

Engineered HCV E2 Immunogens and Related Vaccine Compositions

CROSS-REFERENCE TO RELATED APPLICATIONS

- 5 **[0001]** The subject patent application claims the benefit of priority to U.S. Provisional Patent Application Number 62/879,100 (filed July 26, 2019; now pending). The full disclosure of the priority application is incorporated herein by reference in its entirety and for all purposes.

STATEMENT OF GOVERNMENT SUPPORT

- 10 **[0002]** This invention was made with government support under contracts AI129698, AI125078, AI123861, AI079031 and AI106005 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

- 15 **[0003]** Hepatitis C virus (HCV) infects 1-2% of the world population and poses a major health burden that leads to ~500,000 deaths annually and an estimated 1.5-2 million new infections each year. The opioid epidemic, causing over 70,000 overdose-related deaths in 2017 alone, is directly contributing to the rapid rise of HCV infection
- 20 in North America. Most HCV patients (75–85%) will develop a chronic infection resulting in hepatocellular carcinoma, cirrhosis, and other severe liver diseases. Although direct-acting antiviral (DAA) therapies are expected to increase the HCV cure rate, challenges remain because diagnosis is usually at a late stage after liver damage. DAA treatment cannot prevent HCV reinfection nor reduce the risk of liver cancer in
- 25 advanced liver disease and resistance may emerge. Indeed, increased HCV-associated mortality and new infections in injection drug users (IDU) highlights the urgency in developing an effective prophylactic vaccine to combat HCV.
- 30 **[0004]** A major challenge in HCV vaccine development is how to elicit a broadly protective immune response to overcome the high genetic diversity of seven major HCV genotypes and more than 86 subtypes. Moreover, rapid mutation leads to viral quasispecies in infected individuals that result in immune escape. Notwithstanding,

spontaneous viral clearance in 20-30% of acutely infected patients suggests that chronic HCV infection is preventable if an effective immune response can be induced by vaccination.

- [0005]** Despite substantial progresses in vaccine design, there are still needs in the medical field for more effective and potent vaccine immunogens, e.g., for preventing infections from non-viral or viral pathogens (e.g., HCV infection). The present invention addresses unmet needs in the art.

SUMMARY OF THE INVENTION

- [0006]** In one aspect, the invention provides modified HCV E2 ectodomain polypeptides. These polypeptides contain an engineered E2 ectodomain sequence that is redesigned (i.e., additionally modified) by at least one of the following: (a) a truncation in the VR2 disordered region and (b) a deletion in the β -sandwich loop connecting β sheets 6 and 7 of the β -sandwich domain. In some embodiments, the engineered E2 ectodomain sequence prior to the additional modification comprises SEQ ID NO:4, a conservatively modified variant or a substantially identical sequence thereof. In some of these modified HCV E2 ectodomain polypeptides, truncation of the VR2 disordered region comprises replacing residues 41-61 of SEQ ID NO:4, CPERLASCSSGSGCWHYPPRPC (SEQ ID NO:6), with CPERASGHYPPRPC (SEQ ID NO:10). In some modified HCV E2 ectodomain polypeptides, truncation of the VR2 disordered region comprises replacing residues 41-61 of SEQ ID NO:4, CPERLASCSSGSGCWHYPPRPC (SEQ ID NO:6), with a sequence set forth in CXXXXXXHYPPRPC (SEQ ID NO:8) or CXXXXXXHYPPRPC (SEQ ID NO:9), wherein X is any amino acid residue. In some of these embodiments, XXXXXX (SEQ ID NO:75) in SEQ ID NO:8 is QNWDEP (SEQ ID NO:11), KVNIDP (SEQ ID NO:12), EKVEEL (SEQ ID NO:13), PDENMK (SEQ ID NO:14), or KREEKM (SEQ ID NO:15). In some other embodiments, XXXXXX (SEQ ID NO:76) in SEQ ID NO:9 is PKTEV (SEQ ID NO:16), KRVDI (SEQ ID NO:17), PSDMV (SEQ ID NO:18), PNEEE (SEQ ID NO:19), or KKEIR (SEQ ID NO:20).
- [0007]** In some modified HCV E2 ectodomain polypeptides of the invention, deletion of the β -sandwich loop sequence (SEQ ID NO:21) from the engineered E2 ectodomain sequence SEQ ID NO:4 comprises deletion of one or more residues that

form the tip of the β -sandwich loop. In some of these embodiments, residues that form the tip of the β -sandwich loop are residues 543-546 from SEQ ID NO:4. In some of these embodiments, residues 543-546 from SEQ ID NO:4, RRPL (SEQ ID NO:22), are deleted.

- 5 **[0008]** Some modified HCV E2 ectodomain polypeptides of the invention contain both truncation in the VR2 disordered region and deletion in the β -sandwich loop. In some of these embodiments, the modified HCV E2 ectodomain polypeptide comprises an amino acid sequence as set forth in any one of SEQ ID NOs:26-38, a conservatively modified variant or a substantially identical sequence thereof.
- 10 **[0009]** In some related aspects, the invention provides polynucleotides that encode the modified HCV E2 ectodomain polypeptides described herein (polypeptides set forth in SEQ ID NOs:26-38), vectors or expression constructs that harbor such a polynucleotide sequence, as well as pharmaceutical compositions that contain a modified HCV E2 ectodomain polypeptide of the invention.
- 15 **[0010]** In another aspect, the invention provides vaccine compositions that contain a modified HCV E2 ectodomain polypeptide described herein that is displayed on the surface of a self-assembling nanoparticle. In some embodiments, the C-terminus of the modified HCV E2 ectodomain polypeptide is fused to the N-terminus of subunit of the self-assembling nanoparticle via a linker sequence. In some of these nanoparticle
- 20 vaccines, the linker sequence connecting the modified HCV E2 ectodomain polypeptide to the nanoparticle subunit comprises (GGGGS)₂ (SEQ ID NO:42). In various embodiments, the modified HCV E2 ectodomain polypeptide displayed on the nanoparticle comprises an amino acid sequence as set forth in any one of SEQ ID NOs:26-38, a conservatively modified variant or a substantially identical sequence
- 25 thereof. In any of these embodiments, the employed self-assembling nanoparticle can have a subunit sequence as shown in SEQ ID NO:39 (E2p), SEQ ID NO:40 (I3-01), SEQ ID NO:41 (ferritin), a conservatively modified variant or a substantially identical sequence thereof.
- 30 **[0011]** In another related aspect, the invention provides polynucleotide sequences that encode a fusion protein containing a modified HCV E2 ectodomain polypeptide described herein and a self-assembling nanoparticle subunit. In the encoded fusion protein, typically the modified HCV E2 ectodomain polypeptide is fused at its C-

terminus to the N-terminus of the self-assembling nanoparticle subunit. Another related aspect of the invention is directed to vectors or expression constructs harboring such a polynucleotide sequence. In another aspect, the invention provides pharmaceutical compositions that contain a nanoparticle particle displayed modified HCV E2
5 ectodomain polypeptide described herein, and optionally a pharmaceutically acceptable carrier.

[0012] In still another aspect, the invention provides methods for treating or preventing HCV infection in a subject afflicted with or at risk of developing HCV infection. These methods of the invention entail administering to the subject a
10 pharmaceutical composition that contains a modified HCV E2 ectodomain polypeptide described herein or a nanoparticle vaccine containing the polypeptide.

[0013] A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification and claims.

15

DESCRIPTION OF THE DRAWINGS

[0014] Fig. 1 shows rational design of HCV E2 cores. (A) Schematic representation of HCV E2 (amino acids 384-746) showing structural components: variable regions (VRs), antigenic site 412 (AS412), front layer, β -sandwich, CD81
20 binding loop, back layer, stalk trans-membrane (TM) region, and *N*-linked glycans and conserved disulfide bonds. Sequence alignment of the design regions between E2 and E2mc3 (see below) is shown. Sequences of VR2 disordered region (residues 452-494) (SEQ ID NO:5) and β -sandwich loop (residues 540-550) (SEQ ID NO:21) of HCV H77 isolate are shown. (B) Structure-based design of minimal E2 cores. Left: Structure of
25 H77 E2c (PDB ID: 4MWF) with shortened VR2 loop modeled by LOOPY loop prediction program. The redesigned β -sandwich loop and the shortened VR2 disordered region are colored in magenta. Disulfide bonds, C494-C564 and C452-C620, which anchor VR2 to the back layer, are shown. Front layer, CD81 binding loop, and back layer are also labeled. Middle 1: Structure of H77 E2mc3 with tip-truncated β -sandwich
30 loop and further truncated VR2 disordered region (tVR2). Middle 2: root-mean-square fluctuation (RMSF) plot for redesigned tVR2 ensemble is shown with the major steps involved in the ensemble-based *de novo* protein design below. Right: Structure of H77

E2mc3 with five top-ranking tVR2 design variants (E2mc3 v1-v5) highlighted in a transparent molecular surface. (C) SEC profiles of E2mc3 and variants. Left: H77 E2mc3, v1-v5, and v6-v10. Right: HK6a E2mc3 and v1. (D) SDS-PAGE E2mc3 and variants (Left: H77; Right: HK6a). (E) EC₅₀ values of H77 (upper panel) and HK6a (lower panel) E2 cores binding to 12 HCV antibodies, including eight bNAbs (HCV1, HC33, HC84.1, AR3C, HEPC3, HEPC74, 212.1.1, and HC1AM), one NAb (AR2A), and three non-NAbs (AR1A, AR1B, and E1). E2 cores tested here include E2c3, E2mc3, and E2c3 variants (10 for H77 and 1 for HK6a). (F) Binding affinities (K_ds) of H77 and HK6a E2mc3 variants for six selected HCV antibodies. (G) Thermal stability of H77 and HK6a E2c3 and E2mc3 variants measured by DSC. Two thermal parameters, T_m and $\Delta T_{1/2}$, are listed for four H77 E2 cores and three HK6a E2 cores.

[0015] Fig. 2 shows structures of rationally designed HCV E2 cores. (A) Crystal structures of H77/HK6a E2mc3 indicate an overall similar fold to H77 E2c and HK6a E2c3 (PDB: 4MWF and 6BKB). (B) Superposition of the β -sandwich loop from the H77 E2mc3-v1 structure on the HK6a E2c3-Fab E1 complex confirming that loss of binding of E2mc3s to Fab E1 results from truncation of the β -sandwich loop. (C) Superposition of E2 of HK6a E2c3 (PDB 6BKB), H77 E2mc3-v1, H77 E2mc3-v6, and HK6a E2mc3-v1 on the structure of H77 E2c (PDB 4MWF) illustrating the conformation of the redesigned tVR2 (a.a. 452-494). The redesigned tVR2 regions of H77 E2mc3-v1 and HK6a E2mc3-v1 structures are fully modeled but only partly in the H77 E2mc3-v6 structure. (D) Superposition of the H77/HK6a E2mc3 structures to H77 E2c and HK6a E2c3 indicating similar conformation of the neutralization face with only local conformational changes for the redesigned VR2 E2s.

[0016] Fig. 3 shows rational design of self-assembling E2 core nanoparticles. (A) Schematic representation of HCV virion (top) and E2 core-based nanoparticle vaccine (bottom). For the HCV virion, single-stranded (SS)-RNA, capsid, membrane, and envelope glycoproteins E1 and E2 are labeled, while for the vaccine, optimized E2 core and nanoparticle carrier are labeled. (B) Colored surface models of nanoparticle carriers (top) and E2 core-based nanoparticle vaccines (bottom). Three nanoparticle carriers shown here are 24-meric ferritin (FR) and 60-meric E2p and I3-01. Nanoparticle size is indicated by diameter (in nanometers). (C) SEC profiles of H77 E2mc3-v1 nanoparticles obtained from a Superose 6 10/300 GL column. The particle fraction is

indicated by a dotted-line box. While both FR and I3-01 nanoparticles were produced in ExpiCHO cells, E2p nanoparticles were expressed in HEK293 F cells. (D) BN-PAGE of SEC-purified H77 E2mc3-v1 nanoparticles. (E) Negative stain EM images of SEC-purified H77 E2mc3-v1 nanoparticles. (F) EC₅₀ values of H77 (upper panel) and HK6a (lower panel) E2mc3-v1 nanoparticles binding to 12 HCV antibodies listed in Fig. 7C. (G) Antigenic profiles of H77 (left) and HK6a (right) E2mc3-v1 and three nanoparticles against six HCV antibodies. Sensorgrams were obtained from an Octet RED96 using an antigen titration series of six concentrations (3.57-0.11 μ M by twofold dilution for E2mc3-v1 and 52.08-1.63 nM by twofold dilution for nanoparticles) and quantitation biosensors, as shown in Fig. 10H and I. The peak values at the highest concentration are listed in the matrix. Higher color intensity indicates greater binding signal measured by Octet.

[0017] Fig. 4. shows immunogenicity of newly designed E2 cores and nanoparticles in mice. (A) Schematic representation of the mouse immunization protocol. In study #1, mice were immunized with H77 E2mc3-v1 (group 1), H77 E2mc3-v1-10GS-FR (group 2), and H77 E2mc3-v1-10GS-E2p (group3). In study #2, mice were immunized with HK6a E2mc3-v1 (group 1), HK6a E2mc3-v1-10GS-E2p (group 2), and HK6a/H77 E2mc3-v1-10GS-E2p mix (group 3). (B) Longitudinal analysis of E2-specific antibody titers in immunized mouse sera at weeks 2, 5, 8 and 11. Top panel: EC₅₀ values calculated from ELISA binding of mouse sera in study #1 to the coating antigen, H77 E2mc3-v1. Bottom panel: EC₅₀ values calculated from ELISA binding of mouse sera in study #2 to the coating antigens HK6a E2mc3-v1 (groups 1-3) and H77 E2mc3-v1 (group 3). The *P*-values were determined by an unpaired *t* test in Prism and are labeled on the plots, with (*) indicating the level of statistical significance. Detailed serum ELISA data is shown in Fig. 11, A-D. (C) Mouse serum neutralization in study #1. Top panel: Percent (%) neutralization of mouse sera against autologous H77 at weeks 2, 5, 8 and 11. Bottom panel: Percent (%) neutralization of mouse sera against heterologous HCV1, J6, and SA13 at the last time point, week 11. (D) Mouse serum neutralization in study #2. Percent (%) neutralization of mouse sera against heterologous H77 at weeks 2, 5, 8 and 11. (E) Validation of the HCV pseudotyped particle (HCVpp) neutralization assay using five HCV bNABs and an

- HIV-1 bNAb (negative control). Percent (%) neutralization of all antibodies was determined at three concentrations, 10.0µg/ml, 1.0µg/ml, and 0.1µg/ml.
- [0018]** Fig. 5 shows patterns associated with HCV E2-specific B cell response in mouse immunization. (A) Schematic representation of the strategy used to analyze HCV E2-specific B cell response that combines antigen-specific bulk sorting of splenic B cells with next-generation sequencing (NGS) and antibodyomics analysis. (B) Statistical analysis of B cell sorting data obtained for group 1 (H77 E2mc3-v1 monomer) and group 3 (H77 E2mc3-v1-10GS-E2p nanoparticle) in study #1. Left: Frequency of E2-specific B cells. Right: Number of E2-specific B cells per million splenic cells. Five mice from group 1 (M1, M3, M5, M6 and M10) and five mice from group 3 (M5, M7, M8, M9, and M10) were randomly selected and analyzed. (C) Distribution of germline gene usage plotted for group 1 and group 3. Top panel: Germline V_H genes. Bottom panel: Germline V_κ genes. Statistical analysis of number of activated V_H/V_κ genes (≥1% of the total population) is shown on the far right. (D) Distribution of germline divergence or degree of somatic hypermutation (SHM) plotted for groups 1 and 3. For each group, percent (%) mutation is calculated at the nucleotide (nt) level for V_H (left) and V_κ (right). Statistical analysis of germline divergence is shown on the far right. (E) Distribution of CDR3 loop length plotted for groups 1 and 3. For each group, CDR3 length calculated at the amino acid (a.a.) level is shown for heavy (left) and light chains (right). Statistical analysis of root-mean-square fluctuation (RMSF) of CDR3 loop length, which is used as an indicator of how much the CDR3 loop length varies within the E2-specific antibodies from each animal. (F) Neutralization curves using purified IgG for groups 1 (left) and 3 (right) in study #1. Autologous H77 (1a) and heterologous SA13 (5a) were tested in HCVpp assays with a starting IgG concentration of 100µg/ml followed by a series of three-fold dilutions. Structural models of the immunogens are placed next to their neutralization curves. (G) Epitope mapping of polyclonal antibody sera from groups 1 and 3 in study #1. Surface model of a truncated E2 ectodomain (E2_{ECTO}) (PDB: 6MEI) is shown in the middle with the front layer (FL) and AS412. Statistical analysis of EC₅₀ titers of groups 1 and 3 against the FL probe (left) and the AS412 probe (right). Structural models of the designed nanoparticle probes are placed next to their plots. Epitopes on the nanoparticles are colored according to the truncated E2_{ECTO} model (PDB: 6MEI).

[0019] Fig. 6 shows sequence, structural, and computational analyses of HCV envelope glycoprotein E2 and E2 core design variants. (A) Sequence alignment of H77 E2 Δ TM, E2c, E2c3, and E2mc3 constructs (SEQ ID NOs:1-4, respectively). Regions of HVR1, VR2, VR3, and the β -sandwich loop are marked with lines. The E2c and E2c3 mutations are labeled, and the E2mc3 mutations noted with arrows. (B) Structures of H77 E2c (PDB: 4MWF) and 1b09 truncated E2 ectodomain (PDB: 6MEI). The protein chain is represented as a tube with the molecular surface color-coded as in Fig. 1A. (C) Structure of HK6a E2c3 bound to E1 and AR3A. The molecular surfaces of HK6a E2c3, E1, and AR3A are shown. A close-up view of four E1-interacting amino acids at the tip of the β -sandwich loop is shown as an insert. (D) Projection of the E1 epitope onto the E2c3 structure. Left: back layer, β -sandwich loop, and CD81 binding loop are shown as tubes within the transparent molecular surface. The tip of the β -sandwich loop is labeled with a rectangle; Right: E1-interacting amino acids are labeled on the solid molecular surface of E2. (E) Schematic overview of the interactions between E1 mAb HC and LC CDRs (SEQ ID NOs:44-49, respectively) and E2. E2 interacting residues are highlighted for (hydrogen bonds and hydrophobic interactions). (F) Conformational ensembles of redesigned E2mc3 tVR2 loop. E2mc3 structure is shown in ribbons and 1000 modeled loops are shown. Left: loop length #1 (13 a.a.) (SEQ ID NO:50); right: loop length #2 (12 a.a.) (SEQ ID NO:51). (G) Distribution of α root-mean-square (RMS) fluctuation plotted for the two redesigned tVR2 loop ensembles. (H) Five top-ranking designs and their energy scores for the two loop ensembles. Sequences shown are tVR2 sequence of E2mc3 (SEQ ID NO:10), and 10 fragment sequences (SEQ ID NOs:11-20, respectively) that replace the PERASG (SEQ ID NO:52) motif in the tVR2 sequence. (I) Sequence alignment of H77 E2, H77 E2c, and HK6a E2c3 (SEQ ID NOs:53-55, respectively). Amino acids that are disordered in the crystal structures of H77 E2c-AR3C (PDB: 4MWF) and HK6a E2c3-AR3A (PDB: 6BKB) complexes are noted.

[0020] Fig. 7 shows biochemical, biophysical, and antigenic characterization of E2 cores derived from H77(1a) and HK6a(6a). (A) SEC profiles of E2c3 proteins obtained from a Superdex 200 10/300 column after immunoaffinity (AR3A) purification. (B) SDS-page of E2c3 and E2mc3 proteins after immunoaffinity (AR3A) and SEC purification. (C) Antigenic sites and epitopes mapped onto the H77 E2c surface

(publications are listed). (D) ELISA binding of H77 E2c3 and E2mc3 variants to 12 HCV-specific antibodies. (E) EC₅₀ values of H77 E2 core constructs binding to 12 HCV-specific antibodies. (F) ELISA binding of HK6a E2 E2c3 and E2mc3 variants to 12 HCV-specific antibodies. (G) EC₅₀ values of HK6a E2 core constructs to 12 HCV-specific antibodies. In (E) and (G), EC₅₀ values were calculated for all ELISA plots in Prism except where the highest OD₄₅₀ value was below 0.1 or data fitting was ambiguous. (H) Octet binding of H77 E2mc3 variants to six HCV-specific antibodies. (I) Octet binding of HK6a E2mc3 variants to six HCV-specific antibodies. In (H) and (I), sensorgrams were obtained from an Octet RED96 instrument using a titration series of six concentrations (3.57-0.11 μ M by twofold dilution for all E2mc3 variants) and kinetics biosensors (see Methods). (J) K_D values of H77 and HK6a E2mc3 variants measured by Octet. (K) Differential scanning calorimetry (DSC) curves of selected E2 core constructs. Two thermal parameters, T_m and T_{1/2}, are labeled on the DSC profiles.

[0021] Fig. 8 shows crystal structures of H77 E2mc3 and HK6a E2mc3. (A) Conformational flexibility of the front layer 430-438 loop. The 430-438 loop in the H77 and HK6a E2c structures acquire different conformations yet maintain similar interactions with the Fab CDRH3 loop, indicating high flexibility of this region. (B) The conformation of the β -sandwich loop region (a.a. 540-550) in H77 E2c, HK6a E2c3, and H77/HK6a E2mc3.

[0022] Fig. 9 shows structural analysis of the redesigned tVR2 region. (A) 2Fo-Fc and unbiased omit electron density maps of the redesigned VR2 region (1 σ or 0.9 σ , respectively). (B) The intrinsic interactions of the redesigned tVR2 region of H77 E2mc3-v1, H77 E2mc3-v6, and HK6a E2mc3-v1 with the VR3 and post VR3 regions (a.a. 570-597). (C) Comparison of the H77 E2mc3-v1 and HK6a E2mc3-v1 indicating conformation changes of the redesigned VR2 and the neighboring VR3-back layer region (a.a. 565-608). (D) In the crystal structure of the truncated 1b09 E2 ectodomain (left), the full length VR2 wraps around VR3 to form the variable face. Superposition of H77 E2mc3-v1 structure on 1b09 E2 indicates only a minor influence of the VR2 redesign on the E2 overall fold (as indicated by aligning C452 and C494).

[0023] Fig. 10 shows biochemical, biophysical, and antigenic characterization of E2 core nanoparticles derived from H77(1a) and HK6a(6a). (A) SEC profiles of HK6a E2mc3-v1 nanoparticles based on FR, E2p and I3-01 obtained from a Superose 6 200

- Increase 10/300 column after immunoaffinity (AR3A) purification. (B) BN-PAGE of HK6a E2mc3-v1 nanoparticles based on FR, E2p and I3-01. (C) Negative-stain EM images of HK6a E2mc3-v1 nanoparticles based on FR, E2p and I3-01. (D) ELISA binding of H77 E2c3-v1 nanoparticles to 12 HCV-specific antibodies. (E) EC₅₀ values of H77 E2 E2c3-v1 nanoparticles binding to 12 HCV-specific antibodies. (F) ELISA binding of HK6a E2mc3-v1 nanoparticles to 12 HCV-specific antibodies. (G) EC₅₀ values of HK6a E2mc3-v1 nanoparticles with 12 HCV-specific antibodies. In (E) and (G), EC₅₀ values were calculated for all ELISA plots in Prism except where the highest OD₄₅₀ value was below 0.1 or data fitting was ambiguous. (H) Octet binding of H77 E2mc3-v1 nanoparticles to six HCV-specific antibodies. (I) Octet binding of HK7a E2mc3-v1 nanoparticles to six HCV-specific antibodies. In (H) and (I), sensorgrams were obtained from an Octet RED96 instrument using a titration series of six concentrations (3.57-0.11 μ M by twofold dilution for E2mc3-v1 and 52.08-1.63 nM by twofold dilution for E2mc3-v1 nanoparticles) and quantitation biosensors.
- 15 **[0024]** Fig. 11 shows murine antibody response during immunization at w2, w5, w8 and w11. (A) ELISA binding of H77 E2mc3-v1 to mouse sera from groups 1, 2, and 3 in study #1, which were immunized with H77 E2mc3-v1, E2mc3-v1-10GS-FR, and E2mc3-v1-10GS-E2p, respectively, at four time points. (B) EC₅₀ values of study #1 mouse sera binding to H77 E2mc3-v1 at four time points. (C) ELISA binding of HK6a E2mc3-v1 or H77 E2mc3-v1 to mouse sera from groups 1, 2, and 3 in study #2, which were immunized with HK6a E2mc3-v1, HK6a E2mc3-v1-10GS-E2p, and H77/HK6a E2mc3-v1-10GS-E2p mix, respectively, at four time points. Panels 1 and 2: sera from mice immunized with HK6a E2mc3-v1 (group 1) and HK6a E2mc3-v1-10GS-E2p (group 2) were tested against HK6a E2mc3-v1. Panels 3 and 4: sera from mice immunized with H77/HK6a E2mc3-v1-10GS-E2p mix (group 3) were tested against H77 E2mc3-v1 (panel 3) and HK6a E2mc3-v1 (panel 4). (D) EC₅₀ values of study #2 mouse sera binding to HK6a or H77 E2mc3-v1 at four time points.
- 25 **[0025]** Fig. 12 shows next-generation sequencing (NGS) analysis of bulk-sorted E2mc3-specific mouse splenic B cells. (A) SEC profile of biotinylated Avi-tagged H77 E2mc3-v1, termed E2mc3-v1-Avi-Biot, obtained from a Superdex 200 10/300 column, with the peak corresponding to biotin ligase labeled on the profile. (B) Summary of H77 E2mc3-v1-specific bulk sorting of mouse splenic B cells from study #1, groups 1

and 3. (C) Antibodyomics analysis of NGS data obtained for E2mc3-v1-sorted mouse splenic B cells. NGS data from groups 1 and 3, a total of 10 mice, were analyzed.

- [0026]** Fig. 13 shows analysis of mouse polyclonal serum antibody response. (A) Neutralization of HCV H77 and SA13 isolates by mouse IgG purified from the H77 E2mc3-v1-10GS-FR group in study #1. The HCVpp neutralization assays were performed with a starting IgG concentration of 100µg/ml and a series of 3-fold dilutions. The full neutralization curves were created to facilitate the comparison between different groups. (B) Design of epitope-specific probes for front layer (FL) (SEQ ID NOs:56-58) and AS412 (SEQ ID NOs:59-61). Left: epitope-scaffold design showing sequence alignment of the epitope, the designed epitope-scaffold, and the original scaffold obtained from the database search (* – match and • – similar with engineered disulfide bonds); Middle: structural model of designed epitope-scaffold, with the scaffold backbone shown in tan and the FL and AS412 epitopes. Right: molecular model of nanoparticle probe, with FL and AS412 epitopes. (C) ELISA binding of mouse sera from groups 1 and 3 in study #1, which were immunized with H77 E2mc3-v1 and E2mc3-v1-10GS-E2p, respectively, to the two epitope probes. Left: ELISA curves; Right: summary of EC₅₀ titers.

DETAILED DESCRIPTION

I Overview

- [0027]** Glycoproteins E1 and E2 form a heterodimer on the HCV envelope that mediates viral entry into host hepatocytes. E2 interacts with host cellular receptors CD81 and SR-B1 and is a major target for neutralizing antibodies (NAb). Crystal structures of an E2 core (E2c) from isolate H77 (genotype 1a) with a broadly neutralizing antibody (bNAb), AR3C, and a truncated E2 from isolate J6 (genotype 2a) bound to a non-NAb, 2A12, provided the first insight into immune recognition of HCV envelope glycoproteins and paved the way for structure-based design of antiviral drugs and vaccines. Diverse vaccine strategies such as viral vectors, DNA vaccines, virus-like particles (VLP), and recombinant E2 and E1E2 proteins have been explored, but no licensed vaccine is available to prevent HCV infection. Although recombinant E1, E2 and E1E2 glycoproteins have elicited NABs in animals and humans, neutralization breadth was limited and directed mainly to the immunodominant variable loops.

Therefore, HCV vaccine efforts should be focused on design and optimization of envelope glycoprotein-based antigens capable of eliciting a bNAb response.

[0028] Over the last decade, several rational vaccine design strategies for human immunodeficiency virus type-1 (HIV) have included epitope-focused and native Env trimer-based approaches, that aim to direct the immune response to bNAb epitopes either by grafting the epitope onto heterologous scaffolds, removing or suppressing immunodominant regions, or stabilizing Env structures. Another major advance was the development of self-assembling nanoparticles (NP) to present stabilized Env trimers and epitope-scaffolds as multivalent VLP vaccines. These general design elements can, in principle, be applied to a wide range of vaccine targets including HCV. Indeed, epitope-scaffolds have been designed for conserved E1 and E2 NAb epitopes, but with no reported *in vivo* data or little improvement in neutralization breadth. See, e.g., He *et al.*, *Sci. Rep.* **5**, 12501, 2015; Pierce *et al.*, *J. Virol.* **91**, e01032-01017, 2017; and Sandomenico *et al.*, *J. Virol.* **90**, 3745-3759, 2016.

[0029] The present invention is predicated in part from the inventors' studies to redesign HCV E2 based vaccine polypeptides. Specifically, the inventors first redesigned E2 core constructs by truncating the VR2 disordered region and the β -sandwich loop, followed by computationally optimizing the truncated VR2 loop (tVR2). The redesigned HCV E2 polypeptides were found to have improved yield, higher purity and increased thermostability. They also have stronger binding to broadly neutralizing antibodies and reduced binding to non-neutralizing antibodies. The inventors further displayed the redesigned HCV E2 polypeptides on nanoparticles of various sizes. The resulting nanoparticle vaccine compositions also demonstrated substantial yield, high purity, and enhanced antigenicity. In addition, mice were immunized with the nanoparticle vaccine constructs, and longitudinal serum analysis confirmed the superior immunogenicity of the nanoparticle vaccines. Moreover, statistical analysis was performed to validate the *in vivo* data, providing a rigorous foundation for future comparison of different types of HCV vaccine candidates.

[0030] In accordance with these studies, the invention provides novel HCV immunogen polypeptides that are derived from the engineered or redesigned HCV E2 core polypeptides described herein. Also provided in the invention are HCV vaccine compositions containing a displaying platform, including a self-assembling

- nanoparticle, that displays one or more of the redesigned HCV E2 polypeptides. Therapeutic applications of the redesigned HCV E2 polypeptides and the related nanoparticle vaccine compositions, e.g., treating or preventing HCV infections, are also provided in the invention.
- 5 **[0031]** Unless otherwise specified herein, the vaccine immunogens of the invention, the encoding polynucleotides, expression vectors and host cells, as well as the related therapeutic applications, can all be generated or performed in accordance with the procedures exemplified herein or routinely practiced methods well known in the art. See, e.g., *Methods in Enzymology*, Volume 289: Solid-Phase Peptide
- 10 Synthesis, J. N. Abelson, M. I. Simon, G. B. Fields (Editors), Academic Press; 1st edition (1997) (ISBN-13: 978-0121821906); U.S. Pat. Nos. 4,965,343, and 5,849,954; Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou ed., 2003); Davis et al., *Basic Methods in Molecular Biology*,
- 15 Elsevier Science Publishing, Inc., New York, USA (1986); or *Methods in Enzymology: Guide to Molecular Cloning Techniques Vol. 152*, S. L. Berger and A. R. Kimmerl Eds., Academic Press Inc., San Diego, USA (1987); *Current Protocols in Protein Science (CPPS)* (John E. Coligan, et. al., ed., John Wiley and Sons, Inc.), *Current*
- 20 *Protocols in Cell Biology (CPCB)* (Juan S. Bonifacino et. al. ed., John Wiley and Sons, Inc.), and *Culture of Animal Cells: A Manual of Basic Technique* by R. Ian Freshney, Publisher: Wiley-Liss; 5th edition (2005), *Animal Cell Culture Methods (Methods in Cell Biology, Vol. 57*, Jennie P. Mather and David Barnes editors, Academic Press, 1st edition, 1998). The following sections provide additional guidance for practicing the compositions and methods of the present invention.

25

II. Definitions

- [0032]** Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention pertains. The following references provide one of skill with a
- 30 general definition of many of the terms used in this invention: *Academic Press Dictionary of Science and Technology*, Morris (Ed.), Academic Press (1st ed., 1992); *Oxford Dictionary of Biochemistry and Molecular Biology*, Smith et al. (Eds.), Oxford

- University Press (revised ed., 2000); *Encyclopaedic Dictionary of Chemistry*, Kumar (Ed.), Anmol Publications Pvt. Ltd. (2002); *Dictionary of Microbiology and Molecular Biology*, Singleton et al. (Eds.), John Wiley & Sons (3rd ed., 2002); *Dictionary of Chemistry*, Hunt (Ed.), Routledge (1st ed., 1999); *Dictionary of Pharmaceutical Medicine*, Nahler (Ed.), Springer-Verlag Telos (1994); *Dictionary of Organic Chemistry*, Kumar and Anandand (Eds.), Anmol Publications Pvt. Ltd. (2002); and *A Dictionary of Biology (Oxford Paperback Reference)*, Martin and Hine (Eds.), Oxford University Press (4th ed., 2000). Further clarifications of some of these terms as they apply specifically to this invention are provided herein.
- 10 **[0033]** As used herein, the singular forms "a," "an," and "the," refer to both the singular as well as plural, unless the context clearly indicates otherwise. For example, "an Env-derived trimer" can refer to both single or plural Env-derived trimer molecules, and can be considered equivalent to the phrase "at least one Env-derived trimer."
- [0034]** As used herein, the terms "antigen" or "immunogen" are used
- 15 interchangeably to refer to a substance, typically a protein, which is capable of inducing an immune response in a subject. The term also refers to proteins that are immunologically active in the sense that once administered to a subject (either directly or by administering to the subject a nucleotide sequence or vector that encodes the protein) is able to evoke an immune response of the humoral and/or cellular type
- 20 directed against that protein. Unless otherwise noted, the term "vaccine immunogen" is used interchangeably with "protein antigen" or "immunogen polypeptide".
- [0035]** The term "conservatively modified variant" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or
- 25 essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For polypeptide sequences, "conservatively modified variants" refer to a variant which has conservative amino acid substitutions, amino acid residues replaced
- 30 with other amino acid residue having a side chain with a similar charge. Families of amino acid residues having side chains with similar charges have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine,

histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).

5 [0036] Epitope refers to an antigenic determinant. These are particular chemical groups or peptide sequences on a molecule that are antigenic, such that they elicit a specific immune response, for example, an epitope is the region of an antigen to which B and/or T cells respond. Epitopes can be formed both from contiguous amino acids or noncontiguous amino acids juxtaposed by tertiary folding of a protein.

[0037] Effective amount of a vaccine or other agent that is sufficient to generate a desired response, such as reduce or eliminate a sign or symptom of a condition or disease, such as a viral infection. For instance, this can be the amount necessary to inhibit viral replication or to measurably alter outward symptoms of the viral infection.

15 In general, this amount will be sufficient to measurably inhibit virus (for example, HCV) replication or infectivity. When administered to a subject, a dosage will generally be used that will achieve a target concentration that has been shown to be sufficient for in vitro inhibition of viral replication. In some embodiments, an "effective amount" is one that treats (including prophylaxis) one or more symptoms and/or underlying causes of

20 any of a disorder or disease, for example to treat HCV infection. In some embodiments, an effective amount is a therapeutically effective amount. In some embodiments, an effective amount is an amount that prevents one or more signs or symptoms of a particular disease or condition from developing, such as one or more signs or symptoms associated with HCV infection.

25 [0038] As used herein, a fusion protein is a recombinant protein containing amino acid sequence from at least two unrelated proteins that have been joined together, via a peptide bond, to make a single protein. The unrelated amino acid sequences can be joined directly to each other or they can be joined using a linker sequence. As used herein, proteins are unrelated, if their amino acid sequences are not normally found

30 joined together via a peptide bond in their natural environment(s) (e.g., inside a cell). For example, the amino acid sequences of bacterial enzymes such as *B.*

stearothermophilus dihydrolipoyl acyltransferase (E2p) and the amino acid sequences of HCV E2 glycoproteins are not normally found joined together via a peptide bond.

[0039] As used herein, “E1” refers to an envelope glycoprotein on hepatitis C virus. The E1 polypeptide is located in the HCV polyprotein immediately after the core protein at amino acid residues corresponding to residues 192-383 of the HCV genotype 1a, H77 isolate polyprotein.

[0040] As used herein, HCV envelope glycoprotein 2 (E2) refers to an envelope glycoprotein on hepatitis C virus. E2 is a type I transmembrane protein with an amino-terminal ectodomain connected to a carboxy-terminal transmembrane helix through an amphipathic, α -helical stem. The E2 polypeptide is located in the HCV polyprotein immediately after the E1 polypeptide at amino acid residues corresponding to residues 384-746 (SEQ ID NO:1) of the HCV genotype 1a, H77 isolate. See, e.g., Goffard et al., *Biochimie* 85, 295-301, 2003. Unless otherwise noted, reference to an E2 polypeptide herein also includes truncated forms thereof, such as soluble forms that are truncated at their C-terminus to remove the transmembrane domain, that retain a property, such as a binding property, and variants, such as those from different HCV isolates and quasispecies, including polypeptides that have at least 40%, 45%, 50%, 55%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity to the E2 polypeptide from HCV genotype 1a, isolate H77, set forth in SEQ ID NO:1.

[0041] Unless otherwise noted, E2 polypeptides include synthetic molecules prepared by translation of nucleic acids, proteins generated by chemical synthesis, such as syntheses that include ligation of shorter polypeptides, through recombinant methods, proteins isolated from HCV-infected human and non-human tissue and cells, chimeric E2 polypeptides and modified forms thereof. E2 polypeptides also include fragments or portions of E2 that are of sufficient length or include appropriate regions to retain at least one property of a full-length E2 polypeptide, such as the ability to bind to anti-E2 antibodies. E2 polypeptides also include those that contain chemical or posttranslational modifications and those that do not contain chemical or posttranslational modifications. Such modifications include, but are not limited to, pegylation, albumination, glycosylation, farnesylation, carboxylation, hydroxylation, phosphorylation, and other polypeptide modifications known in the art.

[0042] As used herein, the HCV E2 ectodomain (E2_{ECTO} or E2 Δ TM) refers to the soluble portion of the E2 polypeptide that is outside the viral membrane, i.e., the N-terminal portion of E2 protein that does not include the transmembrane helix. In the prototype H77 isolate, E2 Δ TM or E2_{ECTO} consists of amino acids 384-717 (SEQ ID NO:2) and is stabilized by nine conserved disulfide bonds. It contains three variable regions including hypervariable region 1 (HVR1, a.a. 384-411), VR2 (a.a. 460-484), and VR3 (a.a. 570-580) and is covered with ~11 *N*-linked glycans. As used herein, the VR2 disordered region refers to a specific E2 fragment which spans VR2, and is bordered by two cysteine residues that form two disulfide bonds for anchoring to the back layer and the β -sandwich domain, respectively. In the H77 isolate, the VR2 disordered region corresponds to residues 452-494, and the two anchoring disulfide bonds are C452-C620 and C494-C564.

[0043] As used herein, reference to an amino acid position in any of the mature, processed HCV proteins, such as E2, is made with numbering relative to the full length HCV polypeptide, and not with reference to the mature polypeptide. Thus, for example, reference to amino acid position 425 in the E2 polypeptide from HCV genotype 1a, isolate H77 corresponds to position 42 of the mature E2 polypeptide set forth in SEQ ID NO:1.

[0044] As used herein, corresponding residues refers to residues that occur at aligned loci. Related or variant polypeptides are aligned by any method known to those of skill in the art. Such methods typically maximize matches, and include methods such as using manual alignments and by using the numerous alignment programs available (for example, BLASTP) and others known to those of skill in the art. By aligning the sequences of polypeptides, one skilled in the art can identify corresponding residues, using conserved and identical amino acid residues as guides. For example, by aligning the sequences of E2 polypeptides from different HCV isolates, one of skill in the art can identify corresponding residues, using conserved and identical amino acid residues as guides. Corresponding positions also can be based on structural alignments, for example by using computer simulated alignments of protein structure.

[0045] As used herein, the term “hepatitis C virus,” “HCV,” or “HCVs” includes different viral genotypes, subtypes, quasispecies and isolates. It includes any noncytopathic RNA virus that has a single and positive-stranded RNA genome

belonging to the *Hepacivirus* genus of the Flaviviridae family. The term includes different isolates of HCV such as, without limitation, those having polypeptide sequences and accession numbers shown above, as well as any others in the NCBI database. Examples of different genotypes encompassed by this term include, without
5 limitation, genotype 1, 2, 3, 4, 5 and 6, as described in Simmonds et al. (*Hepatology* 42:962-973, 2005). Reference to HCV also includes those of any additional genotypes that are established. Examples of different subtypes include, without limitation, 1a, 1b, 1c, 2a, 2b, 2c, 2i, 2k, 3a, 3b, 3k, 4a, 4d, 4f, 5a, 6a, 6b, 6c, 6d, 6e, 6f, 6g, 6h, 6i, 6j, 6k, 6l, 6m, 6n, 6o, 6q, 6p and 6t. The term also includes cell culture HCVs (HCVcc) and
10 pseudotype HCVs (HCVpp), as well as HCV quasispecies. Various HCVs are described by Simmonds P. in *Genetic diversity and evolution of hepatitis C virus--15 years on*, *J Gen Virol* **85**:3173-3188 (2004) and Simmonds et al. in *Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes*, *Hepatology* **42**:962-973 (2005). HCV nucleotide sequences are known in the art. For
15 example, see Viral Bioinformatics Research Center (hcvdb.org) and the Hepatitis C Virus database (hcv.lanl.gov).

[0046] Immunogen as used herein refers to a protein or a portion thereof that is capable of inducing an immune response in a mammal, such as a mammal infected or at
20 risk of infection with a pathogen. Administration of an immunogen can lead to protective immunity and/or proactive immunity against a pathogen of interest.

[0047] Immune response refers to a response of a cell of the immune system, such as a B cell, T cell, or monocyte, to a stimulus. In some embodiment, the response is specific for a particular antigen (an "antigen-specific response"). In some embodiments, an immune response is a T cell response, such as a CD4⁺ response or a CD8⁺ response.
25 In some other embodiments, the response is a B cell response, and results in the production of specific antibodies.

[0048] Immunogenic composition refers to a composition comprising an immunogenic polypeptide that induces a measurable CTL response against virus expressing the immunogenic polypeptide, or induces a measurable B cell response
30 (such as production of antibodies) against the immunogenic polypeptide.

[0049] Sequence identity or similarity between two or more nucleic acid sequences, or two or more amino acid sequences, is expressed in terms of the identity or similarity

between the sequences. Sequence identity can be measured in terms of percentage identity; the higher the percentage, the more identical the sequences are. Two sequences are "substantially identical" if two sequences have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, 60% identity, optionally 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity over a specified region, or, when not specified, over the entire sequence), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Optionally, the identity exists over a region that is at least about 50 nucleotides (or 10 amino acids) in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides (or 20, 50, 200 or more amino acids) in length.

[0050] Homologs or orthologs of nucleic acid or amino acid sequences possess a relatively high degree of sequence identity/similarity when aligned using standard methods. Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman & Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins & Sharp, *Gene*, 73:237-44, 1988; Higgins & Sharp, *CABIOS* 5:151-3, 1989; Corpet et al., *Nuc. Acids Res.* 16:10881-90, 1988; Huang et al. *Computer Appls. in the Biosciences* 8, 155-65, 1992; and Pearson et al., *Meth. Mol. Bio.* 24:307-31, 1994. Altschul et al., *J. Mol. Biol.* 215:403-10, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.

[0051] The term "subject" refers to any animal classified as a mammal, e.g., human and non-human mammals. Examples of non-human animals include dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, and etc. Unless otherwise noted, the terms "patient" or "subject" are used herein interchangeably. Preferably, the subject is human.

[0052] The term "treating" or "alleviating" includes the administration of compounds or agents to a subject to prevent or delay the onset of the symptoms, complications, or biochemical indicia of a disease (e.g., an HCV infection), alleviating the symptoms or arresting or inhibiting further development of the disease, condition, or disorder. Subjects in need of treatment include those already suffering from the disease or disorder as well as those being at risk of developing the disorder. Treatment

may be prophylactic (to prevent or delay the onset of the disease, or to prevent the manifestation of clinical or subclinical symptoms thereof) or therapeutic suppression or alleviation of symptoms after the manifestation of the disease.

[0053] Vaccine refers to a pharmaceutical composition that elicits a prophylactic or therapeutic immune response in a subject. In some cases, the immune response is a protective immune response. Typically, a vaccine elicits an antigen-specific immune response to an antigen of a pathogen, for example a viral pathogen, or to a cellular constituent correlated with a pathological condition. A vaccine may include a polynucleotide (such as a nucleic acid encoding a disclosed antigen), a peptide or polypeptide (such as a disclosed antigen), a virus, a cell or one or more cellular constituents. In some embodiments of the invention, vaccines or vaccine immunogens or vaccine compositions are expressed from fusion constructs and self-assemble into nanoparticles displaying an immunogen polypeptide or protein on the surface.

[0054] Virus-like particle (VLP) refers to a non-replicating, viral shell, derived from any of several viruses. VLPs are generally composed of one or more viral proteins, such as, but not limited to, those proteins referred to as capsid, coat, shell, surface and/or envelope proteins, or particle-forming polypeptides derived from these proteins. VLPs can form spontaneously upon recombinant expression of the protein in an appropriate expression system. Methods for producing particular VLPs are known in the art. The presence of VLPs following recombinant expression of viral proteins can be detected using conventional techniques known in the art, such as by electron microscopy, biophysical characterization, and the like. See, for example, Baker et al. (1991) *Biophys. J.* 60:1445-1456; and Hagensee et al. (1994) *J. Virol.* 68:4503-4505. For example, VLPs can be isolated by density gradient centrifugation and/or identified by characteristic density banding. Alternatively, cryoelectron microscopy can be performed on vitrified aqueous samples of the VLP preparation in question, and images recorded under appropriate exposure conditions.

[0055] A self-assembling nanoparticle refers to a ball-shape protein shell with a diameter of tens of nanometers and well-defined surface geometry that is formed by identical copies of a non-viral protein capable of automatically assembling into a nanoparticle with a similar appearance to VLPs. Known examples include ferritin (FR), which is conserved across species and forms a 24-mer, as well as *B. stearothermophilus*

dihydrolipoyl acyltransferase (E2p), *Aquifex aeolicus* lumazine synthase (LS), and *Thermotoga maritima* encapsulin, which all form 60-mers. Self-assembling nanoparticles can form spontaneously upon recombinant expression of the protein in an appropriate expression system. Methods for nanoparticle production, detection, and
 5 characterization can be conducted using the same techniques developed for VLPs.

III. Redesigned HCV E2 core immunogenic polypeptides

[0056] The invention provides engineered or redesigned HCV E2 core polypeptides or proteins for producing HCV vaccines. E2 is a type I transmembrane
 10 protein with an amino-terminal ectodomain connected to a carboxy-terminal transmembrane helix through an amphipathic, α -helical stem. The E2 polypeptide is located in the HCV polyprotein immediately after the E1 polypeptide at amino acid residues corresponding to residues 384-746 of the HCV genotype 1a, H77 isolate. See, e.g., Goffard et al., *Biochimie* 85, 295-301, 2003. E2 is highly modified post-
 15 translationally with several (9-11) N-linked glycosylation sites and 20 cysteine residues that are conserved across all genotypes. Unless otherwise noted, amino acid residue numbering of HCV genotype 1a, H77 isolate is used herein as the reference.

[0057] E2 core refers to a portion of E2 that forms a 3-dimensional structure that is recognized by a broadly neutralizing antibody, AR3C Fab. See, e.g., Law et al., *Nat.*
 20 *Med.* 2008;14:25, 2008. The E2 core has a compact, globular domain structure, consisting mostly of β -strands and random coil with two small α -helices. The strands are arranged in two, perpendicular sheets (A and B), which are held together by an extensive hydrophobic core and disulfide bonds. The E2 core contains a central immunoglobulin (Ig)-like β -sandwich with front and back layers. As used herein, an
 25 E2 core polypeptide refers to an engineered polypeptide sequence that includes E2 residues or fragments that correspond to the key structural elements of E2 core in its 3-dimensional structure of E2. A prototype E2 core polypeptide, E2c (SEQ ID NO:3), is derived from E2 Δ TM of the H77 isolate. It consists of E2 amino acids 412-459, a GSSG linker replacing residues 460-485 of E2, and amino acids 486-645, with both the
 30 stalk region and the transmembrane (TM) region deleted. Relative to E2 sequence, E2c additionally contains mutations N448D and N576D. A further modified E2 core polypeptide, E2c3 (SEQ ID NO:4), is derived from E2c. Relative to E2c, the E2c3

polypeptide additionally contains a “GGPTDG” (SEQ ID NO:43) linker replacing E2 residues 569-597. Thus, it consists of amino acids 412-459, the “GSSG” linker, amino acids 486-568, the “GGPTDG” (SEQ ID NO:43) linker, and amino acids 598-645. Also, E2c3 only harbors the N448D substitution, with the N576D substitution being removed as a result of residues 569-597 being replaced by the GGPTDG (SEQ ID NO:43) linker. See, e.g., Kong et al., *Science*. 342:1090–1094, 2013; and Kong et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12768–12773, 2016.

[0058] Unless otherwise noted, the amino acid residue numbering of E2 or derivative thereof (e.g., E2ΔTM and E2 core), as well as the various redesigned HCV core polypeptides of the invention, is based on HCV genotype 1a, H77 isolate. Due to substantial sequence conservation among the E2 proteins from different HCV genotypes, corresponding amino acid residues bordering the various domains, regions and loops of an E2 protein from any other HCV genotype and isolate can be readily determined (e.g., by sequence alignment) or otherwise known in the art. For example, while E2 core of the H77 isolate spans residues 412-645, the E2 core of HCV genotype 2a, isolate J6 is comprised of residues spanning 412-645 based on the H77 numbering system (or 412–649 in the actual J6 polypeptide sequence). See, e.g., Kong et al., *Science*. 342:1090–1094, 2013; and Khan et al., *Nature* 509:381-384, 2014.

[0059] In general, the redesigned HCV E2 polypeptides of the invention contain one or more modifications in the VR2 disordered region and the β-sandwich loop. The original VR2 disordered region contains 43 residues, corresponding to residues 452-494 in the H77 isolate,

CPERLASCRRRLTDFAGWGPISYANGSGLDERPYCWHYPPRPC (SEQ ID NO:5). The corresponding region in the engineered E2c proteins known in the art, E2c and E2c3, contain a sequence of 21 residues, CPERLASCSSGSCWHYPPRPC (SEQ ID NO:6). Typically, the redesigned HCV E2 polypeptides of the invention contain a further shortened VR2 disordered region (tVR2) relative to that of the engineered E2c proteins known in the art, E2c and E2c3. In various embodiments, the further shortened VR2 disordered region (“tVR2”) in the redesigned HCV E2 polypeptides contains less than about 20, 19, 18, 17, 16 or less residues. In some preferred embodiments, the tVR2 contains less than about 15, 14, 13, 12 or less residues. As shown in the exemplified tVR2 sequence of redesigned HCV E2 polypeptide E2mc3,

CPERASGHYPRPC (SEQ ID NO:10), the further shortening of the VR2 disordered region of E2c and E2c3 involves deletions of various residues in SEQ ID NO:6. In some embodiments, the tVR2 of the redesigned HCV E2 polypeptides of the invention can have an amino acid sequence set forth in C(X)nHYPRPC (SEQ ID NO:7), wherein X is any amino acid residue, and n is between about 3 to about 10. In some of these 5 embodiments, the tVR2 contains 13 or 12 residues as set forth in CXXXXXXHYPRPC (SEQ ID NO:8) or CXXXXXXHYPRPC (SEQ ID NO:9), wherein X is any amino acid residue. Importantly, it was observed that the prior art E2 polypeptide such as E2c3 could not be displayed on nanoparticles with meaningful expression. In contrast, further 10 modification of the VR2 disordered region in the redesigned HCV E2 polypeptides of the invention enables satisfactory nanoparticle display of the polypeptides, as exemplified herein.

[0060] In addition to the shortened length, modification of the VR2 disordered region can additionally involve optimization of the sequence to further stabilize the 15 HCV E2 polypeptides. Optimization of the tVR2 sequence can be performed using ensemble-based de novo protein design as described in, e.g., Kong et al., Nat. Commun. 7:12040, 2016. Specific guidance for optimizing tVR2 sequence by selecting random sequences, performing simulated annealing and identifying lowest-energy sequences, is also provided in the Examples herein. In some of the redesigned HCV E2 polypeptides 20 of the invention, the tVR2 sequence is derived from optimization of the sequence as set forth in any one of SEQ ID NOs:7-9. As exemplification with redesigned HCV E2 polypeptides set forth in SEQ ID NOs:27-38, the optimized tVR2 sequence can be any of SEQ ID NOs:65-74, a substantially identical sequence or a conservatively substituted variant thereof.

[0061] Additionally or alternatively to the modified VR2 disordered region, the 25 redesigned HCV E2 polypeptides of the invention also contain a modification in the β -sandwich domain. As shown in Figure 1A, the β -sandwich domain is composed of two parts, residues 485-518 and 536-569, in the linear sequence HCV E2. These two parts interact with each other and form the β -sandwich domain in the 3-dimensional 30 structure. This β sandwich domain is composed of 7 β strands, where the CD81 binding loop (residues 519-535) is connecting the two parts of the domain between β strand 5 and 6. Modification of the β -sandwich domain in the redesigned HCV E2

- polypeptides of the invention relates to a β -sandwich loop, H77 isolate E2 residues 540-550 (NNTRPPLGNWF; SEQ ID NO:21), that connects β strand 6 to 7 in the β -sandwich domain. This modification is intended to disrupt the binding site on E2 that is recognized by non-neutralizing antibodies, thereby focusing immune response to
- 5 broadly neutralizing antibodies. The modification involves deletion of one or more residues in the β -sandwich loop sequence that form the tip in the 3-D structure. In the H77 isolate, these are residues 543-546 (RPPL; SEQ ID NO:22). Optionally, one or more of the adjacent residues may also be deleted. Thus, in various embodiments, the modification of the β -sandwich loop can entail deletion of any combination of 1, 2, 3, 4,
- 10 5, or 6 residues from residues 542-547 (TRPPLG; SEQ ID NO:23). In some embodiments, the β -sandwich loop modification includes deletion of residues PP, RPP, PPL, RPPL (SEQ ID NO:22), TRPPL (SEQ ID NO:24), RPPLG (SEQ ID NO:25) or TRPPLG (SEQ ID NO:23). In some preferred embodiments, the modification involves deletion of residues RPPL (SEQ ID NO:22) from the β -sandwich loop.
- 15 **[0062]** Preferably, the redesigned HCV E2 core polypeptides contain both modifications of the VR2 disordered region and the β -sandwich loop as described above. Some specific examples of redesigned HCV E2 core polypeptides containing such modifications are set forth in SEQ ID NOs:26-38. As exemplified herein, while being derived from different HCV isolates and genotypes, these redesigned HCV E2
- 20 core polypeptides all demonstrated substantially improved yield and purity, as well as satisfactory immunogenic activities. In addition to these specific polypeptides, redesigned HCV E2 polypeptides of the invention also encompass polypeptides having an amino acid sequence that is substantially identical to one of these sequences, including conservatively modified variant sequences. In various embodiments, the
- 25 redesigned HCV E2 core polypeptides of the invention can have an amino acid sequence that is identical to any of SEQ ID NOs:26-38, except for one or more amino acid residue substitutions of non-conserved residues among different HCV isolates or genotypes, or substitutions in the non-conserved region or motif of the E2 sequence of different HCV isolates or genotypes.
- 30 **[0063]** As exemplified herein with isolate H77 (genotype 1a) and isolate HK6a (genotype 6a), E2 sequences from various HCV strains can all be readily employed to generate HCV immunogen polypeptides in accordance with the redesign strategy

described herein. HCV isolates have been grouped into seven genotypes and a number of subtypes, which have different geographical distributions. E2 sequences of many of these isolates are already known. For example, HCV related sequence information can be obtained from the Hepatitis C Virus Databases maintained by the National Institute of Allergy and Infectious Diseases (NIAID). While there is a considerable degree of variability among E2 sequences of different HCV isolates, a certain number of conserved motifs or neutralizing epitopes have been identified in the E2 sequences. In redesigning E2 core polypeptides with a given HCV E2 sequence in accordance with the strategy described herein, one can readily determine the appropriate residues for modifications via sequence alignment and also considering the conserved motifs known in the art. See, e.g., Simmonds et al., *Hepatology* 42:962-973, 2005; Wang et al., *Viruses* 3:2127-2145, 2011; and Krey et al., *PLoS Pathog.* 6:e1000762, 2010; Bhattarai et al., *PLoS Pathog.* 11: e1005183, 2015; and Kachko et al., *Vaccine* 30:69-77, 2011. As exemplified herein with the HK6a isolate, redesigned HCV E2 polypeptides from other isolates can be obtained by directly adopting the H77 sequence designs without further modification. This includes replacing the VR2 disordered region with the same optimized tVR2 sequence identified for the H77 isolate and deleting the corresponding residues in the tip of the β -sandwich loop as in the redesigned H77 sequences exemplified herein.

[0064] As detailed below, the resigned HCV E2 polypeptides may be conjugated to a presenting platform (e.g., nanoparticles or VLPs) via various means. Preferably, the conjugation is achieved via covalent linkage, e.g., protein fusions or insertions. In some preferred embodiments, the protein sequence is fused with the presenting platform sequence via a linker sequence. In various embodiments, other modifications can also be made to the redesigned E2 polypeptides or the conjugating partner in order to improve stability or antigenicity.

[0065] The various redesigned HCV E2 polypeptides of the invention can be obtained or generated in accordance with the protocols exemplified herein or methods well known in the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou ed., 2003). Upon recombinant expression (e.g., in HEK293 F cells as detailed herein), the proteins can be

purified by any of the routinely practiced procedures. See, e.g., *Guide to Protein Purification*, Ed. Deutscher, Meth. Enzymol. 185, Academic Press, San Diego, 1990; and Scopes, *Protein Purification: Principles and Practice*, Springer Verlag, New York, 1982. Substantial purification denotes purification from other proteins or cellular components. A substantially purified protein is at least 60%, 70%, 80%, 90%, 95% or 98% pure. Once purified, antigenicity and other properties of the redesigned HCV E2 polypeptides can also be readily examined with standard methods, e.g., antigenic profiling using known bNAbs and non-Nabs, differential scanning calorimetry (DSC), electron microscopy, binding analysis via ELISA, Biolayer Interferometry (BLI), Surface Plasmon Resonance (SPR), and co-crystallography analysis as exemplified herein.

IV. Scaffolded HCV E2 vaccine compositions

[0066] Other than the engineered or redesigned HCV E2 polypeptides described above, the invention also provides HCV E2 polypeptide based vaccine compositions that contain a heterologous scaffold that presents or incorporates a redesigned HCV E2 protein. Any heterologous scaffold can be used to present the redesigned HCV E2 polypeptide in the construction of the vaccines of the invention. These include nanoparticles, virus-like particles, protein carriers (e.g., immunoglobulin chains or domains such as Fc, KLH, BSA, tetanus toxoid, and diphtheria toxoid), as well as various chemical scaffolds. In some embodiments, a virus-like particle (VLP) such as bacteriophage Q β VLP and nanoparticles can be used. In some preferred embodiments, the heterologous scaffold for presenting or displaying the redesigned HCV E2 polypeptide is a self-assembling nanoparticle. Various nanoparticle platforms can be employed in generating the vaccine compositions of the invention. In general, the nanoparticles employed in the invention need to be formed by multiple copies of a single subunit. The nanoparticles are typically ball-like shaped, and/or have rotational symmetry (e.g., with 3-fold and 5-fold axes), e.g., with an icosahedral structure exemplified herein. Additionally or alternatively, the amino-terminus of the particle subunit has to be exposed and in close proximity to the 3-fold axis, and the spacing of three amino-termini has to closely match the spacing of the carboxyl-termini of the HCV E2 polypeptides. In some preferred embodiments, the immunogens comprise

self-assembling nanoparticles with a diameter of about 20nm or less (usually assembled from 12, 24, or 60 subunits) and 3-fold axes on the particle surface.

[0067] In some preferred embodiments, the immunogen protein or polypeptide (e.g., an HCV E2 protein) is presented on self-assembling nanoparticles such as self-assembling nanoparticles derived from E2p, I3-01 and ferritin (FR) as exemplified herein. E2p is a redesigned variant of dihydrolipoyl acyltransferase from *Bacillus stearothermophilus* that has been shown to self-assemble into thermostable 60-meric nanoparticle. See, e.g., He et al., *Nat. Commun.* 7:12041, 2016. Similarly, I3-01 is an engineered protein that can self-assemble into hyperstable nanoparticles. See, e.g., Hsia et al., *Nature* 535, 136-139, 2016. Sequences of the subunits of these proteins are known in the art. See, e.g., WO2017/192434. Ferritin is a globular protein found in all animals, bacteria, and plants. The globular form of ferritin is made up of monomeric subunit proteins (also referred to as monomeric ferritin subunits), which are polypeptides having a molecule weight of approximately 17-20 kDa. Amino acid sequences of E2p, I3-01 and ferritin nanoparticle subunits as exemplified herein are shown in SEQ ID NOs:39-41, respectively. Relative to the original sequence, E2p sequence shown in SEQ ID NO:39 contains an Ala substitution at residue 92 as underscored in the sequence below. In various embodiments, the HCV E2 nanoparticle vaccines of the invention can employ any of these known nanoparticles, as well as their conservatively modified variants or variants with substantially identical (e.g., at least 90%, 95% or 99% identical) sequences.

[0068] E2p subunit sequence (SEQ ID NO:39)

AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLV
AHRKKFKAIAAEKGKLTFLPYVVKALVSALREYPVLNTAIDDETEEIIQKHYY
25 NIGIAADTDRGLLVPVIKHADRKPIFALAEQINELAEKARDGKLTPGEMKGASC
TITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHR
MIDGATAQKALNHIKRLSDPELLLM

[0069] I3-01 subunit sequence (SEQ ID NO:40)

30 MKMEELFKKHKIVAVLRANSVEEAKKKALAVFLGGVHLIEITFTVPDADTVIK
ELSFLKEMGAIIAGAGTVTSVEQCRKAVESGAEFIVSPHLDEEISQFCKEKGVFYM
PGVMTPTLVKAMKLGHTILKLPGEVVGPPQFVKAMKGPFPNVKFPVPTGGVN
LDNVCEWFKAGVLAVGVGSALVKGTPVEVAEKAKAFVEKIRGCTE

35 [0070] Ferritin sequence (SEQ ID NO:41)

MLSKDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYE
 HAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAI
 KSKDHATFNFLQWYVAEQHEEEVLFDKILDKIELIGNENHGLYLADQYVKGIA
 KSRKS

5

[0071] In addition to these exemplified nanoparticle sequences, many other nanoparticles or VLPs known in the art may also be used in the practice of the invention. These include, e.g., *Aquifex aeolicus* lumazine synthase, *Thermotoga Maritima* encapsulin, *Myxococcus xanthus* encapsulin, bacteriophage Qbeta virus particle, Flock House Virus (FHV) particle, ORSAY virus particle, and infectious bursal disease virus (IBDV) particle. Other molecules that may be used as the presenting platform of the nanoparticle vaccines of the invention include, e.g., molecules with the following PDB IDs: 1JIG (12-mer Dlp-2 from *Bacillus anthracis*), 1UVH (12-mer DPS from *Mycobacterium smegmatis*), 2YGD (24-mer eye lens chaperone α B-crystallin), 3CS0 (24-mer DegP24), 3MH6 and 3MH7 (24-mer HtrA proteases), 3PV2 (12-mer HtrA homolog DegQ WT), 4A8C (12-mer DegQ from *E. Coli.*), 4A9G (24-mer DegQ from *E. Coli.*), 4EVE (12-mer HP-NAP from *Helicobacter pylori* strain YS29), and 4GQU (24-mer HisB from *Mycobacterium tuberculosis*).

[0072] Some HCV E2 nanoparticle vaccine compositions can additionally contain other structural components that function to further enhance stability and antigenicity of the displayed immunogen. In some embodiments, a locking protein domain can be inserted into the nanoparticle construct, e.g., by covalently fused to the C-terminus of the nanoparticle subunit. The locking domain can be any dimeric protein that is capable of forming an interface through specific interactions such as hydrophobic (van der Waals) contacts, hydrogen bonds, and/or salt bridges. General guidance on selecting locking domains and specific examples are described in the art, e.g., PCT2019/036917. In some embodiments, HCV E2 nanoparticle vaccines of the invention can also contain a T-cell epitope to promote robust T-cell responses and to steer B cell development towards bNAbs. The T-cell epitope can be located at any position in relation to the other structural components as long as it does not impact presentation of the immunogen polypeptides on the nanoparticle surface. Any T-cell epitope sequences or peptides known in the art may be employed in the practice of the present invention. They include any polypeptide sequence that contain MHC class-II epitopes and can effectively activate CD4⁺ and CD8⁺ T cells upon immunization, e.g.,

T-helper epitope that activates CD4⁺ T helper cells. See, e.g., Alexander et al., *Immunity* 1, 751-761, 1994; Ahlers et al., *J. Clin. Invest.* 108:1677-1685, 2001; Fraser et al., *Vaccine* 32, 2896-2903, 2014; De Groot et al., *Immunol. Cell Biol.* 8:255-269, 2002; and *Gene Ther.* 21: 225-232, 2014. In some embodiments, the HCV E2
5 nanoparticle vaccines of the invention also contain a neck region or domain to facilitate display of the immunogen on the surface of nanoparticles. The neck region can be a three-helix bundle derived from a viral protein, e.g., Hendra virus domain (PDB ID: 4HEO) or Measles virus domain (PDB ID: 1OKS). It is typically inserted between the immunogen and the nanoparticle subunit, thereby elevating the immunogen polypeptide
10 from the nanoparticle surface. In still some embodiments, the nanoparticle vaccines of the invention can contain a protein domain that serves to stabilize the immunogen polypeptide. For example, the protein domain can be the C-terminal trimerization motif of T4 fibritin (foldon) that is well known in the art. This foldon domain constitutes the C-terminal 30 amino acid residues of the trimeric protein fibritin from bacteriophage
15 T4, and functions in promoting folding and trimerization of fibritin. See, e.g., Papanikolopoulou et al., *J. Biol. Chem.* 279: 8991-8998, 2004; and Guthe et al., *J. Mol. Biol.* 337: 905-915, 2004.

[0073] The scaffolded HCV E2 vaccine compositions of the invention can be constructed in accordance with standard recombinant techniques and the methods
20 described herein (e.g., Examples 4 and 8) and/or other methods that have been described in the art, e.g., He et al., *Nat. Comm.* 7, 12041, 2016; Kong et al., *Nat. Comm.* 7, 12040, 2016; and He et al., *Sci Adv.* 4(11):eaau6769, 2018. In various embodiments, nanoparticle displaying any of the redesigned HCV E2 polypeptides can be constructed by fusing the E2 polypeptide to the subunit of the nanoparticle (e.g., E2p
25 subunit). Preferably, C-terminus of the HCV E2 polypeptide is fused to the N-terminus of the nanoparticle subunit. In some embodiments, a short peptide linker can be used to connect the E2 polypeptide and the nanoparticle, e.g., the 10GS linker (GGGGSGGGGS) (SEQID NO:42) exemplified herein. Once constructed, the antigenicity and structural integrity of the nanoparticle displayed HCV E2 polypeptides
30 can be readily analyzed via standard assays, e.g., antibody binding assays, biolayer interferometry, and negative-stain electron microscopy (EM). As exemplified herein, the various fusion molecules can all self-assemble into nanoparticles that display

immunogenic epitopes of the HCV E2 polypeptides. By eliciting a robust neutralizing antibody response, these nanoparticles are useful for vaccinating individuals against a broad range of HCV viruses.

5 V. Polynucleotides and expression constructs

[0074] The redesigned HCV E2 polypeptides and the nanoparticle vaccine compositions of the invention are typically produced by first generating expression constructs (i.e., expression vectors) that contain operably linked coding sequences of the various structural components described herein. Accordingly, in some related aspects, the invention provides substantially purified polynucleotides (DNA or RNA) that encode the redesigned HCV E2 polypeptides and nanoparticle displayed HCV E2 immunogens as described herein, as well as expression vectors that harbor such polynucleotides and host cells for producing the HCV E2 immunogen polypeptides and the vaccine compositions (e.g., HEK293 F cells and ExpiCHO cells exemplified herein). The fusion polypeptides encoded by the polynucleotides or expressed from the vectors are also included in the invention.

[0075] The polynucleotides and related vectors can be readily generated with standard molecular biology techniques or the protocols exemplified herein. For example, general protocols for cloning, transfecting, transient gene expression and obtaining stable transfected cell lines are described in the art, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou edition, 2003). Introducing mutations to a polynucleotide sequence by PCR can be performed as described in, e.g., *PCR Technology: Principles and Applications for DNA Amplification*, H.A. Erlich (Ed.), Freeman Press, NY, NY, 1992; *PCR Protocols: A Guide to Methods and Applications*, Innis et al. (Ed.), Academic Press, San Diego, CA, 1990; Mattila et al., *Nucleic Acids Res.* 19:967, 1991; and Eckert et al., *PCR Methods and Applications* 1:17, 1991.

[0076] The selection of a particular vector depends upon the intended use of the fusion polypeptides. For example, the selected vector must be capable of driving expression of the fusion polypeptide in the desired cell type, whether that cell type be prokaryotic or eukaryotic. Many vectors contain sequences allowing both prokaryotic

vector replication and eukaryotic expression of operably linked gene sequences.

Vectors useful for the invention may be autonomously replicating, that is, the vector exists extrachromosomally and its replication is not necessarily directly linked to the replication of the host cell's genome. Alternatively, the replication of the vector may be linked to the replication of the host's chromosomal DNA, for example, the vector may be integrated into the chromosome of the host cell as achieved by retroviral vectors and in stably transfected cell lines. Both viral-based and nonviral expression vectors can be used to produce the immunogens in a mammalian host cell. Nonviral vectors and systems include plasmids, episomal vectors, typically with an expression cassette for expressing a protein or RNA, and human artificial chromosomes (see, e.g., Harrington et al., Nat. Genet. 15:345, 1997). Useful viral vectors include vectors based on lentiviruses or other retroviruses, adenoviruses, adenoassociated viruses, cytomegalovirus, herpes viruses, vectors based on SV40, papilloma virus, HBP Epstein Barr virus, vaccinia virus vectors and Semliki Forest virus (SFV). See, Brent et al., supra; Smith, Annu. Rev. Microbiol. 49:807, 1995; and Rosenfeld et al., Cell 68:143, 1992.

[0077] Depending on the specific vector used for expressing the fusion polypeptide, various known cells or cell lines can be employed in the practice of the invention. The host cell can be any cell into which recombinant vectors carrying a fusion of the invention may be introduced and wherein the vectors are permitted to drive the expression of the fusion polypeptide is useful for the invention. It may be prokaryotic, such as any of a number of bacterial strains, or may be eukaryotic, such as yeast or other fungal cells, insect or amphibian cells, or mammalian cells including, for example, rodent, simian or human cells. Cells expressing the fusion polypeptides of the invention may be primary cultured cells or may be an established cell line. Thus, in addition to the cell lines exemplified herein (e.g., CHO cells), a number of other host cell lines capable well known in the art may also be used in the practice of the invention. These include, e.g., various Cos cell lines, HeLa cells, HEK293, AtT20, BV2, and N18 cells, myeloma cell lines, transformed B-cells and hybridomas.

[0078] The use of mammalian tissue cell culture to express polypeptides is discussed generally in, e.g., Winnacker, *From Genes to Clones*, VCH Publishers, N.Y., N.Y., 1987. The fusion polypeptide-expressing vectors may be introduced to the

selected host cells by any of a number of suitable methods known to those skilled in the art. For the introduction of fusion polypeptide-encoding vectors to mammalian cells, the method used will depend upon the form of the vector. For plasmid vectors, DNA encoding the fusion polypeptide sequences may be introduced by any of a number of

5 transfection methods, including, for example, lipid-mediated transfection (“lipofection”), DEAE-dextran-mediated transfection, electroporation or calcium phosphate precipitation. These methods are detailed, for example, in Brent et al., supra. Lipofection reagents and methods suitable for transient transfection of a wide variety of transformed and non-transformed or primary cells are widely available, making

10 lipofection an attractive method of introducing constructs to eukaryotic, and particularly mammalian cells in culture. For example, LipofectAMINE™ (Life Technologies) or LipoTaxi™ (Stratagene) kits are available. Other companies offering reagents and methods for lipofection include Bio-Rad Laboratories, Clontech, Glen Research, Life Technologies, JBL Scientific, MBI Fermentas, PanVera, Promega,

15 Quantum Biotechnologies, Sigma-Aldrich, and Wako Chemicals USA.

[0079] For long-term, high-yield production of recombinant fusion polypeptides, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the fusion polypeptide-encoding sequences controlled by appropriate expression control elements (e.g.,

20 promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and selectable markers. The selectable marker in the recombinant vector confers resistance to the selection and allows cells to stably integrate the vector into their chromosomes. Commonly used selectable markers include neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., J. Mol. Biol., 150:1,

25 1981); and hygromycin, which confers resistance to hygromycin (Santerre et al., Gene, 30: 147, 1984). Through appropriate selections, the transfected cells can contain integrated copies of the fusion polypeptide encoding sequence.

VI. Pharmaceutical compositions and therapeutic applications

30 **[0080]** The invention provides pharmaceutical or immunogenic compositions and related methods of using the redesigned HCV E2 polypeptides and nanoparticles displaying the polypeptides as described herein for preventing and treating HCV

infections. In some embodiments, a redesigned HCV E2 polypeptide or a nanoparticle displaying the polypeptide is included in a pharmaceutical composition. The pharmaceutical composition can be either a therapeutic formulation or a prophylactic formulation. Typically, the composition additionally includes one or more
5 pharmaceutically acceptable vehicles and, optionally, other therapeutic ingredients (for example, antibiotics or antiviral drugs). Various pharmaceutically acceptable additives can also be used in the compositions.

[0081] Some of the pharmaceutical compositions of the invention are vaccines. For vaccine compositions, appropriate adjuvants can be additionally included.
10 Examples of suitable adjuvants include, e.g., aluminum hydroxide, lecithin, Freund's adjuvant, MPL™ and IL-12. In some embodiments, the redesigned HCV E2 polypeptides disclosed herein can be formulated as a controlled-release or time-release formulation. This can be achieved in a composition that contains a slow release polymer or via a microencapsulated delivery system or bioadhesive gel. The various
15 pharmaceutical compositions can be prepared in accordance with standard procedures well known in the art. See, e.g., Remington's Pharmaceutical Sciences, 19^{sup}.th Ed., Mack Publishing Company, Easton, Pa., 1995; Sustained and Controlled Release Drug Delivery Systems, J. R. Robinson, ed., Marcel Dekker, Inc., New York, 1978); U.S. Pat. Nos. 4,652,441 and 4,917,893; U.S. Pat. Nos. 4,677,191 and 4,728,721; and U.S.
20 Pat. No. 4,675,189.

[0082] Therapeutic methods of the invention involve administering a redesigned HCV E2 polypeptide of the invention or a pharmaceutical composition containing the polypeptide to a subject having or at risk of developing HCV infection. In some
embodiments, a pharmaceutical compositions of the invention is employed in
25 therapeutic or prophylactic applications for treating HCV infection or eliciting an protective immune response against HCV in a subject. For example, the composition can be administered to a subject to induce an immune response to HCV, e.g., to induce production of broadly neutralizing antibodies to HCV. For subjects at risk of developing an HCV infection, a vaccine composition of the invention can be
30 administered to provide prophylactic protection against viral infection. Depending on the specific subject and conditions, the pharmaceutical compositions of the invention can be administered to subjects by a variety of administration modes known to the

- person of ordinary skill in the art, for example, intramuscular, subcutaneous, intravenous, intra-arterial, intra-articular, intraperitoneal, or parenteral routes. In general, the pharmaceutical composition is administered to a subject in need of such treatment for a time and under conditions sufficient to prevent, inhibit, and/or
- 5 ameliorate a selected disease or condition or one or more symptom(s) thereof. Symptoms of HCV exposure or infection include, e.g., inflammation of the liver, decreased appetite, fatigue, abdominal pain, jaundice, flu-like symptoms, itching, muscle pain, joint pain, intermittent low-grade fevers, sleep disturbances, nausea, dyspepsia, cognitive changes, depression headaches and mood changes.
- 10 **[0083]** Typically, the immunogenic composition of the invention is administered in an amount sufficient to induce an immune response against HCV. For therapeutic applications, the compositions should contain a therapeutically effective amount of a redesigned HCV E2 polypeptide or nanoparticle vaccine composition described herein. For prophylactic applications, the compositions should contain a prophylactically
- 15 effective amount of the HCV E2 polypeptide or a nanoparticle displaying the polypeptide. The appropriate amount of the polypeptide immunogen or the nanoparticle composition can be determined based on the specific disease or condition to be treated or prevented, severity, age of the subject, and other personal attributes of the specific subject (e.g., the general state of the subject's health and the robustness of
- 20 the subject's immune system). Determination of effective dosages is additionally guided with animal model studies followed up by human clinical trials and is guided by administration protocols that significantly reduce the occurrence or severity of targeted disease symptoms or conditions in the subject.
- [0084]** For prophylactic applications, the immunogenic composition is provided in
- 25 advance of any symptom, for example in advance of infection. The prophylactic administration of the immunogenic compositions serves to prevent or ameliorate any subsequent infection. Thus, in some embodiments, a subject to be treated is one who has, or is at risk for developing, an HCV infection, for example because of exposure or the possibility of exposure to HCV. Following administration of a therapeutically
- 30 effective amount of the disclosed therapeutic compositions, the subject can be monitored for HCV infection, symptoms associated with HCV infection, or both.

[0085] For therapeutic applications, the immunogenic composition is provided at or after the onset of a symptom of disease or infection, for example after development of a symptom of HCV infection, or after diagnosis of HCV infection. The immunogenic composition can thus be provided prior to the anticipated exposure to HCV virus so as to attenuate the anticipated severity, duration or extent of an infection and/or associated disease symptoms, after exposure or suspected exposure to the virus, or after the actual initiation of an infection.

[0086] The pharmaceutical composition of the invention can be combined with other agents known in the art for treating or preventing HCV infections. These include, e.g., (1) Daclatasvir, (2) Elbasvir and grazoprevir, (3) Glecaprevir and pibrentasvir, (4) Ledipasvir and sofosbuvir, (5) Ombitasvir, paritaprevir, and ritonavir, (6) Simeprevir (Olysio) and sofosbuvir (Sovaldi), (7) Sofosbuvir and velpatasvir (Epclusa), and (8) Sofosbuvir, velpatasvir, and voxilaprevir (Vosevi). See, e.g., Kish et al., P T. 42:316-329, 2017; and Dahiya et al., BCMJ 61:72-77, 2019. Administration of the pharmaceutical compositions and the known anti-HCV agents can be either concurrently or sequentially.

[0087] Pharmaceutical compositions containing a redesigned HCV E2 polypeptide or nanoparticle vaccine of the invention can be provided as components of a kit. Optionally, such a kit includes additional components including packaging, instructions and various other reagents, such as buffers, substrates, antibodies or ligands, such as control antibodies or ligands, and detection reagents. An optional instruction sheet can be additionally provided in the kits.

EXAMPLES

[0088] The following examples are offered to illustrate, but not to limit the present invention.

Example 1 Structure-based optimization of HCV envelope glycoprotein E2 core

[0089] The HCV E2 ectodomain (E2 Δ TM or E2_{ECTO}) is stabilized by nine conserved disulfide bonds, contains three variable regions including hypervariable region 1 (HVR1, a.a. 384-410), VR2 (a.a. 460-485), and VR3 (a.a. 570-580) and is covered with ~11 *N*-linked glycans (see Goffard et al., Biochimie 85, 295-301, 2003)

(Fig. 1A). HVR1 modulates SR-BI interaction and facilitate host immune evasion by generating escape mutations and shielding neutralizing epitopes. Empirical engineering (Fig. 6A) enabled E2 structure determination by shortening N-/C-termini and VR2, removing glycans at N448 and N576 (see Kong et al., Science 342, 1090-1094, 2013), and additional V3 truncation (see, Tzarum et al., Sci Adv 5, eaav1882, 2019). The E2 core contains a central immunoglobulin (Ig)-like β -sandwich with front and back layers (Fig. 6B). The CD81 receptor binding site is a hydrophobic patch formed by the front layer and CD81 binding loop and overlaps the E2 neutralizing face (Fig. 6B, middle). However, current E2c constructs still exhibit high flexibility involving the front layer C-terminus, shortened VR2, and β -sandwich N-terminus.

[0090] Here, we redesigned the VR2 disordered region that is anchored to the back layer and β -sandwich by two disulfide bonds, C452-C620 and C494-C564 (Fig. 1, A and B). Although this region consists of 43 residues in wild-type E2 and 21 residues in E2c/E2c3, the C α distance between C452 and C494 is 26.3 Å for H77 E2c (Fig. 6B), which could be spanned by only 7 residues. We first manually truncated the VR2 loop to 13 residues (tVR2) (Fig. 1, A and B; Fig. 6A) and deleted the tip (aa 543-546) of the β -sandwich loop (a.a. 540-550) to focus the immune response to bNAbs epitopes, as non-Nabs, such as AR1B, E1 and HEPC46, bind to this region (Fig. 6, C-E). The new E2 core is termed E2 mini-core 3 (E2mc3). Next, we redesigned tVR2 in H77 E2mc3 for two loop lengths, 13 aa (as in E2mc3) and 12 aa, using ensemble-based *de novo* protein design (see Kong et al., Nat. Commun. 7, 12040, 2016) to identify optimal tVR2 sequences that stabilize E2mc3 (Fig. 1B). For each tVR2 length (13 and 12aa), an ensemble (1,000) of loop conformations was generated to connect C452 and C494 (Fig. 6F) with C α root-mean-square fluctuation (RMSF) ranging from 1.9 to 9.8 Å (average 3.6 Å) and 1.6 to 7.8 Å (average 3.2 Å), respectively (Fig. 6G). After extensive Monte Carlo sampling, the five top-ranking sequences for each loop length, E2mc3-v1-v5 and v6-v10 (Fig. 1B, right; Fig. 6H), were selected for further characterization. As HK6a E2c3 and H77 E2c share high structural similarity and disordered regions when bound to AR3 bNAbs, we designed HK6a E2mc3 and E2mc3-v1 constructs (Fig. 6I) without further modification. In total, eleven H77 E2 cores and two HK6a E2 cores were advanced to experimental evaluation.

Example 2 Biochemical, biophysical, and antigenic assessment of HCV E2mc3 designs

[0091] 13 E2mc3 constructs from H77 and HK6a and two parental E2c3 constructs (see, Tzarum et al., Sci Adv 5, eaav1882, 2019; and Kong et al., Proc. Natl. Acad. Sci. U.S.A. 113, 12768–12773, 2016) were transiently expressed in HEK293 F cells and purified using immunoaffinity (see Kong et al., Science 342, 1090-1094, 2013) followed by size exclusion chromatography (SEC). Overall, the purified E2mc3 variants showed greater yield than their respective E2c3 constructs ranging from 5.0 to 11.5 mg from 1L 293F transfection. AR3A-purified E2mc3 was mostly in monomeric form with a small aggregate peak (Fig. 1C and Fig. 7A) and SEC-purified proteins ran as a single band (~50 kDa) on SDS-PAGE (Fig. 1D and Fig. 7B). We then tested H77 and HK6a E2mc3 variants by ELISA on a panel of HCV antibodies that includes (b)NAb targeting antigenic site 412 (AS412), AS434, antigenic region 3 (AR3), AR2, and non-NAb targeting AR1 (Fig. 7C). H77 E2mc3 showed greater binding than E2c3 for most bNAbs (excluding HEPC3/74) and NAb AR2A, with further improvements for some E2mc3 variants (Fig. 1E, upper panel; Fig. 7, D and E). As expected, truncation of the β -sandwich loop reduced binding to non-NAb AR1B and E1 with negligible effect on AR1A, which recognizes AR1 but not the β -sandwich loop. Similar patterns were observed for HK6a E2mc3 and E2mc3-v1 except for 212.1.1 (Fig. 1E, lower panel; Fig. 7, F and G) with no detectable binding by genotype-specific AR2A and AR1A/B. By biolayer interferometry (BLI), HK6a E2mc3 variants exhibited similar antigenic profiles (Fig. 1F; Fig. 7, H-J). Differential scanning calorimetry (DSC) showed a 4.2°C increase in T_m for E2mc3 (H77) relative to E2c3 that was further increased by 1°C and 0.2°C for E2mc3-v1 and v6, respectively (Fig. 1G and Fig. 7K).

25

Example 3 Characterization of minimized cores derived from H77 and HK6a

[0092] Crystallization of H77 and HK6a E2 cores with Fabs from AR3A/B/C/D bNAbs led to structures of H77 E2mc3-v1 and E2mc3-v6 with AR3C and HK6a E2mc3-v1 with AR3B at 1.90 Å, 2.85 Å, and 2.06 Å, respectively (Fig. 2). The overall fold of E2mc3 variants is highly similar to H77 E2c and HK6a E2c3 (PDB: 4MWF and 6BKB) (Fig. 8A), but with differences in the back layer C-terminus (a.a. 629-640) and a front layer loop (a.a. 430-438) (Fig. 8A) that interacts with HCDR3 of bNAbs AR3A-

30

D. However, similar (hydrophilic) contacts are maintained with HCDR3 (Fig. 8A) further supporting the conformational plasticity of the E2 front layer. The shortened β -sandwich loop can be fully modeled in the H77 E2mc3-v1 complex with bNAb AR3C (Fig. 8B), confirming that its truncation results in loss of key interactions with non-
 5 NAb E1 (Fig. 1E and Fig. 2B). The redesigned tVR2 can be fully modeled in bNAb-bound H77 and HK6a E2mc3-v1 structures but is only partially visible in the AR3C-bound H77 E2mc3-v6 structure (Fig. 2C and Fig. 9A). The tVR2 redesign does not introduce any conformational changes to the E2 neutralizing face (Fig. 2D), which is anchored to the back layer and β -sandwich by C452-C620 and C494-C564 and interacts
 10 with the truncated VR3 and β -sandwich (Fig. 9B). While H77 and HK6a E2mc3-v1 constructs share the same tVR2 sequence, a significant difference in conformation (Fig. 2C and Fig. 9C) likely results from differences in sequence and structure of adjacent VR3 and β -sandwich loop in genotypes 1 and 6 (Fig. 8B and Fig. 9C). The tVR2 redesign has minimal effect when compared to recent truncated E2_{ECTO} structures from
 15 isolates 1a53 and 1b09 of genotype 1 with bNAbs HEPC3/74 (Fig. 6B and Fig. 9D).

Example 4 Design and characterization of nanoparticles presenting optimized E2 cores

[0093] Recently, improved immunogenicity in mice was reported for a ferritin nanoparticle carrying soluble E2 (sE2, a.a. 384-661) that contains three full-length immunodominant variable loops (Yan et al., J. Infect. Dis. [Epub ahead of print], 2019). Here, we displayed our E2mc3 variants, which only present the conserved bNAb epitopes, on self-assembling nanoparticles as multivalent HCV vaccine candidates (Fig. 3A). We tested three nanoparticle platforms: 24-meric ferritin (FR, as control) and 60-
 25 meric E2p and I3-01, ranging in size from 24.5-37.5 nm (Fig. 3B). We genetically fused the C-terminus of E2mc3-v1 to the N-terminus of a nanoparticle subunit via a 10-residue linker, (G4S)₂. Six constructs were transiently expressed in ExpiCHO or 293 F cells and purified on the AR3A column followed by SEC (Fig. 3C and Fig. 10A). For H77, the SEC profiles demonstrated substantial yield and purity for all E2 core
 30 nanoparticles with different patterns for 24- vs. 60-mers (Fig. 3C). For HK6a, reduced nanoparticle yield and purity was accompanied by increased low-molecular-weight species (Fig. 10A), suggesting that H77 tVR2 may be incompatible with HK6a and

hinder particle assembly. However, effective particle assembly was observed in both blue-native PAGE (BN-PAGE) and negative-stain electron microscopy (EM) (Fig. 3, D and E; Fig. 10, B and C), as well as enhanced bNAb binding to H77 nanoparticles (up to 100-fold change in EC₅₀) and no binding to non-NAbs targeting the β -sandwich loop (Fig. 3F and Fig. 10 D-G). HK6a E2mc3-v1 nanoparticles exhibited similar but genotype-specific profiles. In BLI, correlation between peak signal and antigen valency was observed irrespective of genotype with 60-mers > 24-mer > E2 core monomer (Fig. 3G; Fig. 10, H and I), consistent with our HCV gp140 nanoparticles (He et al., Sci. Adv. 4, eaau6769, 2018).

10

Example 5 E2 core nanoparticles elicit stronger immune responses than E2 cores
[0094] We assessed H77 and HK6a E2 core nanoparticles in wild-type BALB/c mice in studies #1 and #2, respectively, using a short regimen (Fig. 4A). N. 2, I3-01-based constructs were not included due to the difficulties in producing HK6a E2mc3-v1-10GS-I3-01 nanoparticles, as indicated by SEC (Fig. 10A). In study #1, three H77-based vaccines showed a correlation between E2-specific EC₅₀ titer and antigen valency at week 2 with significant *P*-values (Fig. 4B, upper panel; Fig. 11, A and B). While E2-specific antibody titers continued to rise, the difference between the three vaccine groups diminished at week 11. In study #2, we compared HK6a E2mc3-v1, its E2p nanoparticle, and HK6a/H77 E2mc3-v1 E2p nanoparticle mix (Fig. 4B, lower panel; Fig. 11, C and D). The HK6a E2mc3-v1 E2p group retained its advantage in antibody titer until week 8, whereas its H77 counterpart did until week 11. The E2p mix elicited significantly higher titers to H77 than to HK6a throughout the immunization. Overall, E2 core nanoparticles induced greater antibody titers than E2 cores, although E2 only accounts for 42% (E2p) to 51% (FR) of the protein mass. We then evaluated serum neutralization using HCV pseudoparticles (HCVpp). In study #1, autologous neutralization increased steadily over time with distinct temporal patterns (Fig. 4C, upper panel). At week 2, the FR group showed the highest H77 neutralization, whereas the E2p group was unexpectedly the lowest (Fig. 4B, upper panel). From week 5, the FR group showed lower neutralizing activity with a significant *P*-value at week 11, whereas the E2p group became the best performer with statistical significance at weeks 8 and 11. Week 11 sera also neutralized heterologous isolates HCV1 (1a), J6 (2), and

30

SA13 (5a), with significant *P*-values for HCV1 and J6 (Fig. 4C, lower panel). A similar trend was observed for group #2 (Fig. 4D), and the E2p mix group was equivalent to the HK6a-only E2p group. Five HCV bNAbs and HCV bNAb VRC01 validated the HCVpp assays (Fig. 4E).

5

Example 6 Distinctive patterns of E2 core and nanoparticle induced B cell response
[0095] We combined antigen-specific B cell sorting and antibody NGS to obtain a quantitative readout of vaccine-induced B cell responses and determine B cell patterns associated with different vaccine platforms (Fig. 5A). We used an H77 E2mc3-v1 probe
 10 with an C-terminal Avi-tag (Fig. 12A) to sort E2-specific splenic B cells from mice in the H77 E2mc3-v1 and E2p nanoparticle groups by flow cytometry. A greater frequency/number of E2-specific B cells was observed for the E2p group with significant *P*-values (Fig. 5B and Fig. 12B). Sorted B cells from 10 mice, five per group, were subjected to next-generation sequencing (NGS) analysis and repertoire
 15 profiling (Fig. 12C). The E2p group used significantly more (5 to 8) heavy-chain variable (V_H) genes than the E2 core group (~1) (Fig. 5C). Antibodies elicited by E2p contained more V_H mutations with a significant *P*-value of 0.0268 (Fig. 5D). Distinct patterns of HCDR3 length were observed for the two groups with the E2 core group showing two dominant HCDR3 lengths, while the E2p group produced a much broader
 20 distribution (Fig. 5E). The E2p group yielded a greater average RMSF – the range in which the loop length varies – than the E2 core group (4.5aa vs. 1.1aa) with a *P*-value of <0.0001.

Example 7 NAbs induced by E2 core and E2 nanoparticle target different epitopes
[0096] Mouse serum contains nonspecific antiviral activity, which may interfere with HCVpp assays. We purified mouse IgG from study #1 at week 11 for
 25 neutralization of H77 and SA13 HCVpps with starting IgG at 100 µg/ml followed by a series of three-fold dilutions (Fig. 5F and Fig. 13A). For H77, while no mice sera in the E2 core group neutralized >60% viruses at the first concentration, mice #9 and #10 in
 30 the E2p group showed plateaued curves, suggesting potent NAb in the IgG, with a similar but less pronounced trend for SA13. Unpaired *t* tests indicated a significant

difference between E2p and E2 core groups for H77 ($P < 0.0001$), but not for SA13 ($P = 0.1680$). The FR group ranked the lowest in serum neutralization but slightly outperformed E2 core in IgG neutralization (Fig. 13A). Nonetheless, ferritin (FR) may not be an optimal platform for HCV nanoparticle vaccine design. Here, two epitope-specific probes were used to examine vaccine-induced antibody responses to two prominent bNAbs epitopes: front layer (FL, a.a. 421-459), integral to the E2 neutralizing face (Tazrum et al., Front. Immunol. 9, 1315, 2018), and AS412 (Fig. 5G, middle). A trimeric scaffold was designed to present FL, which was anchored to each subunit via an engineered disulfide bond (Fig. 13B). This trimer FL-scaffold was displayed on FR.

10 In ELISA, the E2p group yielded an average EC_{50} titer of 4281 compared to 3044 for the E2mc3-v1 group (Fig. 5G, left and Fig. 13C). However, unpaired t test reported a non-significant P -value, 0.1192, between the two groups (Fig. 5G, left). Nonetheless, nanoparticle display improved recognition of FL including antigenic site 434 (AS434, a.a. 434-446). We then used a previously designed FR nanoparticle (He et al., Sci. Rep.

15 5:12501, 2015) (Fig. 13B, bottom) to probe AS412-specific response. In ELISA, the E2p group demonstrated a uniform, robust response to the β -hairpin of AS412 with an average EC_{50} titer of 5584, which is 38-fold greater than the E2 core group with a P -value of < 0.0001 (Fig. 5G, right and Fig. 13C, bottom). Thus, particulate display focuses the response on conserved bNAbs epitopes. These epitope probes also provide

20 valuable tools for epitope mapping for assessment of HCV vaccine candidates.

Example 8 Some exemplified materials and methods

[0097] Structural design of truncated VR2 (tVR2): Based on the structure of bNAbs AR3C-bound H77 E2c (PDB ID: 4MWF), the already shortened VR2 loop in E2c, i.e.

25 the segment between C452 and C494, was manually truncated by removing exposed hydrophobic residues and a disulfide bond (C459-C486), resulting in the H77 E2mc3 construct (Fig. 6A). The truncated VR2 (tVR2) was modeled by the LOOPY (Xiang et al., Proc. Natl. Acad. Sci. U.S.A. 99:7432-7437, 2002), a torsion-space loop modeling and prediction program. Computational redesign of tVR2 was then performed using an

30 ensemble-based *de novo* protein design method (Kong et al., Nat. Commun. 7:12040, 2016) with a focus on the N-terminal region of the peptide sequence PERASGHYPRP (SEQ ID NO:62) between C452 and C494. Briefly, an ensemble of 11-aa (G₆HYPRP)

(SEQ ID NO:63) or 10-aa (G₅HYPRP) (SEQ ID NO:64) loop conformations was generated to connect C452 and C494 using LOOPY. For each loop conformation, a starting sequence for the multi-glycine (G_n) region was selected from a pool of 50 random sequences based on the RAPDF potential (Zhu et al., *Proteins* 65, 463-479, 2006) and subjected to 500 steps of Monte Carlo simulated annealing (MCSA) with the temperature linearly decreasing from 300 to 10K. The lowest-energy sequence for each loop was recorded and all MCSA-derived designs were ranked based on energy at the completion of the process. The top 5 designs from the G₆ and G₅ ensembles, termed H77 E2mc3-v1-v5 and v6-10, respectively, were selected for experimental validation.

HK6a E2mc3 and E2mc3-v1 constructs were designed by directly adopting the H77 sequence designs without further modification.

[0098] Expression and purification of E2 antigens: All E2 cores (E2c3, E2mc3, and E2mc3 v1-v10) and E2p-based nanoparticles were transiently expressed in HEK293 F cells (Thermo Fisher) for biochemical, biophysical, and antigenic analyses.

Briefly, 293 F cells were thawed and incubated with FreeStyle™ 293 Expression Medium (Life Technologies, CA) in a shaker incubator at 37°C, 135 rpm and 8% CO₂. When the cells reached a density of 2.0×10⁶/ml, expression medium was added to reduce cell density to 1.0×10⁶/ml for transfection with polyethyleneimine (PEI) (Polysciences, Inc). Next, 900 µg of plasmid in 25 ml of Opti-MEM transfection medium (Life Technologies, CA) was mixed with 5 ml of PEI-MAX (1.0 mg/ml) in 25 ml of Opti-MEM. After 30-min incubation, the DNA-PEI-MAX complex was added to 1L 293 F cells. Culture supernatants were harvested five days after transfection, clarified by centrifugation at 1800 rpm for 20 min, and filtered using a 0.45 µm filter (Thermo Scientific). E2 proteins were extracted from the supernatants using an AR3A antibody column as previously described (Long et al., *Science* 342, 1090-1094, 2013; and Tzarum *et al.*, *Sci Adv* 5, eaav1882, 2019). Bound proteins were eluted three times, each with 5ml of 0.2M Glycine (pH=2.2) and neutralized with 0.5ml of Tris-Base (pH=9.0). The proteins were further purified by size exclusion chromatography (SEC) on a Superdex 75 Increase 10/300 GL column (GE Healthcare) for E2 cores and on a Superose 6 10/300 GL column (GE Healthcare) for E2p nanoparticles. E2mc3-v1-attached ferritin and I3-01 nanoparticles were produced in ExpiCHO cells (Thermo Fisher). Briefly, ExpiCHO cells were thawed and incubated with ExpiCHO™

Expression Medium (Thermo Fisher) in a shaker incubator at 37 °C, 135 rpm and 8% CO₂. When the cells reached a density of 10×10^6 ml⁻¹, ExpiCHO™ Expression Medium was added to reduce cell density to 6×10^6 ml⁻¹ for transfection. The ExpiFectamine™ CHO/plasmid DNA complexes were prepared for 100-ml transfection in ExpiCHO cells following the manufacturer's instructions. For these two nanoparticles, 100 µg of plasmid and 320 µl of ExpiFectamine™ CHO reagent were mixed in 7.7 ml of cold OptiPRO™ medium (Thermo Fisher). After the first feed on day one, ExpiCHO cells were cultured in a shaker incubator at 32 °C, 120 rpm and 8% CO₂ following the Max Titer protocol with an additional feed on day five (Thermo Fisher). Culture supernatants were harvested 13 to 14 days after transfection, clarified by centrifugation at 4000 rpm for 20 min, and filtered using a 0.45 µm filter (Thermo Fisher). The AR3C antibody column was used to extract E2mc3-attached nanoparticles from the supernatants, which was followed by SEC on a Superose 6 10/300 GL column. For E2 cores and nanoparticles, protein concentration was determined using UV₂₈₀ absorbance with theoretical extinction coefficients.

[0099] *Blue native* polyacrylamide gel electrophoresis (BN-PAGE): HCV E2 core nanoparticles were analyzed by *blue native* polyacrylamide gel electrophoresis (BN-PAGE) and stained with Coomassie blue. The protein samples were mixed with G250 loading dye and added to a 4-12% Bis-Tris NuPAGE gel (Life Technologies). BN-PAGE gels were run for 2.5 hours at 150 V using the NativePAGE™ running buffer (Life Technologies) according to the manufacturer's instructions.

[00100] Enzyme-linked immunosorbent assay (ELISA): Each well of a Costar™ 96-well assay plate (Corning) was first coated with 50 µl PBS containing 0.2 µg of the appropriate antigens. The plates were incubated overnight at 4 °C, and then washed five times with wash buffer containing PBS and 0.05% (v/v) Tween 20. Each well was then coated with 150 µl of a blocking buffer consisting of PBS, 20 mg ml⁻¹ blotting-grade blocker (Bio-Rad), and 5% (v/v) FBS. The plates were incubated with the blocking buffer for 1 hour at room temperature, and then washed five times with wash buffer. In the mouse sample analysis, serum or plasma was diluted by 50-fold in the blocking buffer and subjected to a 10-fold dilution series. For each sample dilution, a total of 50 µl volume was added to the wells. Each plate was incubated for 1 hour at room temperature, and then washed five times with wash buffer. A 1:2000 dilution of

horseradish peroxidase (HRP)-labeled goat anti-mouse IgG antibody (Jackson ImmunoResearch Laboratories) was then made in the wash buffer, with 50 µl of this diluted secondary antibody added to each well. The plates were incubated with the secondary antibody for 1 hour at room temperature, and then washed five times with wash buffer. Finally, the wells were developed with 50 µl of TMB (Life Sciences) for 3-5 min before stopping the reaction with 50 µl of 2 N sulfuric acid. The resulting plate readouts were measured at a wavelength of 450 nm.

[00101] Bio-layer interferometry (BLI): The kinetics of E2 cores and nanoparticle binding to HCV-specific antibodies was measured using an Octet Red96 instrument (fortéBio, Pall Life Sciences). All assays were performed with agitation set to 1000 rpm in fortéBio 1× kinetic buffer. The final volume for all the solutions was 200 µl per well. Assays were performed at 30 °C in solid black 96-well plates (Geiger Bio-One). 5 µg ml⁻¹ of antibody in 1× kinetic buffer was loaded onto the surface of anti-human Fc Capture Biosensors (AHC) for E2 cores and of anti-human Fc Quantitation Biosensors (AHQ) for nanoparticles for 300 s. A 60 s biosensor baseline step was applied prior to the analysis of the association of the antibody on the biosensor to the antigen in solution for 200 s. A two-fold concentration gradient of antigen, starting at 3.57 µM for E2 cores and 52.08 nM for nanoparticles, depending on the size, was used in a titration series of six. The dissociation of the interaction was followed for 300 s. Correction of baseline drift was performed by subtracting the mean value of shifts recorded for a sensor loaded with antibody but not incubated with antigen and for a sensor without antibody but incubated with antigen. Octet data were processed by FortéBio's data acquisition software v.8.1. Experimental data were fitted with the binding equations describing a 2:1 interaction to achieve optimal fitting. Of note, E2mc3-v1 binding was also measured using AHQ to facilitate the comparison of antibody binding signals with nanoparticles.

[00102] Differential scanning calorimetry (DSC): Thermal melting curves of HCV E2 core glycoproteins were obtained with a MicroCal VP-Capillary calorimeter (Malvern). The purified E2 glycoproteins produced from 293S cells were buffer exchanged into 1×PBS and concentrated to 27–50 µM before analysis by the instrument. Melting was probed at a scan rate of 90 °C·h⁻¹ from 25 °C to 110 °C. Data

processing, including buffer correction, normalization, and baseline subtraction, was conducted using the standardized protocol from the Origin 7.0 software.

[00103] Protein expression and purification for crystallization: The E2 constructs were expressed and purified as described above. Fabs AR3A and AR3B were expressed and purified as previously described by Giang et al., Proc. Natl. Acad. Sci. U.S.A. 109:6205-6210, 2012). The mAbs were purified on a protein G affinity column followed by size exclusion chromatography using a Superdex-200 column (Pharmacia) in 50 mM NaCl, 20 mM Tris-HCl (pH=7.2) buffer.

[00104] Crystallization and structural determination of HK6a E2c3 – Fab E1 – AR3A – protein G complex. The HK6a E2c3 – Fab E1 – AR3A complex was formed by overnight incubation of purified E2 and Fabs in a molar ratio of 1:1.2:1.25 (E2:Fab E1:Fab AR3A) at room temperature followed by size exclusion chromatography (Superdex-200) to remove unbound Fabs using 20 mM Tris and 50 mM NaCl (pH=7.2) buffer. Crystallization experiments were performed using our high-throughput CrystalMation™ robotic system (Rigaku) using the vapor diffusion sitting drop method (drop size 0.3 µl) at 20 °C and resulted with crystals that diffracted to ~5 Å. To improve crystal resolution, prior to the crystallization experiment, domain III of protein G (PDB entry 1IGC) was added to the HK6a E2c3 – Fab E1 – AR3A complex in a molar ratio of 1:2 (complex: protein G). These experiments resulted in crystals of HK6a E2c3 – Fab E1 – AR3A – protein G that diffracted to 3.40 Å. Crystals were obtained using a reservoir solution of 0.2M magnesium chloride, 10% (w/v) PEG 3000, 15% ethylene glycol, 0.1M Na-cacodylate, pH=6.5. Prior to data collection, crystals were flash cooled in liquid nitrogen. Diffraction data sets were collected at Stanford Synchrotron Radiation Lightsource (SSRL). Data were integrated and scaled using HKL2000 and structure was solved by molecular replacement method using Phaser with the HK6a E2c3 - AR3A (PDB entry 6BKB) as a search model. Structure refinement was carried out in Phenix and model building with COOT.

[00105] Crystallization and structural determination of E2mc3 - Fab complexes: Crystallization experiments were performed for H77 E2mc3, H77 E2mc3v-1, H77 E2mc3v-6, HK6a E2mc3, and HK6a E2mc3v-1 in complex with AR3A, AR3B, AR3C, and AR3D Fabs. The E2-Fab complexes were formed by overnight incubation of purified E2 and Fabs in a molar ratio of 1:1.25 (E2:Fab) at room temperature followed

by size exclusion chromatography (Superdex-200) to remove unbound Fabs using 20 mM Tris and 50 mM NaCl (pH=7.2) buffer. Crystallization screening using our high-throughput CrystalMation™ robotic system (Rigaku) using the vapor diffusion sitting drop method (drop size 0.3 µl) at 20 °C resulted in crystals of H77 E2mc3-v1 - AR3C, H77 E2mc3-v6 - AR3C, and HK6a E2mc3-v1 - AR3B that diffracted to 1.90 Å, 2.85 Å, and 2.06 Å, respectively. Crystals of the H77 E2mc3-v1 - AR3C complex were obtained using a reservoir solution of 20% (w/v) PEG 3500, 0.2M di-ammonium hydrogen phosphate; H77 E2mc3-v6 -AR3C complex from 20% (w/v) PEG 3500, 0.2M Na-thiocyanate, pH=6.9; and HK6a E2mc3-v1 - AR3B complex from 20% (w/v) PEG 8000, 0.1M HEPES pH=7.5. Prior to data collection, H77 E2mc3-v6 - AR3C and HK6a E2mc3-v1 - AR3B crystals were cryoprotected with 10-15% ethylene glycol and flash cooled in liquid nitrogen. Diffraction data sets were collected at the Advanced Photon Source (APS). Data were integrated and scaled using HKL2000. Structures were solved by molecular replacement method using Phaser with the H77 E2c - AR3C or HK6a E2c3 - AR3B (PDB entry 4MWF or 6BKC) as a search model. Structure refinement was carried out in Phenix and model building with COOT.

[00106] Negative stain electron microscopy (EM): The EM experiments were conducted at the Scripps Core Microscopy Facility. Briefly, nanoparticle samples were prepared at the concentration of 0.01 mg/ml. Carbon-coated copper grids (400 mesh) were glow-discharged and 8 µL of each sample was adsorbed for 2 minutes. Excess sample was wicked away and grids were negatively stained with 2% uranyl formate for 2 minutes. Excess stain was wicked away and the grids were allowed to dry. Samples were analyzed at 80kV with a Talos L120C transmission electron microscope (Thermo Fisher) and images were acquired with a CETA 16M CMOS camera.

[00107] Mouse immunization and sample collection: The Institutional Animal Care and Use Committee (IACUC) guidelines were followed with animal subjects tested in the immunization study. Eight-week-old BALB/c mice were purchased from The Jackson Laboratory. The mice were housed in ventilated cages in environmentally controlled rooms at TSRI, in compliance with an approved IACUC protocol and AAALAC guidelines. At week 0, each mouse was immunized with 200 µl of antigen/adjuvant mix containing 50 µg of antigen and 100 µl AddaVax adjuvant (Invivogen) per manufacturer's instruction via the intramuscular (*i.m.*) route. At weeks

3, 6 and 9, the animals were boosted with 50 µg of antigen formulated in AddaVax adjuvant. At week 11, the animals were terminally bled through the retro orbital membrane using heparinized capillary tubes. Samples were diluted with an equal volume of PBS and then overlaid on 4.5 ml of Ficoll/Histopaque in a 15 ml SepMate tube (StemCell) and spun at 1200 RPM for 10 min at 20 °C to separate plasma and cells. The plasma was heat inactivated at 56 °C for 30 min, spun at 1200 RPM for 10 min, and sterile filtered. The cells were washed once in PBS and then resuspended in 1 ml of ACK Red Blood Cell lysis buffer (Lonza). After two rounds of washing with PBS, PBMCs were resuspended in 2 ml of Bambanker Freezing Media (Lymphotec Inc.). Spleens were also harvested and grounded against a 40-µm cell strainer (BD Falcon) to release the splenocytes into a cell suspension. The cells were centrifuged, washed in PBS, treated with 10 ml of RBC lysis buffer per manufacturer's specifications, and resuspended in Bambanker Freezing Media for cell freezing. While serum and plasma were used in HCVpp neutralization assays, 80% of the serum from individual mice in study #1 (9, 10, and 10 in groups 1, 2, and 3, respectively) was purified using a 0.2-ml protein G spin kit (Thermo Scientific) following the manufacturer's instructions. Purified IgGs were used to assess the polyclonal serum NAb response in HCVpp assays.

[00108] HCV neutralization assay: HCV pseudotyped particle (HCVpp) assays were utilized to assess the neutralizing activity of vaccine-induced antibody response in mouse sera, as well as synthesized antibodies from the next-generation sequencing (NGS) analysis of bulk-sorted mouse splenic B cells. Briefly, HCVpps were generated by co-transfection of 293T cells with pNL4-3.lucR-E- plasmid and the corresponding expression plasmids encoding the E1E2 genes at a 4:1 ratio by polyethylenimine as previously described by Bazzill et al., *Nano Lett.* **18**:7832-7838, 2018. In vitro neutralization was carried on Huh7.5 cells using a single dilution of 1:50 for mouse sera and three concentrations (10µg/ml, 1.0µg/ml, and 0.1µg/ml) for antibodies. Full neutralization curves were determined for IgGs purified from mice in study #1 against autologous H77 (1a) and heterologous SA13 (5a), with a starting IgG concentration of 100 µg/ml and a series of three-fold dilutions.

[00109] Bulk sorting of HCV E2-specific mouse B cells: Spleens were harvested from immunized mice 15 days after the last immunization and cell suspension was

prepared. Cells were stained as follows: dead cells were excluded by staining with Fixable Aqua Dead Cell Stain kit (Thermo Fisher L34957). Receptors FcγIII (CD16) and FcγII (CD32) were blocked by adding 20 μl of 2.4G2 mAb (BD Pharmingen N553142). Cells were then incubated with 10 μg/ml of biotinylated HCV E2mc3 protein. Briefly, E2mc3 was generated by biotinylation of the individual Avi-tagged HCV E2mc3 using biotin ligase BirA according to the manufacturer's instructions (Avidity LLC). Biotin excess was removed by SEC on a Superdex 200 column (GE Healthcare). In the SEC profile, the Avi-tagged E2mc3 peak is centered at 14.5 ml, while a broader peak of biotin ligase can be found at 18-23 ml. Cells and biotinylated proteins were incubated for 5 min at 4 °C, followed by the addition of 2.5 μl of anti-mouse IgG fluorescently labeled with FITC (Jackson ImmunoResearch 115-095-071) and incubated for 15 min at 4 °C. Finally, 5 μl of premium-grade allophycocyanin (APC)-labeled streptavidin were added to the cells and incubated for 15 min at 4 °C. In each step, cells were washed with DPBS and the sorting buffer was 0.5 ml FACS buffer. FITC⁺ APC⁺ E2mc3 specific B cells were sorted using BD FACS Aria II into Eppendorf tube with 500 μl of FACS buffer.

[00110] Next-generation sequencing (NGS) and bioinformatics analysis of mouse B cells: A 5'-rapid amplification of cDNA ends (RACE) protocol has been reported for unbiased sequencing of mouse B cell repertoires. See He *et al.*, *Sci. Adv.* 4, eaau6769, 2018; and Morris *et al.*, *mBio* 8, e00036-00017, 2017. Here, this protocol was applied to bulk-sorted, E2-specific mouse splenic B cells. Briefly, 5'-RACE cDNA was obtained from bulk-sorted splenic B cells of each mouse with SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (TaKaRa). The immunoglobulin PCRs were set up with Platinum *Taq* High-Fidelity DNA Polymerase (Life Technologies) in a total volume of 50 μl, with 5 μl of cDNA as template, 1 μl of 5'-RACE primer, and 1 μl of 10 μM reverse primer. The 5'-RACE primer contained a PGM/S5 P1 adaptor, while the reverse primer contained a PGM/S5 A adaptor. We adapted the mouse 3'-C_γ1-3/3'-C_μ inner primers and 3'-mC_κ outer primer as reverse primers for 5'-RACE PCR processing of heavy and light (κ) chains. A total of 25 cycles of PCR was performed and the expected PCR products (500-600 bp) were gel purified (Qiagen). NGS was performed on the Ion S5 GeneStudio system. Briefly, heavy and light (κ) chain libraries from the same mouse were quantitated using Qubit® 2.0 Fluorometer with Qubit® dsDNA HS

Assay Kit, and then mixed using a ratio of 3:1 before being pooled with antibody libraries of other mice at an equal ratio for sequencing. Template preparation and (Ion 530) chip loading were performed on Ion Chef using the Ion 520/530 Ext Kit, followed by sequencing on the Ion S5 system with default settings. The mouse *Antibodyomics* pipeline was used to process the raw data and to determine distributions for germline gene usage, somatic hypermutation (SHM), germline divergence, and H/KCDR3 loop length.

[00111] List of some sequences disclosed herein.

[00112] E2 (H77 genotype 1a) (SEQ ID NO:1), HCV polyprotein residues 384–746
 10 ETHVTGGSAGHTTAGLVGLLTPGAKQNIQLINTNGSWHINSTALNCNDSLTTG
 WLAGLFYRHKFNSSGCPERLASCRLTDFAQGWGPISYANGSGLDERPYCWH
 YPPRPGIVPAKSVCGPVYCFTSPVVGTTDRSGAPTYSWGANDTDVFLNN
 TRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGNNTLLCPTDCFRKHPEATY
 SRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGVEHRLEAACNW
 15 TRGERCDLEDRDRSELSPLLLSTTQWQVLPSCFTTLPALSTGLIHLHQNIVDVQY
 LYGVGSSIASWAIKWEYVLLFLLADARVCSCWMLLISQAEA

[00113] E2ΔTM (H77 isolate) (SEQ ID NO:2), HCV polyprotein residues 384–717
 ETHVTGGSAGHTTAGLVGLLTPGAKQNIQLINTNGSWHINSTALNCNDSLTTG
 WLAGLFYRHKFNSSGCPERLASCRLTDFAQGWGPISYANGSGLDERPYCWH
 20 YPPRPGIVPAKSVCGPVYCFTSPVVGTTDRSGAPTYSWGANDTDVFLNN
 TRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGNNTLLCPTDCFRKHPEATY
 SRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGVEHRLEAACNW
 TRGERCDLEDRDRSELSPLLLSTTQWQVLPSCFTTLPALSTGLIHLHQNIVDVQY
 LYGVGSSIASWAIKWE

25 [00114] E2c (H77 isolate) (SEQ ID NO:3), engineered core sequence
 QLINTNGSWHINSTALNCNESLNTGWLGLFYQHKFDSSGCPERLASCGSSGC
WHYPPRPGIVPAKSVCGPVYCFTSPVVGTTDRSGAPTYSWGANDTDVFL
 NNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGDNTLLCPTDCFRKHPE
 ATYSRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGVEHRLEAA
 30 CN (the VR2 disordered region underlined)

[00115] E2c3 (H77 isolate) (SEQ ID NO:4), engineered core sequence

- QLINTNGSWHINSTALNCNESLNTGWLGLFYQHKFDSSGCPERLASCGSSGC
WHYPPRPCGIVPAKSVCGPVYCFTPSPVVVGTTDRSGAPTYSWGANDTDVFL
 NNTRPPLGNWFGCTWMNSTGFTKVCGAPPGPTDGGSGPWITPRCMVDYPYR
 LWHYPCTINYTIFKVRMYVGGVEHRLEAACN (the VR2 disordered region
 5 underlined)
- [00116] E2mc3 (SEQ ID NO:26), redesigned E2 core
 QLINTNGSWHINSTALNCNESLNTGWLGLFYQHKFDSSGCPERASGHYPRPC
 GIVPAKSVCGPVYCFTPSPVVVGTTDRSGAPTYSWGANDTDVFLNNTGNWF
 GCTWMNSTGFTKVCGAPPGPTDGGSGPWITPRCMVDYPYRLWHYPCTINYTI
 10 FKVRMYVGGVEHRLEAACN (tVR2 sequence underlined)
- [00117] E2mc3 variants 1-10 (SEQ ID NOs:27-36): the PERSAG (SEQ ID NO:52)
 motif is replaced respectively with the 6 aa or 5 aa sequences shown in SEQ ID
 NOs:11-20.
- [00118] E2mc3 (HK6a genotype 6a), redesigned E2 core (SEQ ID NO:37)
 15 QLINTNGSWHINRTALNCNDSLQTFITSLFYAKNVDSSGCPERASGHYPRPCD
 VVSARTVCGPVYCFTPSPVVVGTTDKLGIPTYNWGENETDVFMLESLGGWFG
 CTWMNSTGFTKTCGAPPGPTDGGSGPWITPRCLVDYPYRLWHYPCTVNFTL
 HKVRMFVGGIEHRFDAACN (tVR2 sequence underlined)
- [00119] E2mc3 (HK6a genotype 6a) variant 1 (SEQ ID NO:38):
 20 QLINTNGSWHINRTALNCNDSLQTFITSLFYAKNVDSSGCQNWDEPHYPRPC
 DVVSARTVCGPVYCFTPSPVVVGTTDKLGIPTYNWGENETDVFMLESLGGWF
 GCTWMNSTGFTKTCGAPPGPTDGGSGPWITPRCLVDYPYRLWHYPCTVNFT
 LHKVRMFVGGIEHRFDAACN (changed residues in tVR2 underlined)
- [00120] tVR2 sequences of E2mc3 variants 1-10: CQNWDEPHYPRPC (SEQ ID
 25 NO:65), CKVNIDPHYPRPC (SEQ ID NO:66), CEKVEELHYPRPC (SEQ ID NO:67),
 CPDENMKHYPRPC (SEQ ID NO:68), CKREEKMHYPRPC (SEQ ID NO:69),
 CPKTEVHYPRPC (SEQ ID NO:70), CKRVDIHYPRPC (SEQ ID NO:71),
 CPSDMVHYPRPC (SEQ ID NO:72), CPNEEEHYPRPC (SEQ ID NO:73), and
 CKKEIRHYPRPC (SEQ ID NO:74).

30

[00121] The invention thus has been disclosed broadly and illustrated in reference to
 representative embodiments described above. It is understood that various

modifications can be made to the present invention without departing from the spirit and scope thereof.

- [00122]** It is further noted that all publications, sequence accession numbers, patents and patent applications cited herein are hereby expressly incorporated by reference in
- 5 their entirety and for all purposes as if each is individually so denoted. Definitions that are contained in text incorporated by reference are excluded to the extent that they contradict definitions in this disclosure.