YTKN, FC, RLS, SS, CC, CTW, ML, IAW

Competing interests
The authors declare no competing financial or non-financial interests.

Abbreviations
Hepatitis C virus (HCV), non-human primates (NHP), immunoglobulin (IgG), rhesus macaques (RM), broadly neutralizing antibodies (bnAbs), direct-acting antiviral drug (DAA), individualized germline set (IGS), multiple sequence alignment (MSA), antibody heavy-chain (HC), antibody

ARTICLE IN PRESS

light-chain (LC), front layer (FL), back layer (BL), CD81 binding loop (CD81BL), hydrogen-bonds (H-bond), van der Waals (vdW), somatic hypermutation (SHM), buried surface area (BSA), root-mean-square deviation (RMSD), variable domains of antibody heavy-chain (IGHV), antigenic region 3 (AR3), complementarity determining region (CDR).

References

1 Falade-Nwulia, O. *et al.* Oral direct-acting agent therapy for Hepatitis C virus infection: A systematic review. *Ann. Intern. Med.* **166**, 637-648, doi:10.7326/m16-2575 (2017).

2 Kanwal, F. *et al.* Risk of hepatocellular cancer in HCV patients treated with direct-acting antiviral agents. *Gastroenterology* **153**, 996-1005.e1, doi:10.1053/j.gastro.2017.06.012 (2017).

3 Midgard, H. *et al.* HCV epidemiology in high-risk groups and the risk of reinfection. *J. Hepatol.* **65**, S33-S45, doi:10.1016/j.jhep.2016.07.012 (2016).

4 Echeverría, N. *et al.* In the era of rapid mRNA-based vaccines: Why is there no effective Hepatitis C virus vaccine yet? *World J. Hepatol.* **13**, 1234-1268, doi:10.4254/wjh.v13.i10.1234 (2021).

5 Torrents de la Peña, A. *et al.* Structure of the Hepatitis C virus E1E2 glycoprotein complex. *Science* **378**, 263-269, doi:10.1126/science.abn9884 (2022).

6 Metcalf, M. C. *et al.* Structure of engineered Hepatitis C virus E1E2 ectodomain in complex with neutralizing antibodies. *Nat. Commun.* **14**, 3980, doi:10.1038/s41467-023-39659-z (2023).

7 Augestad, E. H. *et al.* The Hepatitis C virus envelope protein complex is a dimer of heterodimers. *Nature* **633**, 704-709, doi:10.1038/s41586-024-07783-5 (2024).

8 Pileri, P. *et al.* Binding of Hepatitis C virus to CD81. *Science* **282**, 938-41, doi:10.1126/science.282.5390.938 (1998).

9 Evans, M. J. *et al.* Claudin-1 is a Hepatitis C virus co-receptor required for a late step in entry. *Nature* **446**, 801-5, doi:10.1038/nature05654 (2007).

10 Scarselli, E. *et al.* The human scavenger receptor class B type I is a novel candidate receptor for the Hepatitis C virus. *Embo J.* **21**, 5017-25, doi:10.1093/emboj/cdf529 (2002).

11 Ploss, A. *et al.* Human occludin is a Hepatitis C virus entry factor required for infection of mouse cells. *Nature* **457**, 882-6, doi:10.1038/nature07684 (2009).

12 Lindenbach, B. D. & Rice, C. M. The ins and outs of Hepatitis C virus entry and assembly. *Nat. Rev. Microbiol.* **11**, 688-700, doi:10.1038/nrmicro3098 (2013).

13 Owsianka, A. *et al.* Monoclonal antibody AP33 defines a broadly neutralizing epitope on the Hepatitis C virus E2 envelope glycoprotein *J. Virol.* **79**, 11095-11104, doi:10.1128/jvi.79.17.11095-11104.2005 (2005).

14 Law, M. *et al.* Broadly neutralizing antibodies protect against Hepatitis C virus quasispecies challenge. *Nat. Med.* **14**, 25-7, doi:10.1038/nm1698 (2008).

15 Harman, C. *et al.* A view of the E2-CD81 interface at the binding site of a neutralizing antibody against Hepatitis C virus. *J. Virol.* **89**, 492-501, doi:10.1128/jvi.01661-14 (2015).

16 Kong, L. *et al.* Hepatitis C virus E2 envelope glycoprotein core structure. *Science* **342**, 1090-4, doi:10.1126/science.1243876 (2013).

17 Tzarum, N., Wilson, I. A. & Law, M. The neutralizing face of Hepatitis C virus E2 envelope glycoprotein. *Front. Immunol.* **9**, 1315, doi:10.3389/fimmu.2018.01315 (2018).

ARTICLE IN PRESS

ARTICLE IN PRESS

18 Kong, L., Jackson, K. N., Wilson, I. A. & Law, M. Capitalizing on knowledge of Hepatitis C virus neutralizing epitopes for rational vaccine design. *Curr. Opin. Virol.* **11**, 148-57, doi:10.1016/j.coviro.2015.04.001 (2015).

19 Giang, E. *et al.* Human broadly neutralizing antibodies to the envelope glycoprotein complex of Hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **109**, 6205-6210, doi:10.1073/pnas.1114927109 (2012).

20 Weber, T. *et al.* Analysis of antibodies from HCV elite neutralizers identifies genetic determinants of broad neutralization. *Immunity* **55**, 341-354.e7, doi:10.1016/j.immuni.2021.12.003 (2022).

21 Tzarum, N. *et al.* An alternate conformation of HCV E2 neutralizing face as an additional vaccine target. *Sci. Adv.* **6**, eabb5642, doi:10.1126/sciadv.abb5642 (2020).

22 Kumar, A. *et al.* Structural insights into Hepatitis C virus receptor binding and entry. *Nature* **598**, 521-525, doi:10.1038/s41586-021-03913-5 (2021).

23 Kumar, A. *et al.* Regions of Hepatitis C virus E2 required for membrane association. *Nat. Commun.* **14**, 433, doi:10.1038/s41467-023-36183-y (2023).

24 Tzarum, N. *et al.* Genetic and structural insights into broad neutralization of Hepatitis C virus by human V_H1-69 antibodies. *Sci. Adv.* **5**, eaav1882, doi:10.1126/sciadv.aav1882 (2019).

25 Chen, F., Tzarum, N., Wilson, I. A. & Law, M. V_H1-69 antiviral broadly neutralizing antibodies: genetics, structures, and relevance to rational vaccine design. *Curr. Opin. Virol.* **34**, 149-159, doi:10.1016/j.coviro.2019.02.004 (2019).

26 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.e3, doi:10.1016/j.chom.2018.10.009 (2018).

27 Ogega, C. O. *et al.* Convergent evolution and targeting of diverse E2 epitopes by human broadly neutralizing antibodies are associated with HCV clearance. *Immunity* **57**, 890-903.e6, doi:10.1016/j.immuni.2024.03.001 (2024).

28 Wu, N. C. & Wilson, I. A. Influenza hemagglutinin structures and antibody recognition. *Cold Spring Harb. Perspect. Med.* **10**, a038778, doi:10.1101/cshperspect.a038778 (2020).

29 Corti, D. & Lanzavecchia, A. Broadly neutralizing antiviral antibodies. *Annu. Rev. Immunol.* **31**, 705-42, doi:10.1146/annurev-immunol-032712-095916 (2013).

30 Barnes, E., Cooke, G. S., Lauer, G. M. & Chung, R. T. Implementation of a controlled human infection model for evaluation of HCV vaccine candidates. *Hepatology* **77**, 1757-1772, doi:10.1002/hep.32632 (2023).

31 Liang, T. J., Feld, J. J., Cox, A. L. & Rice, C. M. Controlled human infection model - fast track to HCV vaccine? *N. Engl. J. Med.* **385**, 1235-1240, doi:10.1056/NEJMsb2109093 (2021).

32 Gibbs, R. A. *et al.* Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**, 222-34, doi:10.1126/science.1139247 (2007).

33 Lemaitre, J. *et al.* Non-human primate models of human respiratory infections. *Mol. Immunol.* **135**, 147-164, doi:10.1016/j.molimm.2021.04.010 (2021).

34 Yu, G. Y. *et al.* Cynomolgus macaque (*Macaca fascicularis*) immunoglobulin heavy chain locus description. *Immunogenetics* **68**, 417-428, doi:10.1007/s00251-016-0921-2 (2016).

35 Antinori, S. *et al.* Non-human primate and human malaria: Past, present and future. *J. Travel Med.* **28**, taab036, doi:10.1093/jtm/taab036 (2021).

ARTICLE IN PRESS

ARTICLE IN PRESS

36 Roark, R. S. *et al.* Recapitulation of HIV-1 Env-antibody coevolution in macaques leading to neutralization breadth. *Science* **371**, eabd2638, doi:10.1126/science.abd2638 (2021).
37 He, W. T. *et al.* Broadly neutralizing antibodies to SARS-related viruses can be readily induced in rhesus macaques. *Sci. Transl. Med.* **14**, eabl9605, doi:10.1126/scitranslmed.abl9605 (2022).
38 Vázquez Bernat, N. *et al.* Rhesus and cynomolgus macaque immunoglobulin heavy-chain genotyping yields comprehensive databases of germline VDJ alleles. *Immunity* **54**, 355-366.e4, doi:10.1016/j.immuni.2020.12.018 (2021).
39 Lefranc, M. P. IMGT, the international immunogenetics database. *Nucleic Acids Res.* **31**, 307-10, doi:10.1093/nar/gkg085 (2003).
40 Ramesh, A. *et al.* Structure and diversity of the rhesus macaque immunoglobulin loci through multiple de novo genome assemblies. *Front. Immunol.* **8**, 1407, doi:10.3389/fimmu.2017.01407 (2017).
41 Kaduk, M., Corcoran, M. & Karlsson Hedestam, G. B. Addressing IGHV gene structural diversity enhances immunoglobulin repertoire analysis: Lessons from rhesus macaque. *Front. Immunol.* **13**, 818440, doi:10.3389/fimmu.2022.818440 (2022).
42 Corcoran, M. M. *et al.* Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity. *Nat. Commun.* **7**, 13642, doi:10.1038/ncomms13642 (2016).
43 Guo, Y., Waltari, E., Lu, H., Sheng, Z. & Wu, X. Novel rhesus macaque immunoglobulin germline genes identified by three sequencing approaches. *Front. Immunol.* **15**, 1506348, doi:10.3389/fimmu.2024.1506348 (2024).
44 Chen, F. *et al.* Antibody responses to immunization with HCV envelope glycoproteins as a baseline for B-cell-based vaccine development. *Gastroenterology* **158**, 1058-1071.e6, doi:10.1053/j.gastro.2019.11.282 (2020).
45 Chen, F. *et al.* Functional convergence of a germline-encoded neutralizing antibody response in rhesus macaques immunized with HCV envelope glycoproteins. *Immunity* **54**, 781-796.e4, doi:10.1016/j.immuni.2021.02.013 (2021).
46 Frey, S. E. *et al.* Safety and immunogenicity of HCV E1E2 vaccine adjuvanted with MF59 administered to healthy adults. *Vaccine* **28**, 6367-73, doi:10.1016/j.vaccine.2010.06.084 (2010).
47 Ray, R. *et al.* Characterization of antibodies induced by vaccination with Hepatitis C virus envelope glycoproteins. *J. Infect. Dis.* **202**, 862-6, doi:10.1086/655902 (2010).
48 Peres, A. *et al.* Population-level genomic analysis of immunoglobulin loci variation in rhesus macaques reveals extensive germline diversity. *Immunity* **59**, 213-228 e6, doi:10.1016/j.immuni.2025.12.002 (2026).
49 Lees, W. D., Saha, S., Yaari, G. & Watson, C. T. Digger: directed annotation of immunoglobulin and T cell receptor V, D, and J gene sequences and assemblies. *Bioinformatics* **40**, btae144, doi:10.1093/bioinformatics/btae144 (2024).
50 Frumento, N., Flyak, A. I. & Bailey, J. R. Mechanisms of HCV resistance to broadly neutralizing antibodies. *Curr. Opin. Virol.* **50**, 23-29, doi:10.1016/j.coviro.2021.07.003 (2021).
51 Connolly, M. L. The molecular surface package. *J. Mol. Graph.* **11**, 139-41, doi:10.1016/0263-7855(93)87010-3 (1993).

ARTICLE IN PRESS

ARTICLE IN PRESS

52 Wu, T. T. & Kabat, E. A. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* **132**, 211-50, doi:10.1084/jem.132.2.211 (1970).

53 Olson, R. D. *et al.* Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Res.* **51**, D678-D689, doi:10.1093/nar/gkac1003 (2023).

54 Goffard, A. & Dubuisson, J. Glycosylation of hepatitis C virus envelope proteins. *Biochimie* **85**, 295-301, doi:10.1016/s0300-9084(03)00004-x (2003).

55 He, L. *et al.* Proof of concept for rational design of Hepatitis C virus E2 core nanoparticle vaccines. *Sci. Adv.* **6**, eaaz6225, doi:10.1126/sciadv.aaz6225 (2020).

56 Gomez-Escobar, E., Roingeard, P. & Beaumont, E. Current Hepatitis C vaccine candidates based on the induction of neutralizing antibodies. *Viruses* **15**, 1151, doi:10.3390/v15051151 (2023).

57 Burm, R., Collignon, L., Mesalam, A. A. & Meuleman, P. Animal models to study Hepatitis C virus infection. *Front. Immunol.* **9**, 1032, doi:10.3389/fimmu.2018.01032 (2018).

58 Bailey, J. R., Barnes, E. & Cox, A. L. Approaches, progress, and challenges to Hepatitis C vaccine development. *Gastroenterology* **156**, 418-430, doi:10.1053/j.gastro.2018.08.060 (2019).

59 Ströh, L. J., Nagarathinam, K. & Krey, T. Conformational flexibility in the CD81-binding site of the Hepatitis C virus glycoprotein E2. *Front. Immunol.* **9**, 1396, doi:10.3389/fimmu.2018.01396 (2018).

60 Nabel, G. J. Designing tomorrow's vaccines. *N. Engl. J. Med.* **368**, 551-60, doi:10.1056/NEJMra1204186 (2013).

61 Burton, D. R. & Hangartner, L. Broadly neutralizing antibodies to HIV and their role in vaccine design. *Annu. Rev. Immunol.* **34**, 635-59, doi:10.1146/annurev-immunol-041015-055515 (2016).

62 Burton, D. R. Antibodies, viruses and vaccines. *Nat. Rev. Immunol.* **2**, 706-13, doi:10.1038/nri891 (2002).

63 Dormitzer, P. R., Ulmer, J. B. & Rappuoli, R. Structure-based antigen design: a strategy for next generation vaccines. *Trends Biotechnol.* **26**, 659-67, doi:10.1016/j.tibtech.2008.08.002 (2008).

64 Schultheiß, C. *et al.* B cells expressing mutated IGHV1-69-encoded antigen receptors related to virus neutralization show lymphoma-like transcriptomes in patients with chronic HCV infection. *Hepatol. Commun.* **8**, e0503, doi:10.1097/hc9.0000000000000503 (2024).

65 Warren, W. C. *et al.* Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. *Science* **370**, eabc6617, doi:10.1126/science.abc6617 (2020).

66 Cirelli, K. M. *et al.* Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance. *Cell* **177**, 1153-1171.e28, doi:10.1016/j.cell.2019.04.012 (2019).

67 Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170-175, doi:10.1038/s41592-020-01056-5 (2021).

68 Kent, W. J. BLAT— the BLAST-like alignment tool. *Genome Res.* **12**, 656-64, doi:10.1101/gr.229202 (2002).

ARTICLE IN PRESS

ARTICLE IN PRESS

69 Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276-W279, doi:10.1093/nar/gkac240 (2022).
70 Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-8, doi:10.1093/bioinformatics/15.4.305 (1999).
71 Abhinandan, K. R. & Martin, A. C. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol. Immunol.* **45**, 3832-9, doi:10.1016/j.molimm.2008.05.022 (2008).
72 Tiller, T. *et al.* Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112-24, doi:10.1016/j.jim.2007.09.017 (2008).
73 Lieu, R. *et al.* Rapid and robust antibody Fab fragment crystallization utilizing edge-to-edge beta-sheet packing. *PLoS One* **15**, e0232311, doi:10.1371/journal.pone.0232311 (2020).
74 Bailey, J. R., Urbanowicz, R. A., Ball, J. K., Law, M. & Foung, S. K. H. Standardized method for the study of antibody neutralization of HCV pseudoparticles (HCVpp). *Methods Mol. Biol.* **1911**, 441-450, doi:10.1007/978-1-4939-8976-8_30 (2019).
75 Chen, F. *et al.* The conserved bridging domain on HCV E1E2 glycoprotein complex is targeted by neutralizing antibodies from diverse lineages. *bioRxiv*, 2025.11.05.686883, doi:10.1101/2025.11.05.686883 (2025).
76 Bartosch, B., Dubuisson, J. & Cosset, F. L. Infectious Hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes. *J. Exp. Med.* **197**, 633-42, doi:10.1084/jem.20021756 (2003).
77 Lavillette, D. *et al.* Characterization of host-range and cell entry properties of the major genotypes and subtypes of Hepatitis C virus. *Hepatology* **41**, 265-74, doi:10.1002/hep.20542 (2005).
78 Meunier, J. C. *et al.* Evidence for cross-genotype neutralization of Hepatitis C virus pseudo-particles and enhancement of infectivity by apolipoprotein C1. *Proc. Natl. Acad. Sci. USA.* **102**, 4560-5, doi:10.1073/pnas.0501275102 (2005).
79 Li, Y. P., Ramirez, S., Mikkelsen, L. & Bukh, J. Efficient infectious cell culture systems of the Hepatitis C virus (HCV) prototype strains HCV-1 and H77. *J. Virol.* **89**, 811-23, doi:10.1128/jvi.02877-14 (2015).
80 Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-26, doi:10.1016/s0076-6879(97)76066-x (1997).
81 McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674, doi:10.1107/s0021889807021206 (2007).
82 Schritt, D. *et al.* Repertoire builder: High-throughput structural modeling of B and T cell receptors. *Mol. Syst. Des. Eng.* **4**, 761-768, doi:10.1039/C9ME00020H (2019).
83 Adams, P. D. *et al.* PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948-54, doi:10.1107/s0907444902016657 (2002).
84 Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-32, doi:10.1107/s0907444904019158 (2004).
85 Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12-21, doi:10.1107/s0907444909042073 (2010).

ARTICLE IN PRESS

86 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774-97, doi:10.1016/j.jmb.2007.05.022 (2007).

Figure Legends

Fig. 1: Overall genetic analysis of AR3-targeting nAbs from NHP. a) Overview: sample collection, single-cell sorting, and individualized germline set (IGS) generation. Two rhesus monkeys (RMs #30734 and #31881) were vaccinated with Chiron HCV E1E2, as previously reported [44,45]. Their peripheral blood mononuclear cells (PBMCs) were collected, followed by single-cell sorting to isolate antigen-specific B cells and mAbs. In parallel, targeted single-molecule real-time (SMRT; Pacific Biosciences) long-read sequencing was performed on PBMC-derived genomic DNA from these same vaccinated RMs, generating an IGS. The Figure was created with BioRender.com. **b, c)** General information of five RM nAbs that were isolated from RMs #30734 and #31881. Their genetic germline allele features are assigned by KIMDB [38], IMGT [39], and MUSA database [48]. NA means no complete, matching germline allele was found in that database. The germline gene assignments were determined by three criteria: (1) matching of V and J gene usage, (2) identical CDR3 length, and (3) CDR3 nucleotide sequence homology >80%. **d)** Frequency of IGHV1-BSGF*02 (orange) and IGHV4-5ARF*17 (cyan) germline alleles. The Y-axis represents the detection frequency of these germline alleles among 106 monkeys from the MUSA database [48]. **e)** The V, D, and J germline allele assignment by our IG long-read sequencing and IGS. **f**) IGS reveals the closest putative IGHD gene for RM10-30 Ab. Sequence alignment was performed between mature Ab and identified IGHD germline gene by IGS. Sequence of IGHD gene and P/N nucleotide insertion region are in red and green, respectively. Dots denote the matching nucleotides with mature Ab sequence. **g, h)** HC amino acid sequence alignment of human mAb1382 [20] and RM Abs with IGHV1-69*10, IGHV1-138*01, and IGHV4-NL_5*01 germline alleles. The variant residues of RM Abs in comparison with encoded germline alleles are in bold and shaded in yellow. The CDRH3 basic, long side chain residues of IGHV1-138*01 encoded RM Abs are shaded in cyan. The key aromatic residue of CDRH3 RM5-16 is shaded in cyan. CDR1-3 length and sequence are defined according to Kabat numbering.

Fig. 2: Overall structure and affinity breadth of E2 with RM IGHV1-138*01 (RM11-73, RM11-48, RM10-30, RM1-36) and IGHV4-NL_5*01 RM5-16 Abs. **a)** E2 subregions are colored as in schematic, including HVR1 (aa 384-420), AS412 (aa 412-420), front layer (FL, aa 421–458), AS434 (aa 434-446), variable region 2 (VR2, aa 459–483), β-sandwich core (aa 484–517 and 535–568), CD81 binding loop (CD81BL, aa 518–534), variable region 3 (VR3, aa 569–579), post-variable 3 region (pVR3, aa 580–595), back layer (BL, aa 596–645), stem (aa 646–717), transmembrane (TM, aa 718-746). E2 FL, beta-sandwich (aa G504-Y508), CD81BL, and

BL subregions are colored in green, brown, salmon, and teal, respectively with the remainder of HK6a E2c3 colored grey. **b)** The crystal structures reveal all IGHV1-138*01 RM Abs target AR3 region (AR3: FL, N-terminus (aa 421-436), helix α1 C-terminus (aa 438-443), middle of β-sandwich (aa 504-508), CD81BL (aa 528-531), and helix α2 of BL (aa 614-622)). AR3 is highlighted in a blue dotted circle. HC and LC of RM Abs are shown in magenta and light orange, respectively. **c)** Buried surface area (BSA) analysis of human IGHV1-69, and RM IGHV1-138*01 HC and LC CDRs indicate the HC makes most of the E2 interactions. Structures of E2 with RM1-73 (PDB ID 9MNT), RM11-48 (9MNU), RM10-30 (9MNQ), RM1-36 (9MNS), AR3A (6BKB), AR3B (6BKC), AR3C (6UYD), AR3D (6BKD), AR3X (6URH), U1 (6WO3), HC11 (6WO4), HC1AM (6WOQ), 212.1.1 (6WO5), AT12009 (7T6X), mAb1198 (7RFB), mAb1382 (7RFC), HEPC3 (6MEI), HEPC74 (6MEH), RM2-01 (7JTF), and RM11-43 (7JTG) are used for BSA analysis. Left panel, BSA of HC and LC are in magenta and wheat, respectively. Right panel, BSA of CDRH1 (H1), CDRH2 (H2), CDRH3 (H3), CDRL1 (L1), CDRL2 (L2), and CDRL3 (L3) are in pink, magenta, purple, cyan, blue, and green, respectively. **d)** Crystal structure of IGHV4-NL_5*01 RM5-16 with HCV-1 E2ecto and HEPC46. HC and LC of HEPC46 are shown in deep olive and slate, respectively. **e)** BSA analysis of RM5-16 HC/LC CDRs upon binding to E2ecto. **f)** HCVpp neutralization of RM antibodies. RM antibodies (IgG) were tested for neutralizing activity against five HCVpp representing tier 1–4 resistance, along with the autologous virus HCV-1, at a single concentration of 25 μg/mL. AR3A was included as a positive control. The vesicular stomatitis virus envelope glycoprotein G (VSVpp) served as a negative control for nonspecific inhibition. Viral genotypes are shown in parentheses. All measurements were performed in triplicate, and mean values are displayed. Neutralization was defined using a 50% cutoff. **g)** BLI binding of all RM Fabs with HCV E2 proteins from different isolates and genotypes. The strongest binding is in red and no binding (NB) in dark green. The gradient yellow-green color indicates the mid-range binding affinity.

ARTICLE IN PRESS

Fig. 3: Interaction of RM Abs with HCV E2 proteins. a-e) Left panel, E2 epitope residues (with BSA > 0 Å^2 according to PISA analysis) are in green sticks, and E2 protein backbones are shown in grey tubes. Mid to right panel: E2 (green) detailed interactions with the RM Ab HC (magenta) and LC (gold). The E2 and RM Ab residues with BSA > 0 Å^2 and within 3.9 Å are labeled and shown as sticks. The H-bonds and charged interactions are in dashed black lines. The π-cation interaction is highlighted in an orange circle.

Fig. 4: RM IGHV1-138*01 Ab recognition of E2 mimics human elite neutralizers mAbs 1198 and 1382, while IGHV4-NL_5*01 RM5-16 has a new binding mode. **a)** The E2 from all complex structures were superimposed to compare the binding modes of RM Fabs and human mAbs 1198/1382, HEPC74, and AR3C. Fab backbones are colored with HC (magenta) and LC (light orange). E2 proteins are shown as surfaces with different colors of the subregions as in Fig. 2. **b-c)** The positions of H1-3 and L1-3 of RM Abs on E2 proteins are compared with those of the human mAbs. H and L are shown in magenta and gold ribbon representations, respectively. **d-e)** The binding mode of RM Ab CDRH1-2 are compared with those of human mAbs 1198 and 1382. Key E2 residues are in green sticks and CDRH1-2 are in magenta sticks. Sequences of CDRH1,2 of RM Abs are defined in Figs. 1, 2. Bottom-right panel: CDRH1-2 sequence alignment of mature RM Abs, their inferred germline counterparts, and corresponding human germline homologs. Residues conserved between RM and human germlines are boxed in cyan; somatic mutations in the mature RM Abs relative to their RM germlines are highlighted in yellow.

ARTICLE IN PRESS

Fig. 5: Interaction patterns of AR3-targeting Abs and CD81 receptor in complex with HCV E2. **a)** The IGHVs of AR3-targeting Fab are compared with host CD81 receptor bound to E2 (PDB ID 7MWX). The CD81BL in all Fab-E2 complexes is retracted, which differs from its extended conformation in the CD81 receptor-E2 complex [23]. The color of E2 subregions and HC and LC of Fab are as in Fig. 2. CD81 is shown in a lime cartoon representation. **b)** The positions of H1-3 of AR3-targeting Abs relative to CD81. CDRH3 loops of RM IGHV1-138*01 Abs (RM1-73, RM11-48, RM10-30, RM1-36, RM2-01, and RM11-43) and human IGHV1-69 Abs (mAbs 1198 and 1382) bind to similar regions as CD81 loop D. Additionally, CDRH1-2 of these Abs interact with a region on E2 targeted by CD81 helix C. In contrast, the CDRH3 loop of RM5-16 binds near the location bound by CD81 helix C. H1-H3 and CD81 are shown in magenta and lime ribbons, respectively. **c)** The binding approach and interaction of Fab CDRH3 and CD81 loop D with E2. The key interface residues of E2, Fab CDRH3, and CD81 loop D are in green, magenta, and yellow sticks, respectively. CDRH3 and CD81 loop D are shown as backbone ribbons with stick side chains. E2 proteins are shown as light grey cartoons.

ARTICLE IN PRESS

a)

Chiron HCV E1E2
Chen *et al.*, 2020
Chen *et al.*, 2021
#30734 #31881
Blood
PBMC
single cell sorting
antigen specific
antigen-antibody structure for RM2-01, RM11-43

This study
PBMC
long read sequencing of IG-enriched DNA
Individualized germline set (IGS)
MUSA
KIMDB
IMGT
antigen-antibody structure for RM1-73, RM11-48, RM10-30, RM1-36, RM5-16

Genetic and structural feature of RM AR3-targeting antibodies

b)

NHP mAb	Animal	Neutralization	KIMDB V gene	KIMDB D gene	KIMDB J gene	IMGT V gene	IMGT D gene	IMGT J gene
RM1-73	#30734	Broad	IGHV1-138*01_S6073	IGHD6-36*01	IGHJ6-6*01	IGHV1-151*01	IGHD6-25*01	IGHJ6*01
RM11-48	#30734	Broad	IGHV1-138*01_S6073	IGHD6-36*01	IGHJ5-4*02	IGHV1-151*01	IGHD6-25*01	IGHJ5-1*02
RM10-30	#30734	Broad	IGHV1-138*01_S6073	NA	IGHJ5-4*02	IGHV1-151*01	NA	IGHJ5-1*02
RM1-36	#30734	Limit	IGHV1-138*01_S6073	IGHD5-32*01_S2987	IGHJ5-4*02	IGHV1-151*01	NA	IGHJ5-1*02
RM5-16	#31881	Limit	IGVH4-NL_5*01_S5158	IGHD1-8*01	IGHJ5-5*01	NA	IGHD1-44*02	IGHJ5-2*02

c)

NHP mAb	MUSA V gene	MUSA D gene	MUSA J gene
RM1-73	IGHV1-BSGF*02	IGHD1-DWON*05	IGHJ6-GAIV*01
RM11-48	IGHV1-BSGF*02	IGHD1-DWON*05	IGHJ4-MIQL*02
RM10-30	IGHV1-BSGF*02	NA	IGHJ4-MIQL*02
RM1-36	IGHV1-BSGF*02	IGHD18-X76R*05	IGHJ4-MIQL*02
RM5-16	IGHV4-5ARF*17	IGHD1-BAAI*02	IGHJ4-XD6S*03

d)

Total of 106 Monkeys
% of animals from MUSA cohort
0 10 20 30 40 50 60
IGHV1-BSGF*02
IGHV4-5ARF*17

e)

NHP mAb	IGS Source V gene	IGS Source D gene	IGS Source J gene	% Identity V	% Identity D	% Identity J	Alignment Length V	Alignment Length D	Alignment Length J
RM1-73	IGHV4-2HU4*01	IGHD9-T3KD*01	IGHJ6-GAIV*01	98.6	93.8	100	296	16	48
RM11-48	IGHV4-2HU4*01	IGHD9-T3KD*01	IGHJ5-MIQL*02	98.0	100	100	296	19	49
RM10-30	IGHV4-2HU4*01	IGHD1-C7RJ*01_i0gga_c1g_t2g_a3g_c5g_t6a_g8a_12g12	IGHJ5-MIQL*02	97.6	100	100	296	6	46
RM1-36	IGHV4-2HU4*01	IGHD18-X76R*05	IGHJ5-MIQL*02	97.0	90	100	296	20	47
RM5-16	IGHV6-2BVG*16	IGHD25-BAAI*02	IGHJ5-XD6S*03	96.0	100	98	296	7	49

f)

```
                                                   Q  A  E  G  S  L  R  L  R
Mature RM10-30 CDRH3                               CCAGGCAGAGGGAAGCTTGCGACTCCGA
IGHD1-C7RJ*01_i0gga_c1g_t2g_a3g_c5g_t6a_g8a_12g12         G......GAGAGGAG
```

100% identity (6/6 bp)
Putative IGHD gene
P/N nucleotide insertions

g)

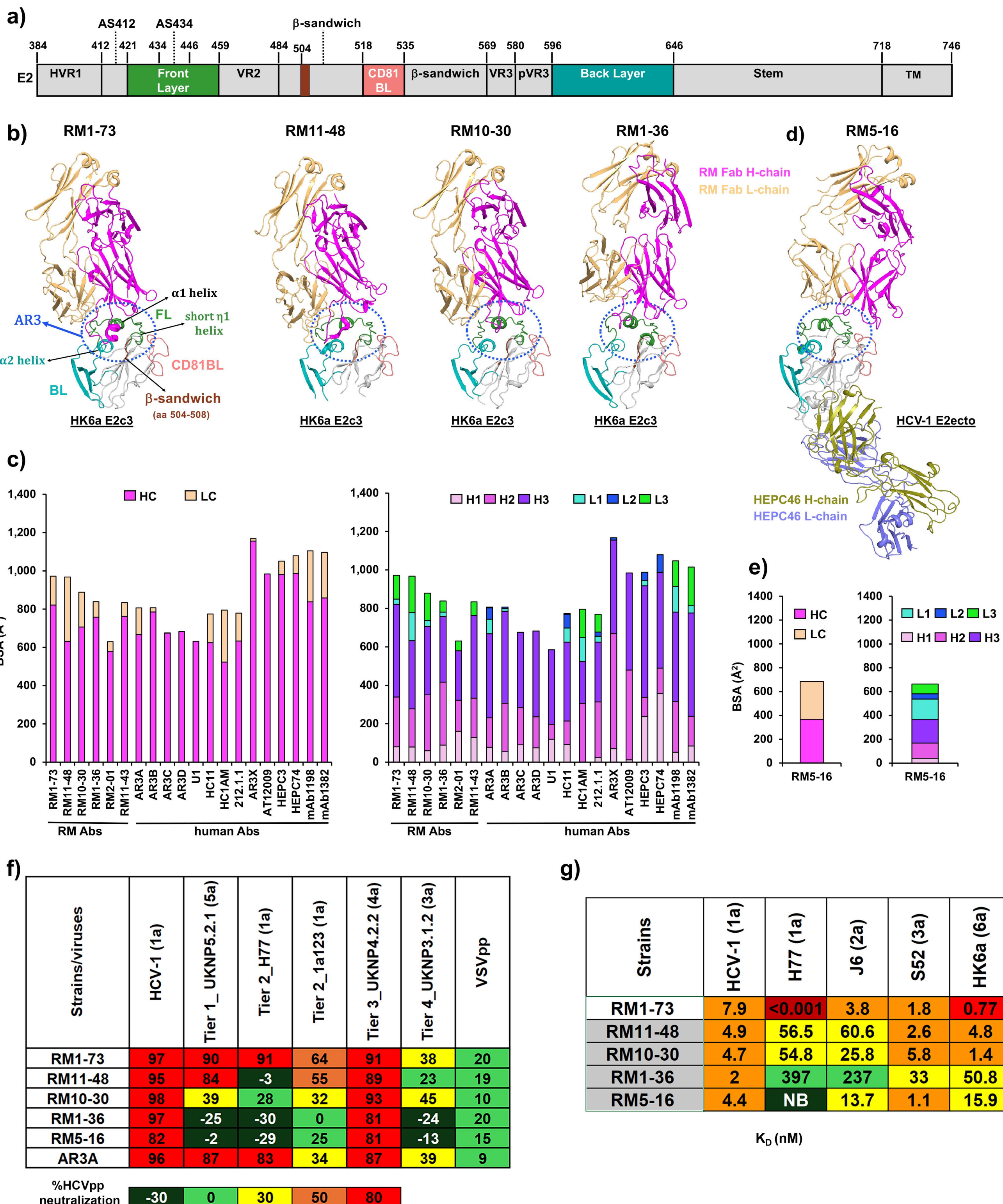
```
             HFR1                          CDRH1      HFR2          CDRH2                  HFR3                                    CDRH3
IGHV1-69*10  QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGGIIPILGIANYAQKFQGRVTITADKSTSTAYMELSSLRSEDTAVYYCAR-----------------------
mAb1382      QVQLVQSGAEVKKPGSSVKVSCTTSGGSFGRDAVSWVRQAPGQGLEPMGGIIPIFGTSSYAQTFQGRVTFTAGKSTSTAYMELSSLKSEDTAVYYCVRWSFGDYGLFLT------EGDYW

IGHV1-138*01 QVQLVQSGAEVKKPGASVKLSCKASGFTFSIYAISWVRQAPGQGLEWMGEIIPLVGITNYAQKFQGRVTITADTSTSTAYMELSSLRSEDTAVYYCARG----------------------
RM1-36       QVQLVQSGAEVKKPGASVKLSCKASGFTFSIYAICWVRQAPGQGLEWMGEIIPRVGLTTYAQKFQGRVTITADTSTTTVYMELSSLRSEDTAVYYCVRGPRYSEY-----SYRASWFDVW
RM1-73       QVQLVQSGAEVKKPGASVRLSCKASGFTFSISAIGWVRQAPGQGLEWMGEIIPLVGITNYAQKFQGRVTITADTSTSTAYMDLSSLRSEDTAVYYCARGDSKLQSVAAGGTEGGNGLDSW
RM10-30      QVQLVQSGAEVKKPGASVKLSCKASGFTFSIDAISWVRQAPGQGLEWMGEIIPRVGITVYAQKFHDRVTITADTSTSTAYMELSSLRSEDTAVYYCARGQAEG-------SLRLRWFDVW
RM11-48      QVQLVQSGAEVKKPGASVKLSCKASGFTFSVYAITWVRQAPGQGLEWMGEIIPLVGITNYAQKFQGRVTITADTSTNTAYMELTSLRSEDTAVYYCARGVTRIAAA-----GTGNWFDVW
RM2-01       QVQLVQSGAEVKKPGASVKLSCKASGFTFRIYAISWVRQAPGQGLEWMGEIVPLGGITNYAQNFQGRAAITADTSINTAYMELSSLRSDDTAVYYCARVRTALVG------PREGIFDFW
RM11-43      QVQLVQSGAEVKKPGASVKLSCKASGFTFRIYAISWVRQAPGQGLEWMGEIIPLGGITNYAQKFQGRVTITADTSTSTAYMELSSLRSEDTAVYYCARGETKMITVLAGTT--GGGLDSW
```

*IGHV1-138*01_S6073

h)

```
              HFR1                          CDRH1      HFR2          CDRH2                  HFR3                                    CDRH3
IGHV4-NL_5*01 QVQLQESGPGLVKPSETLSLTCAVSGGSISSNYWSWIRQSPGKGLEWIGYIYGGSGSTSYNPSLKSRVTISTDTSKNQFSLKLSSVTAADTAVYYCAR-----------------------
RM5-16        QLQLQESGPGLMKPSETLSLTCAVSGGSISSHYWSWIRQSPGKGLEWIGYIYGGTGRTNYNPSLKSRVTISTDTSKNQFSLNLSSVTAADTAMYYCARYPGSHW---------ENSLDVW
```

*IGHV4-NL_5*01_S5158



f)

Strains/viruses	HCV-1 (1a)	Tier 1_ UKNP5.2.1 (5a)	Tier 2_H77 (1a)	Tier 2_1a123 (1a)	Tier 3_UKNP4.2.2 (4a)	Tier 4_UKNP3.1.2 (3a)	VSVpp
RM1-73	97	90	91	64	91	38	20
RM11-48	95	84	-3	55	89	23	19
RM10-30	98	39	28	32	93	45	10
RM1-36	97	-25	-30	0	81	-24	20
RM5-16	82	-2	-29	25	81	-13	15
AR3A	96	87	83	34	87	39	9

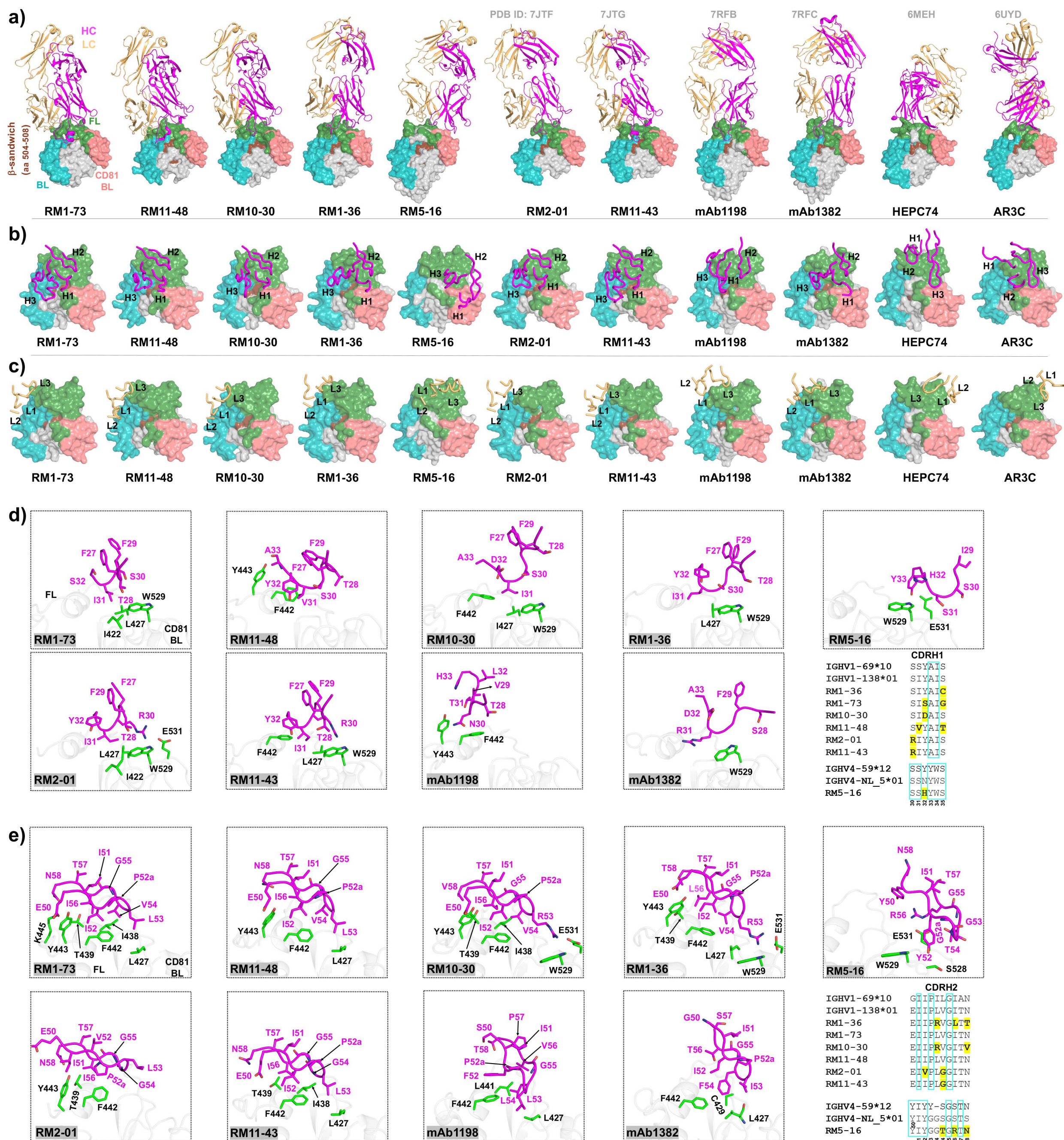
%HCVpp neutralization: -30 | 0 | 30 | 50 | 80

g)

Strains	HCV-1 (1a)	H77 (1a)	J6 (2a)	S52 (3a)	HK6a (6a)
RM1-73	7.9	<0.001	3.8	1.8	0.77
RM11-48	4.9	56.5	60.6	2.6	4.8
RM10-30	4.7	54.8	25.8	5.8	1.4
RM1-36	2	397	237	33	50.8
RM5-16	4.4	NB	13.7	1.1	15.9

K_D (nM)

a)
CDRH1,2
CDRH3
CDRL1,3
RM1-73
b)
RM11-48
c)
CDRL1,3 and LFR1
RM10-30
d)
CDRL3
RM1-36
e)
CDRL1, 2, 3, and LFR3
RM5-16


a)
b)
c)
d)
e)
PDB ID: 7JTF
7JTG
7RFB
7RFC
6MEH
6UYD
HC
LC
FL
β-sandwich (aa 504-508)
BL
CD81 BL
RM1-73
RM11-48
RM10-30
RM1-36
RM5-16
RM2-01
RM11-43
mAb1198
mAb1382
HEPC74
AR3C
CDRH1
IGHV1-69*10 SSYAIS
IGHV1-138*01 SIYAIS
RM1-36 SIYAIC
RM1-73 SISAIG
RM10-30 SIDAIS
RM11-48 SVYAIT
RM2-01 RIYAIS
RM11-43 RIYAIS
IGHV4-59*12 SSYYWS
IGHV4-NL_5*01 SSNYWS
RM5-16 SSHYWS
CDRH2
IGHV1-69*10 GIIPILGIAN
IGHV1-138*01 EIIPLVGITN
RM1-36 EIIPRVGLTT
RM1-73 EIIPLVGITN
RM10-30 EIIPRVGITV
RM11-48 EIIPLVGITN
RM2-01 EIVPLGGITN
RM11-43 EIIPLGGITN
IGHV4-59*12 YIYY-SGSTN
IGHV4-NL_5*01 YIYGGSGSTS
RM5-16 YIYGGTGRTN

a)

CD81BL (extended)
FL
BL
β-sandwich (aa 504-508)
CD81BL (retracted)

CD81 (7MWX) RM1-73 RM11-48 RM10-30 RM1-36 RM5-16 mAb1198 (7RFB) mAb1382 (7RFC)

b)

Helix A
Helix B
Helix C
Helix D
Loop D
H1 H2 H3

CD81 (7MWX) RM1-73 RM11-48 RM10-30 RM1-36 RM5-16 mAb1198 (7RFB) mAb1382 (7RFC)

180°

CD81 (7MWX) RM1-73 RM11-48 RM10-30 RM1-36 RM5-16 mAb1198 (7RFB) mAb1382 (7RFC)

c)

RM1-73
S97 G100e I422 A100d Y443 L99 W420 A444 S419 K98 V100b Y613 G418 P612 N417 Y507

RM11-48
N100f F442 A100 Y443 I99 I422 L441 A444 A100b S440 R98 Y613 G100c P612

RM10-30
R100d Y443 F442 L100a I422 A444 L441 S440 R100b Y613

RM1-36
R97 Y443 F442 Y100a I422 A444 S99 Y98 H421 S440 L441 E100 Y613 P612

RM2-01 (7JTF)
F442 T97 Y443 E100d A98 L99 L441 Y613

RM11-43 (7JTG)
I422 T97 I100 L427 F442 M99 T425 A100d L441 L100c K98 V100b Y613 P505 Y507

mAb1198 (7RFB)
D95 W100g Y443 W99 Q100f F442 I422 L441 T444 P100d H100b W420 Y613

mAb1382 (7RFC)
W95 T100e Y443 D99 F97 H445 F442 F100c T444 L100b G100a T425 A440 Y100 Y613 L615

CD81 loop D (7MWX)
Q89 D86 S80 F442 L82 F83 H445 I422 Y443 H421 N81 Y613 Y507 P612

CDRH3
CD81
E2