

MODIFIED FLAVIVIRUS POLYPEPTIDE AND USES THEREOF

Technical field

- 5 The present invention relates, in general terms, to virology. Disclosed herein are modified flavivirus envelope polypeptides and uses thereof in treating and preventing flaviviral infections.

Background

10

The dengue virus (DENV) is an enveloped flavivirus which infects millions of people worldwide and poses a significant human health and economic burden. There are four circulating DENV serotypes with near identical structures and similar protein sequences. Infection with one serotype may lead to the production of cross-reactive but suboptimal neutralising antibodies which exacerbate subsequent DENV infections in a process known as antibody-dependent enhancement (ADE). Furthermore, changes in environmental conditions such as pH and temperature during the viral life cycle can trigger the viral envelope (E) and membrane (M) proteins to adopt numerous conformations with varying exposure of epitopes. These factors make dengue vaccine development highly challenging from both an immunological as well as a structural point of view. They have also limited the efficacy of current live attenuated vaccines. There is thus a need for more effective vaccines against flaviviruses like dengue virus.

- Virus-like particles (VLPs) can be used to present viral epitopes to the immune system without the risk of infection associated with live attenuated vaccines or the risk of inducing cross-reactive non-neutralising antibodies due to epitopes being presented on immature virions. This makes VLPs promising candidates for use in vaccines. Mature VLPs from dengue virus (DENV) serotype 2 can induce broadly neutralising antibodies against all four serotypes of DENV. However, challenges remain in the production of VLP-based vaccines. The lack of genomic RNA in VLPs tends to reduce particle stability, one of the key factors required for efficacious VLP-based vaccines. Enveloped VLPs such as those derived from flaviviruses are very sensitive to external environmental conditions such as temperature, shear forces and purification processes, which can affect their assembly and epitope display. Furthermore, native DENV VLPs are not secreted efficiently, and it can be costly to produce
- 25
- 30

them in sufficient quantities.

It would be desirable to ameliorate at least one of the above-described problems, or at least to provide a useful alternative.

5

Summary

Disclosed herein is an isolated flavivirus envelope (E) polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain is distinguished from a wild-type stem-transmembrane domain by two or more amino acid substitutions in the wild-type stem-transmembrane domain, wherein the two or more amino acid substitutions are at positions corresponding to positions 8, 11 or 15 of SEQ ID NO: 1.

Disclosed herein is an isolated flavivirus envelope (E) polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain comprises at least two of the following: (a) a non-polar amino acid residue at a position corresponding to position 8 of SEQ ID NO: 1; (b) a non-polar amino acid residue at a position corresponding to position 11 of SEQ ID NO: 1; and (c) a non-polar amino acid residue at a position corresponding to position 15 of SEQ ID NO: 1.

15
20

Disclosed herein is a chimeric flavivirus E polypeptide comprising an ectodomain from a first flavivirus and a stem-transmembrane domain from a second flavivirus, wherein the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to positions 8, 11 or 15 of SEQ ID NO: 1.

25

Disclosed herein is a virus-like particle (VLP) comprising a flavivirus E polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to positions 8, 11 or 15 of SEQ ID NO: 1.

30

Disclosed herein is an isolated polynucleotide comprising a nucleic acid sequence encoding a flavivirus E polypeptide as defined herein.

Disclosed herein is a vector comprising an isolated polynucleotide as defined herein.

Disclosed herein is a host cell comprising an isolated polynucleotide or a vector as defined herein.

- 5 Disclosed herein is a method for producing a VLP, comprising culturing a host cell as defined herein under conditions suitable for expressing the encoded flavivirus E polypeptide to assemble the VLP.

- Disclosed herein is an immunogenic composition comprising a flavivirus E polypeptide, a
10 VLP, an isolated polynucleotide, or a vector as defined herein.

Disclosed herein is a flavivirus E polypeptide, a VLP, an isolated polynucleotide, a vector, or an immunogenic composition as defined herein, for use as a medicament or vaccine.

- 15 Disclosed herein is a method of modulating an immune response in a subject, the method comprising administering a therapeutically effective amount of an immunogenic composition as defined herein to the subject.

- Disclosed herein is a method of preventing or treating a flaviviral infection in a subject, the
20 method comprising administering a therapeutically effective amount of an immunogenic composition as defined herein to the subject.

- Disclosed herein is an immunogenic composition as defined herein, for use in preventing or
treating a flaviviral infection in a subject.

25

Disclosed herein is the use of an immunogenic composition as defined herein in the manufacture of a medicament for preventing or treating a flaviviral infection in a subject.

Brief description of the drawings

30

Embodiments of the present invention will now be described, by way of non-limiting example, with reference to the drawings in which:

Fig. 1 shows electron micrographs of (a) mature DENV-2 VLPs purified from stable

HEK293T cells, (b) JEV VLPs purified from HEK293T cells, (c) mature DENV-2 VLPs purified from transient COS-1 cell transfection, and (d) ZIKV VLPs purified from transient COS-1 cell transfection.

5 **Fig. 2** shows the lipid composition of VLPs produced using different mammalian hosts. (a) Proportions of positive (+) and negative (-) lipids based on total amount of lipids. (b) Proportions of positive (+) and negative (-) lipids based on total M/z.

Fig. 3 shows a comparison of the amino acid sequence of the E protein H1–T2 region of
10 JEV, ZIKV and the four DENV serotypes.

Fig. 4 shows secretion of VLPs containing E protein mutants from HEK-293T cells.

Fig. 5 shows immunofluorescence imaging of HEK-293T cells expression various E protein
15 mutants.

Fig. 6 shows all atom MD simulations of chimeric E-M protein dimer in two different lipid compositions. (A) A cartoon representation of Envelope (E)–M protein dimer showing top and side view. Inset shows a magnified E-H1 helix and the amino acid residues facing the
20 membrane. (B) All atom model of E-M protein dimer embedded in a lipid membrane. The dominant lipid molecules found in the viral lipid envelope are also shown. (C) and (D) RMSD of last 100 ns of 500 ns long simulation trajectory of protein backbone atoms from E_{chi} and E_{chi} mutants embedded in POPC:POPE:POPS (C) and DAG:SAPC:FA (right) lipid membranes. POPC: Phosphatidylcholine (16:0/18:1); POPE: Phosphatidylethanolamine
25 (16:0/18:1); POPS: Phosphatidylserine(16:0/18:1); DAG: Diacylglycerol (18:0/20:4); SAPC: Phosphatidylcholine (18:0/20:4); FA: Fatty acid (18:1).

Fig. 7 shows the per-residue insertion of depth of the E-H1 helix in lipid bilayers. Interactions averaged over last 100 ns of 500 ns of the trajectory for (A) POPC:POPE:POPS
30 or (B) DAG:POPC:FA lipid membranes. (C) Helix insertion depth of DAG-dominant E-M dimer simulation. Each point represents the average over last 200 ns insertion depth (along Z-axis) with respect to lipid head groups. Lipid head groups are defined by a plane corresponding to lipid phosphates in case of POPC and O31 groups in case DAG. Alongside each plot, a snapshot of final coordinates of E-H1 is shown. Error bars were calculated as

standard deviation over 200 ns of the trajectory.

Fig. 8 shows helix-lipid interactions of native dengue E protein and E_{chi} protein in two lipid environments. (A) Per-residue secondary structure propensity over the simulation time for H1 helix (residues 397 to 415) from E and E_{chi} dimer simulations. (B) Per-residue insertion depth of E-H1 helix within lipid bilayers for POPC-dominant (top) and DAG-dominant (bottom) E/E_{chi}-M protein simulations. Each point represents the average over 200 ns insertion depth (along Z-axis) with respect to lipid head groups. (C) Lipid contacts over the simulation time shown as 100 ns block averages in two lipid bilayers: POPC (left) and DAG (right).

Fig. 9. (A) E-M protein structure shown in cartoon representation with the inset highlighting the stem helix 1 (E-H1) from native dengue-JEV E_{chi} protein along with single, double and triple mutants. (B) Plot showing the change in Rosetta Energy Units (REU) upon mutation to any of the other 20 amino acids at critical sites (left). Secretion assay comparing the efficacy of secretion between E_{chi} and E_{chi} with various Leu mutations (right). (C) Helix-lipid interactions and helix insertion into lipid bilayer measured from 500 ns simulations performed on E_{chi}-M dimer embedded in DAG:SAPC:FA membrane.

Fig. 10 shows coarse-grained (CG) simulations of dengue VLPs. (A) Workflow of building the CG model of dengue VLPs using experimentally derived low-resolution cryo-EM density map (EMD:6926). Simulation snapshots of VLPs with different lipid compositions are shown for (B) POPC-dominant lipid VLP models and (C) DAG-dominant VLP models. Snapshots shown for different models involving different number and composition of lipids (VLP-1, VLP-2 and VLP-3): starting structure (0 ns), last frame of 500 ns-long equilibration simulation (Eq.), and 1000 ns and 2000 ns frame from unrestrained simulations. Each unrestrained simulation was performed for 2000 ns and in triplicates. VLP models showing the overall shape and cross-section are shown in bead representations with protein components in yellow and lipid components in grey. The composition and number of lipid molecules in each VLP model is shown in (D)(i) and (D)(ii) respectively. (E) RMSD of VLP backbone beads with respect to the initial structure is shown along the 2000 ns of the unrestrained simulations of three replicates of each model from POPC- (i) and DAG- (ii) dominant lipid VLP.

Fig. 11 shows the lipid composition and number-specific morphology and dynamics of VLPs. (A)-(C) Distribution plots of solvent accessible surface area (SASA) of domains (A) DI, (B) DII and (C) DIII from POPC- (top) and DAG- (bottom) dominant VLP models showing the variation with respect to lipid number. The distribution plots were generated combining SASA values measured from 3 independent 2000 ns-long simulations of each VLP model. (D) Number of contacts between E protein and lipid vesicles over the simulation time is shown for POPC- (i) and DAG- (ii) containing VLPs. (E) Violin plot showing the distribution of average distances between diametrically opposite 12 5-fold symmetric axes calculated at every 1 ns from a cumulative trajectory of 6000 ns from each VLP model in POPC- (i) and DAG- (ii) lipid dominant vesicle system. Schematic illustrates the 6 distances. (F)-(G) First principal component mode (PC1) capturing the coordinated motions of each E protein subunit. PCA was performed on every third protein back bone bead of each VLP trajectory and PC1 accounts for of ~45-50% variation in the data.

Fig. 12 shows a comparison of the extraction efficiency using three different lipid extraction methods applied to virus-like particles.

Fig. 13 shows E-H1 helix-lipid contacts over the simulations time shown as 100 ns block averages in two lipid bilayers: (A) POPC- and (B) DAG- dominant lipid bilayers.

Fig. 14 shows heat maps of secondary structure transition of E-H1 helices from 500 ns long dimeric E protein embedded in (A) POPC- and (B) DAG-dominant lipid membrane trajectory. The secondary structures are colour coded as per the legend.

Fig. 15 shows RMSF values of E-H1 helix protein. A-B) Average RMSF values of E_{chi} protein and respective E-H1 helix mutants embedded in POPC and DAG dominant lipid membrane systems plotted with respect to time for the complete 500 ns trajectory.

Fig. 16 A) Phosphate (PO₄ bead) head group density distribution along X, Y and Z coordinate from VLP CG-model equilibration simulation. Initial and final corresponds to starting (0 ns) and final (500 ns) configuration of lipid vesicle in VLP. B) GL1 group head group density distribution along X, Y and Z coordinate from VLP CG-model equilibration simulation. Initial and final corresponds to starting (0ns) and final (500ns) configuration of lipid vesicle in VLP C-D) Lipid vesicle only simulations of DAG and POPC dominant lipid

systems showing starting and final structures.

Figure 17 shows the effect of mutations in the stem-helix-transmembrane region spanning residues 395 to 495 of JEV. Panels (i) show the per-residue insertion of depth of the (A(i))
 5 F402L/G408L/A409L, (B(i)) F402L/A409L/W420F/F422L, (C(i))
 F402L/A409L/F429L/S431L/G433L, (D(i)) F402L/A409L/K434L/A443L/R445L/T446L,
 (E(i)) F402L/A409L/S452L/W453L/T455L/Q456L/G457L, and (F(i))
 F402L/T405L/G408L/A409L mutants in DAG lipid bilayers. Mutations introduced along
 the stem-helix region of DENV E protein with JEV stem helix-transmembrane region
 10 spanning residue 395 to 495. Panels (ii) show the number of contacts between H1 helices of
 the dimeric E protein mutants with membrane lipid molecules. Alongside each plot, panels
 (iii) show snapshots of starting and final coordinates of stem-helix-transmembrane region in
 cartoon with O31 lipid head group beads are shown in grey transparent spheres is shown.
 Error bars were calculated as standard deviation over 200 ns of the trajectory.

15

Figure 18 are heat maps showing free energy differences ($\Delta\Delta G$) between mutant and wild-type E protein. Residues in the stem helix and transmembrane region were replaced with one of the 19 amino acid residues (except proline). A) $\Delta\Delta G$ heat map for stem helix 1 and 2 spanning residues 396–413 and 427–445 respectively B) $\Delta\Delta G$ heat map of transmembrane
 20 region i.e., TM1 and TM2 spanning residues 451–469 and 476–495 respectively. Negative $\Delta\Delta G$ values (black or darker shades) indicate enhanced protein stability (and thus better secretion) as a result of the mutation, while positive values (white or lighter shades) indicate reduced stability (and thus diminished secretion).

25 Detailed description

The flavivirus envelope (E) protein is translated as part of a single polypeptide encoded by the flaviviral genome and is cleaved by a combination of viral and host proteases to form the E protein in mature virions. Previous studies have shown that mature dengue VLPs are
 30 icosahedral with a smooth surface covered by membrane-anchored E and M proteins. The E proteins are present as homodimers, and each E protein comprises an ectodomain with DI, DII and DIII subdomains; an α -helical stem region with three stem helices (E-H1, E-H2, and E-H3) which lie on the lipid membrane; and a C-terminal transmembrane region containing two transmembrane helices (E-T1 and E-T2). The E protein stem region is believed to

interact with the lipid membrane and thus may influence the stability and secretion of flavivirus VLPs. However, this region is not well characterised structurally, and the contribution of particular residues to protein interactions with the lipid membrane is not well understood.

5

The inventors have discovered that flavivirus virus-like particles (VLPs) may be engineered for improved stability and secretion by introducing specific mutations to the envelope (E) protein to modify protein-protein and protein-lipid interactions. Without being bound by theory, three structural motifs of flavivirus VLPs predominantly interact with the envelope
10 lipids: i) the envelope protein H1 helix (E-H1, corresponding to residues 2–26 of SEQ ID NO: 1); ii) the envelope protein H2 helix (E-H2, corresponding to residues 34–53 of SEQ ID NO: 1); and iii) the TM helices (E-TM, corresponding to residues 56–101 of SEQ ID NO: 1). CryoEM data has shown that the E-H1 helix interacts with the outer leaflet and is positioned with a slight “tilt” with respect to the membrane surface. Altering residues in the
15 H1 helix that are pointing towards the lipid membrane may strengthen protein-lipid interactions to provide a more robust VLP envelope and thus a more stable VLP.

The inventors used a combination of empirical and computational analyses to elucidate lipid-E protein interactions and their contribution to VLP stability and secretion. The dengue VLP
20 lipidome was profiled to generate near-native structural models of VLP lipid membranes with embedded E-M protein dimers. Site-directed mutagenesis of residues in the stem-transmembrane region and molecular dynamics simulations of the mutants showed that having non-polar residues in this region can improve interactions of the E protein with the VLP lipid membrane. Without being bound by theory, the presence of non-polar amino acids
25 like leucine, isoleucine and phenylalanine at these sites can increase protein-lipid interactions and stabilise the protein within the lipid bilayer during maturation and secretion of VLPs. Substitution of select residues in E-H1 with non-polar amino acids was found to particularly improve E protein stability.

30 Accordingly, this disclosure provides flavivirus envelope (E) polypeptides containing stem-transmembrane domains in which two or more amino acid residues have been substituted. In preferred embodiments, the substitutions are to non-polar amino acid residues. Also provided are virus-like particles (VLPs) containing the modified E polypeptides, and use of

the modified E polypeptides and VLPs for treating and preventing flaviviral infections.

- As used herein, the term “flavivirus” encompasses all viruses within the Flaviviridae family as classified under the Baltimore Classification System (BCS) and/or the International Committee on Taxonomy of Viruses (ICTV). Non-limiting examples of suitable flaviviruses include the following: Alfuy virus, Aroa virus, Bagaza virus, Banzi virus, Bouboui virus, Bussuquara virus, Cacipacore virus, Dengue virus (DENV, e.g., Dengue virus serotype I, Dengue virus serotype II, Dengue virus serotype III, Dengue virus serotype IV), Edge Hill virus, Gadgets Gully virus, Kadam virus, Kunjin virus, Kokobera virus, Kyasanur Forest disease virus, Iguape virus, Ilheus virus, Israel turkey meningoencephalomyelitis virus, Japanese encephalitis virus (JEV), Jugra virus, Kedougou virus, Koutango virus, Langat virus, Louping ill virus, Meaban virus, Modoc Virus, Murray Valley encephalitis virus (MVEV), Naranjal virus, Negeishi Virus, Ntaya virus, Omsk hemorrhagic fever virus, Potiskum virus, Powassan virus, Rio Bravo Virus, Rocio virus, Russian Spring Summer Encephalitis Virus, Royal Farm virus, Saboya virus, Saumarez Reef virus, Sepik virus, St. Louis encephalitis virus, Spondweni virus, Stratford virus, Tick-borne encephalitis virus, Tembusu virus, Tyuleny virus, Uganda S virus, Usutu virus, Wesselsbron virus, West Nile virus (WNV), Yellow fever virus (YFV), Yaounde virus, and Zika virus (ZIKV).
- The term “isolated” refers to a biological material, such as a polynucleotide or polypeptide, which is substantially free from components that normally accompany or interact with it in its naturally occurring environment. The isolated material optionally comprises material not found with the material in its natural environment, e.g., a cell or a virion.
- Typically, a polypeptide is isolated when it is at least 70%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 80%, more preferably at least 90%, and most preferably at least 99%, by weight, of a polypeptide of this disclosure. An isolated polypeptide of this disclosure may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesising the polypeptide. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

The term “flavivirus envelope (E) polypeptide” or “flavivirus E protein” refers to the

structural envelope protein found on the surface of a flavivirus or VLP which is responsible for fusion of the virus or VLP with the host cell membrane. The E protein ectodomain is the N-terminal region of the E protein which is exposed on the surface of the virion or VLP and primarily responsible for host cell receptor binding and fusion with host cell membranes.

- 5 The stem-transmembrane domain at the C-terminus connects and anchors the ectodomain to the viral or VLP membrane. The ectodomain contains the DI, DII and DIII subdomains and may be about 380–450 amino acids in length. The stem-transmembrane domain contains the H1, H2, H3, T1 and T2 helices and may be about 80–150 amino acids in length. The lengths of these domains can vary among different flavivirus species. Ectodomains and stem-transmembrane domains of various flavivirus E proteins (such as JEV and DENV) are known in the art. Exemplary E protein ectodomain and stem-transmembrane amino sequences are provided in Tables 1 and 2. The skilled person can identify homologous domains from other flaviviruses by sequence alignment or domain homology analysis.

- 15 As used herein, the term “wild-type” refers to a gene or gene product (e.g., protein) that has the characteristics (e.g., sequence) of that gene or gene product isolated from a naturally occurring source. In contrast, the term “mutant” or “variant” refers to a gene or gene product that displays modifications in sequence when compared to the wild-type gene or gene product. Mutant genes or gene products may be naturally-occurring or synthetic (i.e., containing altered sequences that do not occur in nature).

- 25 As used herein, “a position corresponding to” or recitation that amino acid positions “correspond to” amino acid positions in a disclosed sequence, such as set forth in the sequence listing, refers to amino acid positions identified upon alignment with the disclosed sequence to maximise identity using a standard alignment algorithm or software (such as the BLAST, ClustalW, ClustalOmega, MUSCLE, TCOFFEE or ProbCons programmes). By aligning the sequences, one skilled in the art can identify corresponding residues.

- 30 As used herein, the term “sequence identity” refers to the extent that sequences are identical on an amino acid-by-amino acid basis over a window of comparison. Thus, a “percentage of sequence identity” is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical amino

acid residue (e.g., Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr and Val) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. Sequence identity may be measured using sequence analysis software (for example, BLAST, ClustalW, ClustalOmega, MUSCLE, TCOFFEE or ProbCons programmes).

Sequence variations may arise from conservative amino acid substitutions. A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, which can be generally sub-classified as follows:

Table 1. Amino acid sub-classification.

Sub-classes	Amino acids
Acidic	Aspartic acid (D), Glutamic acid (E)
Basic	Noncyclic: Arginine (R), Lysine (K); Cyclic: Histidine (H)
Charged	Aspartic acid (D), Glutamic acid (E), Arginine (R), Lysine (K), Histidine (H)
Small	Glycine (G), Serine (S), Alanine (A), Threonine (T), Proline (P)
Polar/neutral	Asparagine (N), Histidine (H), Glutamine (Q), Cysteine (C), Serine (S), Threonine (T)
Polar/large	Asparagine (N), Glutamine (Q)
Non-polar	Tyrosine (Y), Valine (V), Isoleucine (I), Leucine (L), Methionine (M), Phenylalanine (F), Tryptophan (W)
Aromatic	Tryptophan (W), Tyrosine (Y), Phenylalanine (F)
Residues that influence chain orientation	Glycine (G) and Proline (P)

Conservative amino acid substitution also includes groupings based on side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is

asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. For example, it is reasonable to expect that replacement
5 of a valine with an methionine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid will not have a major effect on the properties of the resulting mutant polypeptide.

As used herein, a “chimeric polypeptide” refers to a polypeptide containing sequences
10 derived from two or more different proteins or protein domains. These sequences can come from homologous proteins (e.g., proteins or protein domains from different species), or may be artificially combined to confer new or enhanced functions, e.g., improved stability or activity.

As used herein, the term “virus-like particle (VLP)” refers to a structure that in at least one
15 attribute resembles a virus, but which is non-infectious. In general, a VLP closely resembles a virus in the assembly of its outer envelope or protein coat. VLPs of this disclosure are non-replicative (i.e., they are incapable of replicating any component genetic information) and do not carry genetic information in a form that would enable expression of their component
20 proteins. VLPs can often be produced in large quantities by heterologous expression of viral coat or envelope proteins in a suitable host cell.

By “isolated polynucleotide” is meant a nucleic acid molecule (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic
25 acid molecule is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or that exists as a separate molecule independent of other sequences (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion). In addition, the
30 term includes an RNA molecule that is transcribed from a DNA.

An isolated polynucleotide can be, but need not be, substantially purified. For example, a polynucleotide that is isolated within a cloning or expression vector is not pure in that it may comprise only a small percentage of the material in the vector or in the cell in which the

vector resides.

The term “construct” generally refers to recombinant nucleic acid, generally recombinant DNA, that has been generated for the purpose of the expression of a specific nucleic acid sequence (i.e., an “expression construct”), or is to be used in the construction of other recombinant nucleotide sequences (i.e., a “cloning construct”). The construct may be contained within a vector. Two or more constructs can be contained within a single nucleic acid molecule, such as a single vector, or can be contained within two or more separate nucleic acid molecules, such as two or more separate vectors.

An “expression construct” generally includes at least an element operably linked to a nucleotide sequence of interest to direct expression of the nucleic acid sequence in an organism or part thereof including a host cell. Such elements may include control elements such as a promoter that is operably linked to (so as to direct transcription of) the nucleic acid sequence of interest, and often includes a polyadenylation sequence as well. For the practice of the present invention, conventional compositions and methods for preparing and using constructs and host cells are well known to one skilled in the art, see for example, *Molecular Cloning: A Laboratory Manual*, 3rd edition Volumes 1, 2, and 3. J. F. Sambrook, D. W. Russell, and N. Irwin, Cold Spring Harbor Laboratory Press, 2000.

As used herein “operably linked” is the association of two or more nucleic acid fragments in a construct such that the function of one is controlled by the other, for example DNA encoding a protein associated with DNA encoding a promoter.

By “control element” or “control sequence” is meant nucleic acid sequences (e.g., DNA) necessary for expression of an operably linked coding sequence in a particular host cell. The control sequences that are suitable for prokaryotic cells for example, include a promoter, and optionally a cis-acting sequence such as an operator sequence and a ribosome binding site. Control sequences that are suitable for eukaryotic cells include transcriptional control sequences such as promoters, polyadenylation signals, transcriptional enhancers, translational control sequences such as translational enhancers and internal ribosome binding sites (IRES), nucleic acid sequences that modulate mRNA stability, as well as targeting sequences that target a product encoded by a transcribed polynucleotide to an intracellular compartment within a cell or to the extracellular environment. Promoters suitable for use

with expression constructs or vectors of the present disclosure include but are not limited to the CMV promoter, phage lambda PL promoter, the *E. coli* lac, phoA and tac promoters, the SV40 early and late promoters, and promoters of retroviral LTRs.

- 5 By “vector” is meant a polynucleotide molecule, suitably a DNA molecule derived, for example, from a plasmid, bacteriophage, yeast or virus, into which a polynucleotide can be inserted or cloned. A vector may contain one or more unique restriction sites or multiple cloning sites and can be capable of autonomous replication in a defined host cell including a target cell or tissue or a progenitor cell or tissue thereof, or be integrated with the genome
- 10 of the defined host such that the cloned sequence is reproducible. Accordingly, the vector can be an autonomously replicating vector, i.e., a vector that exists as an extra-chromosomal entity, the replication of which is independent of chromosomal replication, e.g., a linear or closed circular plasmid, an extra-chromosomal element, a mini-chromosome, or an artificial chromosome. The vector can contain any means for assuring self-replication. Alternatively,
- 15 the vector can be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. A vector system may comprise a single vector or plasmid, two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The choice of the vector will typically depend on the compatibility of the vector
- 20 with the host cell into which the vector is to be introduced. The vector may also include a selection marker such as an antibiotic resistance gene that can be used for selection of suitable transformants. Examples of such resistance genes are well known to those of skill in the art.
- 25 The terms “host cell”, “host cell line” and “host cell culture” are used interchangeably and refer to cells into which exogenous nucleic acid has been introduced, including the progeny of such cells. Host cells include “transformants” and “transformed cells”, which include the primary transformed cell and progeny derived therefrom without regard to the number of passages. Progeny may not be completely identical in nucleic acid content to a parent cell
- 30 but may contain mutations. Mutant progeny that have the same function or biological activity as screened or selected for in the originally transformed cell are included herein. A host cell is any type of cellular system that can be used to generate the flavivirus polypeptides and/or VLPs of the present invention. Host cells include cultured cells, e.g., mammalian cultured cells, such as CHO cells, BHK cells, NS0 cells, SP2/0 cells, YO myeloma cells, P3X63

mouse myeloma cells, PER cells, PER.C6 cells or hybridoma cells, yeast cells, insect cells, and plant cells, to name only a few, but also cells comprised within a transgenic animal, transgenic plant or cultured plant or animal tissue.

- 5 The term “subject” as used herein refers to an animal, in particular a mammal and more particularly a primate including a lower primate and even more particularly, a human who can benefit from the present disclosure. A subject regardless of whether a human or non-human animal may be referred to as an individual, subject, animal, patient, host or recipient. For convenience, an “animal” specifically includes livestock animals such as cattle, horses,
10 sheep, pigs, camelids, goats and donkeys, as well as domestic animals, such as dogs and cats. Examples of laboratory test animals include mice, rats, rabbits, guinea pigs and hamsters. Rabbits and rodent animals, such as rats and mice, provide a convenient test system or animal model as do primates and lower primates. In one embodiment, the subject is human. Subjects in need of treatment include, but are not limited to, individuals already having a particular
15 infection or condition as well as individuals who are at risk of acquiring a particular infection or condition.

- The term “treating” as used herein may refer to (1) inhibiting the development of the infection or one or more symptoms of the infection; (2) relieving the infection, i.e., causing
20 regression of the infection or at least one or more symptoms of the infection; and/or (3) causing a decrease in the severity of one or more symptoms of the infection. The term “preventing” refers to inhibiting or delaying the appearance of one or more symptoms of the infection. It is to be understood that terms “treating and “preventing” do not imply that the infection, or a symptom thereof, is permanently inhibited, relieved, reduced or alleviated and
25 therefore also encompasses the temporary inhibition, relief, reduction or alleviation of the disease, or of a symptom thereof.

- The term “modulating an immune response” may comprise inducing an immune response in a subject. This includes stimulating an immune response and/or enhancing an existing
30 immune response.

As used herein, a “therapeutically effective amount” or “effective amount” is an amount that is non-toxic to the subject and sufficient to effect a desired outcome in a subject. For example, an effective amount of a VLP may refer to an amount of VLPs necessary or

sufficient to induce immunity or to prevent or treat an infection in the subject. An effective amount may also be the amount sufficient to enhance a subject's own immune response against a subsequent exposure to an infectious agent. The effective amount of an active agent will vary depending upon the manner of administration, the severity of the condition being
5 treated, and the age, body weight, and general health of the subject. A clinician or veterinarian can determine the appropriate dose and dosage regimen.

The term "pharmaceutically acceptable" herein refers to compositions that do not substantially produce adverse biological effects, e.g., toxic, allergic, or immunological
10 reactions, when administered to a subject.

As used herein, the term "pharmaceutically acceptable carrier" refers to any of the standard pharmaceutical carriers including, but not limited to, water, phosphate buffered saline solution, emulsions (e.g., such as an oil/water or water/oil emulsions), various types of
15 wetting agents, any and all solvents, dispersion media, coatings, sodium lauryl sulfate, isotonic and absorption delaying agents, disintegrants (e.g., potato starch or sodium starch glycolate), and the like. The compositions also can include stabilisers and preservatives. For examples of carriers, stabilisers and adjuvants, see, e.g., Martin, *Remington's Pharmaceutical Sciences*, 15th Ed., Mack Publ. Co., Easton, Pa. (1975), incorporated herein
20 by reference in its entirety.

As used herein, the term "adjuvant" refers to a compound that, when used in combination with a specific immunogen (e.g., a VLP) in a formulation, will augment or otherwise alter or modify the resultant immune response. Modification of the immune response includes
25 intensification or broadening the specificity of either or both humoral and cellular immune responses. Modification of the immune response can also mean decreasing or suppressing certain antigen-specific immune responses.

As used herein, the term "multivalent" refers to an immunogenic composition or VLP which
30 contains multiple antigenic components against multiple types or strains of infectious agents.

Envelope protein stem-transmembrane modifications

Flavivirus E polypeptides of this disclosure are modified to contain two or more amino acid

substitutions in the stem-transmembrane domain, e.g., in the stem and/or transmembrane helices. Preferred substitutions are non-polar amino acid substitutions. The substitutions may be at any position corresponding to positions 1 to 101 of SEQ ID NO: 1. Alternatively or in addition, the substitutions may be at the interface of the ectodomain and stem-transmembrane domain, for example, at positions corresponding to positions 1–46, 5 313–321, 339–351, 368–380 of SEQ ID NO: 25.

Disclosed herein is an isolated flavivirus envelope (E) polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain 10 is distinguished from a wild-type stem-transmembrane domain by two or more amino acid substitutions in the wild-type stem-transmembrane domain, wherein the two or more amino acid substitutions are at positions corresponding to position 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 15 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or 101 of SEQ ID NO: 1.

Wild-type E proteins of various flavivirus such as JEV, ZIKV and DENV serotypes 1–4 have been identified and are available at various public databases such as the GenBank 20 database. For example, DENV type 1: Accession No. AAN32773; DENV type 2: Accession No. AIW42846; DENV type 3: Accession No. ADI78875; DENV type 4: Accession No. AFV69144; ZIKV: Accession No. ANF28857; JEV: Accession No. AEV57502. The E protein of a flavivirus such as dengue virus (DENV) may be obtained from a database and aligned with SEQ ID NO: 1 to identify a position corresponding to an amino acid position 25 of SEQ ID NO: 1, such as a position corresponding to positions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 and/or 101 of SEQ ID 30 NO: 1.

In some embodiments, the E polypeptide comprises a stem-transmembrane domain derived from JEV (e.g., with the sequence of SEQ ID NO: 1, or a sequence having at least 70%

sequence identity to a sequence in SEQ ID NO: 1), in which two or more amino acids have been substituted. In some embodiments, the E polypeptide comprises a stem-transmembrane domain derived from DENV (e.g., with the sequence of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 or SEQ ID NO: 5, or a sequence having at least 70% sequence identity to a
5 sequence in SEQ ID NO: 2–5), in which two or more amino acids have been substituted. In some embodiments, the E polypeptide comprises a stem-transmembrane domain derived from ZIKV (e.g., with the sequence of SEQ ID NO: 6, or a sequence having at least 70% sequence identity to a sequence in SEQ ID NO: 6), in which two or more amino acids have been substituted.

10

In some embodiments, the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to position 8, 11, 14, 15, 19, 26, 28, 35, 37, 39, 40, 49, 51, 52, 58, 59, 60, 61, 62 or 63 of SEQ ID NO: 1.

15 In some embodiments, the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to position 8, 11 or 15 of SEQ ID NO: 1.

In some embodiments, the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to position 8, 11 or 15 of SEQ ID NO: 1, and further

20 comprises one or more amino acid substitutions at positions corresponding to position 1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or 101 of
25 SEQ ID NO: 1.

In some embodiments, the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to position 8, 11 or 15 of SEQ ID NO: 1, and further comprises one or more amino acid substitutions at positions corresponding to position 14,

30 19, 26, 28, 35, 37, 39, 40, 49, 51, 52, 58, 59, 60, 61, 62 or 63 of SEQ ID NO: 1.

In some embodiments, the E polypeptide comprises an amino acid substitution at positions corresponding to one of the following sets of positions of SEQ ID NO: 1:

(a) 8 and 11;

- (b) 8 and 15;
- (c) 11 and 15;
- (d) 8, 11 and 15;
- (e) 8, 14 and 15;
- 5 (f) 8, 11, 14 and 15;
- (g) 8, 15, 26 and 28;
- (h) 8, 15, 35, 37 and 39;
- (i) 8, 15, 40, 49, 51 and 52; and
- (j) 8, 15, 58, 59, 61, 62 and 63.

10

In some embodiments, each of the two or more amino acid substitutions is a substitution to a non-polar amino acid residue. Thus, the stem-transmembrane domain may comprise two or more non-polar amino acid substitutions, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 15 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or 101 non-polar amino acid substitutions at positions corresponding to positions 1–101 of SEQ ID NO: 1.

- 20 Disclosed herein is an isolated flavivirus envelope (E) polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain comprises a non-polar amino acid residue at two or more positions corresponding position 8, 11, 14, 15, 19, 26, 28, 35, 37, 39, 40, 49, 51, 52, 58, 59, 60, 61, 62 or 63 of SEQ ID NO: 1.

25

In some embodiments, the stem-transmembrane domain comprises at least two of the following: (a) a non-polar amino acid residue at a position corresponding to position 8 of SEQ ID NO: 1; (b) a non-polar amino acid residue at a position corresponding to position 11 of SEQ ID NO: 1; and (c) a non-polar amino acid residue at a position corresponding to position 15 of SEQ ID NO: 1.

30

In some embodiments, the stem-transmembrane domain comprises a non-polar amino acid residue at two or more positions corresponding to position 8, 11 or 15 of SEQ ID NO: 1, and further comprises a non-polar amino acid residue at one or more positions corresponding to

position 14, 19, 26, 28, 35, 37, 39, 40, 49, 51, 52, 58, 59, 60, 61, 62 or 63 of SEQ ID NO: 1.

In some embodiments, the stem-transmembrane domain comprises a non-polar amino acid at positions corresponding to one of the following sets of positions of SEQ ID NO: 1:

- 5 (a) 8 and 11;
- (b) 8 and 15;
- (c) 11 and 15;
- (d) 8, 11 and 15;
- (e) 8, 14 and 15;
- 10 (f) 8, 11, 14 and 15;
- (g) 8, 15, 26 and 28;
- (h) 8, 15, 35, 37 and 39;
- (i) 8, 15, 40, 49, 51 and 52; and
- (j) 8, 15, 58, 59, 61, 62 and 63.

15

Non-polar amino acid residues herein may be alanine (A), phenylalanine (F), isoleucine (I), leucine (L), methionine (M), proline (P), valine (V) or tryptophan (W). In some embodiments, the non-polar amino acid residue is isoleucine (I), leucine (L) or phenylalanine (F).

20

In some embodiments, the stem-transmembrane domain comprises at least two of the following:

- (a) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 8 of SEQ ID NO: 1;
- 25 (b) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 11 of SEQ ID NO: 1;
- (c) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 14 of SEQ ID NO: 1;
- (d) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 15 of SEQ ID NO: 1;
- 30 (e) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 26 of SEQ ID NO: 1;
- (f) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 28 of SEQ ID NO: 1;

- (g) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 35 of SEQ ID NO: 1;
- (h) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 37 of SEQ ID NO: 1;
- 5 (i) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 39 of SEQ ID NO: 1;
- (j) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 40 of SEQ ID NO: 1;
- (k) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 10 49 of SEQ ID NO: 1;
- (l) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 51 of SEQ ID NO: 1;
- (m) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 52 of SEQ ID NO: 1;
- 15 (n) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 58 of SEQ ID NO: 1;
- (o) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 59 of SEQ ID NO: 1;
- (p) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 20 61 of SEQ ID NO: 1;
- (q) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 62 of SEQ ID NO: 1; or
- (r) a F, I or L residue, or a substitution to F, I or L at a position corresponding to position 63 of SEQ ID NO: 1;

25

In some embodiments, the stem-transmembrane domain comprises at least two of the following:

- (a) a L or a substitution to L at a position corresponding to position 8 of SEQ ID NO: 1;
- (b) a L or a substitution to L at a position corresponding to position 11 of SEQ ID NO: 1;
- 30 (c) a L or a substitution to L at a position corresponding to position 14 of SEQ ID NO: 1;
- (d) a L or a substitution to L at a position corresponding to position 15 of SEQ ID NO: 1;
- (e) a F or a substitution to F at a position corresponding to position 26 of SEQ ID NO: 1;
- (f) a L or a substitution to L at a position corresponding to position 28 of SEQ ID NO: 1;
- (g) a L or a substitution to L at a position corresponding to position 35 of SEQ ID NO: 1;

- (h) a L or a substitution to L at a position corresponding to position 37 of SEQ ID NO: 1;
- (i) a L or a substitution to L at a position corresponding to position 39 of SEQ ID NO: 1;
- (j) a L or a substitution to L at a position corresponding to position 40 of SEQ ID NO: 1;
- (k) a L or a substitution to L at a position corresponding to position 49 of SEQ ID NO: 1;
- 5 (l) a L or a substitution to L at a position corresponding to position 51 of SEQ ID NO: 1;
- (m) a L or a substitution to L at a position corresponding to position 52 of SEQ ID NO: 1;
- (n) a L or a substitution to L at a position corresponding to position 58 of SEQ ID NO: 1;
- (o) a L or a substitution to L at a position corresponding to position 59 of SEQ ID NO: 1;
- (p) a L or a substitution to L at a position corresponding to position 61 of SEQ ID NO: 1;
- 10 (q) a L or a substitution to L at a position corresponding to position 62 of SEQ ID NO: 1;
- or
- (r) a I or a substitution to I at a position corresponding to position 63 of SEQ ID NO: 1;

In some embodiments, the two or more amino acid substitutions herein are selected from

15 F8I, F8L, T11I, T11L, T11F, A15I, A15F and A15L, wherein the first and last letters indicate the wild-type residue and the substituted residue respectively, and the numeral indicates the amino acid position in SEQ ID NO: 1.

In some embodiments, the two or more amino acid substitutions are selected from G14F,

20 G14I, G14L, A19F, A19I, A19L, W26F, W26I, W26L, F28I, F28L, F35I, F35L, S37F, S37I, S37L, G39F, G39I, G39L, K40F, K40I, K40L, A49F, A49I, A49L, R51F, R51I, R51L, T52F, T52I, T52L, S58F, S58I, S58L, W59F, W59I, W59L, I60F, I60L, T61F, T61I, T61L, Q62F, Q62I, Q62L, G63F, G63I or G63L.

25 The E polypeptide may further contain amino acid substitutions at the interface of the ectodomain and transmembrane domain, i.e., at regions of the folded E protein where the ectodomain and stem-transmembrane domain are in close proximity, e.g., within 10 Å of each other. Thus, in some embodiments, the E polypeptide further comprises one or more amino acid substitutions at positions corresponding to positions 1–46, 313–321, 339–351,

30 368–380 of SEQ ID NO: 25.

In some embodiments, the stem-transmembrane domain comprises or consists of an amino acid sequence having at least 80% sequence identity (such as 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity) to an amino acid

sequence set forth in SEQ ID NO: 7–16. In some embodiments, the stem-transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 7–16.

The amino acid substitutions of this disclosure can improve one or more properties of the
5 flaviviral E protein or a virion or virus-like particle (VLP) containing the E protein. Such improvements may include, but are not limited to, increased stability of the E protein in a lipid bilayer or lipid composition; increased efficiency of E protein self-assembly; increased interaction of the E protein with other components of the flaviviral envelope, such as the M protein; increased stability of a flaviviral envelope containing the E protein; increased
10 stability of a virion or VLP containing the E protein; and/or increased production titres of virions or VLPs containing the E protein.

In some embodiments, production of VLPs containing the modified E polypeptide in a suitable host cell is increased compared to production of VLPs containing a corresponding
15 wild-type E polypeptide. The increase in VLP production may be an increase of at least about 10%, for example, about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 100%, or about 2-fold, about 3-fold, about 4-fold, about 5-fold, or greater than 5-fold when compared to production of VLPs containing a corresponding wild-type E polypeptide.

20

In preferred embodiments, amino acid substitutions herein do not affect the immunogenicity of the E protein when the E protein is incorporated into a virion or VLP.

VLP production and immunogenicity can be evaluated using methods known in the art. For
25 example, secreted titres of VLPs can be determined by immunoassay using antibodies targeting the E protein or another component of the VLP envelope. Immunogenicity of an E protein or VLP can be determined using *in vitro* or *in vivo* cytokine release assays or T-cell activation assays.

30

Table 2. Amino acid sequence of exemplary flavivirus E protein stem-transmembrane domains.

SEQ ID NO.	E protein	Stem-transmembrane amino acid sequence
1	Wild-type Japanese encephalitis virus (JEV)	GSTLGKAFSTTLKGAQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
2	Wild-type dengue virus type 1 (DENV1)	GSSIGKMLEATARGARRMAILGDTAWDFGSIGG VFTSVGKLVHQIFGTAYGVLFSGVSWTMKIGIGI LLTWLGLNSRSASLSMTCIAVGMVTLYLGVMV QA
3	Wild-type dengue virus type 2 (DENV2)	GSSIGQMFETTMRGAKRMAILGDTAWDFGSLG GVFTSIGKALHQVFGAIFYGAAFSGVSWTMKILIG VIITWIGMNSRSTLSVSLVLVGVTLYLGVMV QA
4	Wild-type dengue virus type 3 (DENV3)	GSSIGKMFETARGARRMAILGDTAWDFGSVG GVLNSLGKMHQIFGSAYTALFSGVSWIMKIGIG VLLTWIGLNSKNTSMFSCIAIGIITLYLGVVVQA
5	Wild-type dengue virus type 4 (DENV4)	GSSIGKMFESTYRGAKRMAILGETAWDFGSVGG LLTSLGKAVHQVFGSVYTTMFGGVSWMVRILIG LLVLWIGTNSRNTSMAMSCIAVGGITLFLGFTVH A
6	Wild-type Zika virus (ZIKV)	GSTIGKAFEATVRGAKRMAVLGDTAWDFGSVG GALNSLGKGIHQIFGAAFKSLFGGMSWFSQILIG TLLMWLGLNTKNGSISLMCLALGGVLIFLSTAVS A
7	JEV F402L / T405L	GSTLGKALSTLLKGAQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA

8	JEV F402L / A409L	GSTLGKALSTTLKGLQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
9	JEV T405 / A409L	GSTLGKAFSTLLKGLQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
10	JEV F402 / T405 / A409	GSTLGKALSTLLKGLQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
11	F402L / G408L / A409L	GSTLGKALSTTLKLLQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
12	F402L / A409L / W420F / F422L	GSTLGKALSTTLKGLQRLAALGDTAFDLGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
13	F402L / T405L / G408L / A409L	GSTLGKALSTLLKLLQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
14	F402L / A409L / F429L / S431L / G433L	GSTLGKALSTTLKGLQRLAALGDTAWDFGSIGG VLNLILKAVHQVFGGAFRTLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA
15	F402L / A409L / K434L / A443L / R445L / T446L	GSTLGKALSTTLKGLQRLAALGDTAWDFGSIGG VFNSIGLAVHQVFGGFLLLFGGMSWITQGLMG ALLLWMGVNARDRSIALAFLATGGVLVFLATN VHA

16	F402L / A409L / S452L / W453L / T455L / Q456L / G457I	GSTLGKALSTTLKGLQRLAALGDTAWDFGSIGG VFNSIGKAVHQVFGGAFRTLFGGM <u>LLILL</u> LMGA LLLWMGVNARDRSIALAFLATGGVLVFLATNV HA
----	--	---

Table 3. Amino acid sequence of exemplary flavivirus E protein ectodomains.

SEQ ID NO.	E protein	Stem-transmembrane amino acid sequence
17	Japanese encephalitis virus (JEV)	FHCLGMGNRDFIEGASGATWVDLVLEGDSCLTI MANDKPTLDVRMINIEASQLAEVRSYCYHASVT DISTVARCPTTGEAHNEKRADSSYVCKQGFTDR GWGKGCGLFGKGSIDTCAKFSCTSKAIGRTIQPE NIKYEVGIFVHGTTTTSENHGNYSAQVGASQAAK FTVTPNAPSITLKLGDYGEVTLDCPRSGLNTEA FYVMTVGSKSFLVHREWFHDLPLPWTSPTSSTAW RNRELLMEFEEAHATKQSVVALGSQEGSLHQAL AGAIVVEYSSSVKLTSGHLKCRLKMDKLALKGT TYGMCTEKFSFAKNPADTGHGTVVIELTYSGSD GPCKIPIVSVASLNDMTPVGRLVTNPFVATSSA NSKVLLEIEPPFGDSYIVVGRKDKQINHHRHKA
18	Dengue virus type 1 (DENV1)	MRCVGIGNRDFVEGLSGATWVDVVLEHGSCVT TMAKDKPTLDIELLKTEVTNPAVLRKLCIEAKIS NTTDSRCPTQGEATLVEEQDANFVCRRTFVDR GWGNGCGLFGKGSITCAKFKCVTKLEGKIVQY ENLKYSVIVTVHTGDQHQVGNETTEHGTIATITP QAPTSEIQLTDYGALTLDCSPRTGLDFNEMVLLT MKEKSWLVHKQWFLDLPLPWTSFASTPQETWN REDLLVTFKTAHAKKQEVAVLGSQEGAMHTAL TGATEIQTSGTTKIFAGHLKCRLKMDKLTLKGM SYVMCTGSFKLEKEVAETQHGTVLVQVKYEGT DAPCKIPFSTQDEKGVTONGRRLITANPIVTDKEK PVNIEAEPFPGESYIVVGAGEKALKLSWFKK

19	Dengue virus type 2 (DENV2)	MRCIGMSNRDFVEGVSGGSWVDIVLEHGSCVTT MAKNKPTLDFELIKTEAKQPATLRKYCIEAKLTN TTTESRCPTQGEPSLNEEQDKRFVCKHSMVDRG WGNGCGLFGKGGIVTCAMFRCKKNMEGKV VQ PENLEYTIVITPHSGEEHAVGNDTGKHGKEIKITP QSSITEAELTGYGTVTMECSPRTGLDFNEMVLLQ MENKAWLVHRQWFLDLPLPWLPGADTQGSNWI QKETLVTFKNPHAKKQDVVVLGSQEGAMHTAL TGATEIQMSSGNLLFTGHLKCRLRMDKLQLKG MSYSMCTGKFKVVKEIAETQHGTIVIRVQYEGD GSPCKIPFEIMDLEKRHVLGRLITVNPVTEKDSP VNIEAEPFPGDSYIIIGVEPGQLKLNWFKK
20	Dengue virus type 3 (DENV3)	MRCVGVGNRDFVEGLSGATWVDVLEHGGCV TTMAKNKPTLDIELQKTEATQLATLRKLCIEGKI TNITDSRCPTQGEAILPEEQDQNYVCKHTYVDR GWGNGCGLFGKGSVLTCAKFQCLESEGKV VQ HENLKYTVIITAHTGDQHQVGNETQGVTAETSQ ASTAEAILPEYGTGLGLECSPRTGLDFNEMILLTM KNKAWMVHRQWFFDLPLPWTSGATTKTPTWN RKELLVTFKNAHAKKQEVVVLGSQEGAMHTAL TGATEIQTSGGTSIFAGHLKCRLKMDKLELKGM SYAMCLNTFVLKKEVSETQHGTILIKVEYKGED APCKIPFSTEDGQGKAHNGRLITANPVVTKKEEP VNIEAEPFPGESNIVIGIGDKALKINWYRK

21	Dengue virus type 4 (DENV4)	MRCVGVGNRDFVEGVSGGAWVDLVLEHGGCV TTMAQGKPTLDFELIKTTAKEVALLRTYCIASIS NITTATRCPTQGEPYLKEEQDQQYICPRDVVDRG WGNGCGLFGKGGVVTCAKFSCSGKITGHLVQIE NLEYTAVVTVHNGDTHAVGNDPSNPGEPAITP RSPSVEVKLPDYGELTLDCEPRSGIDFNEMILMK MKTKTWLVHKQWFLDLPLPWTAGADTSEVHW NHKERMVTFKVPFAKRQDVTVLGSQEGAMHSA LAGATEVDSGDGNHMFAGHLKCKVRMEKLRIK GMSYTMCSGKFSIDKEMAETQHGTTVVKVYE GTGAPCKVPIEIRDVNKEKVVGRIISSTPFAENTN SVTNIELEPPFGDSYIVIGVGDSALTLHWFRK
22	Zika virus (ZIKV)	IRCIQVSNRDFVEGMSGGTWVDVLEHGGCVT VMAQDKPTVDIELVTTTVSNMAEVRSYCYEASI SDMASDSRCPTQGEAYLDKQSDTQYVCKRTL DRGWGNGCGLFGKGSVLTCAKFACSKKMTGKS IQPENLEYRIMLSVHGSQHSGMIVNDTGHETDEN RAKVEITPNSPRAEATLGGFGLDCEPRTGLD FSDLYYLTMNSKHWLVHKEWFHDIPLPWHAGA DTGTPHWNNKEALVEFKDAHAKRQTVVVLGSQ EGAVHTALAGALEAEMDGAKGRLSSGHLKCRL KMDKLRLKGVSYSLCTAAFTFTKIPAETLHGT TVEVQYAGTDGPCKVPAQMAVDMQTLTPVGRL ITANPVITESTENSKMMLELDPPFGDSYIVIGVGE KKITHHWHR

23	West Nile virus (WNV)	FNCLGMSNRDFLEGVSGATWVDLVLEGDS CVTI MSKDKPTIDVKMMNMEAANLAEVRSYCYLATV SDLSTKAACPTMGEAHNDKRADPAFVCRQGVV DRGWGNGCGLFGKGSIDTCAKFACSTKAIGRTL KENIKYEVAIFVHGPTTVESHGNYSTQAGATQA GRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDT NAYYVMTVGTKTFLVHREWFMDLNL PWSSAGS TVWRNRETLMEFEPPHATKQSVIALGSQEGALH QALAGAIPVEFSSNTVKLTSGHLKCRVKMEKLQ LKGT TYGVCSKAFKFLGTPADTGHGTVVLELQY TGTDGPCKVPISSVASLNDLTPVGRLVTNPFVS VATANAKVLIELEPPFGDSYIVVGRGEQQINHH WHKS
24	Yellow fever virus (YFV)	AHCIGITDRDFIEGVHGGTWVSATLEQGKCVTV MAPDKPSLDISLQTVAIDGPAEARKVCYSAVLT HVKINDKCPSTGEAHLAEENDGDNACKRTYSDR GWGNGCGLFGKGSIVACAKFTCAKSMSLFEVD RTKIQYVIRAQLHVGAKQENWNTDIKTLKFDAL SGSQEAEFTGYGKATLECQVQTAVDFGNSYIAE MEKDSWIVDRQWAQDLTPWQSGSGGTWREM HHLVEFEPHAATIRVLALGNQEGSLKTALTGA MRVTKDENDNNLYKLHGHHVSCRVKLSALT GTSYKMCTDKMSFVKNPTDTGHGTVVMQVKV PKGAPCKIPVIVADDLTA AVNKGILVTNPIAST NDDEV LIEVNPFGDSYIIVGTGDSRLTYQWHKE

Chimeric flavivirus E polypeptides

5 With the exception of comprising two or more amino acid substitutions in the stem-transmembrane region, a flavivirus E protein of this disclosure may be a naturally-occurring (i.e., wild-type) E protein, or a mutant which has undergone further modification. One such modification contemplated in the present disclosure is the substitution of an E protein region from one flavivirus with a corresponding E protein region from another flavivirus or flavivirus strain to form a chimeric E protein. The substitution of particular E protein regions

(for example, the stem-transmembrane domain) may be advantageous in enhancing VLP secretion.

Thus, this disclosure also provides chimeric flavivirus E proteins with ectodomains and stem-transmembrane domains from different flaviviruses. The stem-transmembrane domain of the chimeric E protein may be modified to contain two or more amino acid substitutions as provided in this disclosure.

As used herein, a “chimeric flavivirus E polypeptide” or “chimeric flavivirus E protein” encompasses an E protein having amino acid sequences from two or more different flaviviruses, including different flavivirus strains. For example, chimeras can comprise an E protein in which an ectodomain, stem region and/or transmembrane domain (or regions therein) has been replaced with a corresponding domain or region of a second flavivirus. For example, the chimeras can consist of an E protein (e.g., a DENV or ZIKV E protein) in which the stem-transmembrane domain has been replaced with the stem-transmembrane domain of a second virus (e.g., JEV). The chimeric E proteins can be made from any combination of flavivirus E proteins.

Disclosed herein is a chimeric flavivirus E polypeptide comprising an ectodomain from a first flavivirus and a stem-transmembrane domain from a second flavivirus, wherein the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to position 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or 101 of SEQ ID NO: 1.

In some embodiments, the stem-transmembrane domain of the chimeric E polypeptide comprises two or more amino acid substitutions at positions corresponding to position 8, 11, 14, 15, 19, 26, 28, 35, 37, 39, 40, 49, 51, 52, 58, 59, 60, 61, 62 or 63 of SEQ ID NO: 1.

In some embodiments, the stem-transmembrane domain of the chimeric E polypeptide comprises two or more amino acid substitutions at positions corresponding to position 8, 11 or 15 of SEQ ID NO: 1.

In preferred embodiments, each of the two or more amino acid substitutions is a substitution to a non-polar amino acid residue. The E protein may further contain one or more additional mutations as defined herein, e.g., in the interface of the ectodomain and stem-transmembrane domain.

In some embodiments, the first and second flavivirus are different flaviviruses. The different flaviviruses may be different species, strains and/or serotypes.

10 In some embodiments, the ectodomain of the chimeric flavivirus E polypeptide is from dengue virus (DENV), Zika virus (ZIKV), West Nile virus (WNV), yellow fever virus (YFV) or Japanese encephalitis virus (JEV). In one embodiment, the ectodomain is from DENV or ZIKV. Exemplary ectodomain sequences are provided in Table 3.

15 In some embodiments, the stem-transmembrane domain of the chimeric flavivirus E polypeptide is from dengue virus (DENV), Zika virus (ZIKV), West Nile virus (WNV), yellow fever virus (YFV), or Japanese encephalitis virus (JEV). In one embodiment, the stem-transmembrane domain is from DENV or JEV.

20 In one embodiment, the chimeric flavivirus E polypeptide comprises an ectodomain from DENV or ZIKV and a stem-transmembrane domain from JEV.

Virus-like particles (VLPs)

25 Disclosed herein is a virus-like particle (VLP) comprising a flavivirus E polypeptide as defined herein.

In some embodiments, the VLP is a monovalent VLP comprising E proteins with an ectodomain from a single flavivirus strain or serotype. In other embodiments, the VLP is a multivalent VLP comprising E proteins with ectodomains from two or more different flavivirus strains or serotypes. For example, the VLP may comprise E proteins with ectodomains from two or more dengue virus serotypes.

The VLP may also comprise a flavivirus membrane (M) polypeptide. The term “flavivirus

membrane (M) polypeptide” or “flavivirus M protein” refers to the structural membrane protein found on the surface of mature virions. The M protein may be synthesised as a precursor protein (pre-M, also abbreviated prM), which is proteolytically cleaved by furin enzyme in a host cell to generate the mature M protein. The prM protein can help to stabilise the immature viral particle or VLP and prevent it from prematurely fusing with the host cell membrane. An M protein in the VLP is preferably from the same strain or serotype as an E protein present in the VLP.

In some embodiments, the VLP further comprises a targeting moiety (e.g., a peptide or protein domain such as a receptor binding domain) for targeting specific cells or tissues, e.g., immune cells or antigen presenting cells. The targeting moiety may increase VLP uptake into target cells or tissues. The targeting domains can be chemically attached to the surface of VLPs or may be recombinantly fused to a flavivirus polypeptide (e.g., E or M polypeptide) for display on the surface of the VLPs.

In some embodiments, the VLP further comprises a payload for delivery into a cell, wherein the payload is selected from a small molecule chemical compound, an antibody, a nucleic acid molecule, a polypeptide, or fragments thereof. The payload may be loaded into the VLP during VLP assembly, for example, by association of the payload with one or more components of the VLP. Alternatively, the payload may be loaded into the VLP after the VLP has been assembled, for example, by conjugating the payload to one or more components of the VLP. The payload may be delivered into a target cell by contacting the cell with a VLP containing the payload.

Polynucleotides and vectors

The present disclosure also provides isolated polynucleotides, expression constructs and vectors encoding flavivirus polypeptides as defined herein, in particular flavivirus E polypeptides with amino acid modifications in the stem-transmembrane domain. Also provided are host cells containing the polynucleotides or vectors.

Disclosed herein is an isolated polynucleotide comprising a nucleic acid sequence encoding a flavivirus E polypeptide as defined herein.

The isolated polynucleotide may also comprise a nucleic acid sequence encoding a flavivirus prM polypeptide.

The nucleic acid sequence encoding the E or prM polypeptides may each be operably linked to one or more control sequences. In one embodiment, the nucleic acid sequence encoding the E polypeptide and the nucleic acid sequence encoding the prM polypeptide are operably linked to the same control sequence. In another embodiment, the nucleic acid sequence encoding the E polypeptide and the nucleic acid sequence encoding the prM polypeptide are operably linked to different control sequences.

Disclosed herein is a vector comprising an isolated polynucleotide as defined herein. Isolated polynucleotides of this disclosure may be contained in more than one vector. For example, a polynucleotide encoding the E polypeptide may be contained in one vector and a polynucleotide encoding the prM polypeptide may be contained in a different vector.

Disclosed herein is a host cell comprising an isolated polynucleotide or vector as defined herein. The host cell may be any expression host (e.g., a prokaryotic or eukaryotic cell) capable of expressing proteins encoded in the polynucleotide or vector.

Polypeptide and VLP production

Also provided herein are methods for producing a modified flavivirus E polypeptide or a flavivirus VLP of this disclosure. The E protein modifications herein generally improve VLP secretion and enhance extracellular VLP titres.

In general, VLPs comprising one or more flavivirus E polypeptides of this disclosure may be produced by transformation of a suitable host cell with a modified E polypeptide-encoding polynucleotide in a suitable expression vector. The polynucleotide may also encode a flavivirus prM polypeptide for improved VLP stability during VLP assembly and secretion. E and prM polypeptides that are expressed can self-assemble into VLPs, and the VLPs may be collected from the cells and/or from the culture medium.

Disclosed herein is a method for producing a VLP, comprising culturing a host cell comprising an isolated polynucleotide or vector as defined herein, under conditions suitable

for expressing the encoded flavivirus E polypeptide to assemble the VLP.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the modified flavivirus E protein or VLP containing such. An isolated polypeptide of the invention may be produced in a prokaryotic host (e.g., *E. coli*) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae*, insect cells, e.g., Sf21 cells, or mammalian cells, e.g., NIH 3T3, HeLa, COS cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD). Non-limiting examples of insect cells are, *Spodoptera frugiperda* (Sf) cells, e.g., Sf9, Sf21, *Trichoplusia ni* cells, e.g., High Five cells, and *Drosophila* S2 cells. Examples of fungi (including yeast) host cells are *S. cerevisiae*, *Kluyveromyces lactis* (*K. lactis*), species of *Candida* including *C. albicans* and *C. glabrata*, *Aspergillus nidulans*, *Schizosaccharomyces pombe* (*S. pombe*), *Komagataella pastoris* (previously known as *Pichia pastoris*), and *Yarrowia lipolytica*. Examples of mammalian cells are COS cells, baby hamster kidney cells, mouse L cells, LNCaP cells, Chinese hamster ovary (CHO) cells, human embryonic kidney (HEK) cells, African green monkey cells, CV1 cells, HeLa cells, MDCK cells, Vero and Hep-2 cells. *Xenopus laevis* oocytes, or other cells of amphibian origin, may also be used. Prokaryotic host cells include bacterial cells, for example, *E. coli*, *B. subtilis*, and mycobacteria.

20

In one embodiment, the host cell is a mammalian cell. The mammalian cell is preferably a suspension cell line, non-limiting examples of which include the CHO-K1, Vero, PER.C6 and HEK-293 cell lines.

25 Depending on the expression system and host cell selected, the proteins or VLPs are produced by growing host cells transformed by a polynucleotide or an expression vector under conditions whereby the recombinant proteins are expressed and VLPs are formed. The selection of appropriate culture media and growth conditions is within the capacity of one of ordinary skill in the art. The protein or VLP can be produced in the cells and/or culture supernatant, and both may be used to recover the protein or VLP.

30

Methods to grow cells that produce VLPs of the invention include, but are not limited to, batch, batch-fed, continuous and perfusion cell culture techniques. Typically, cell culture is performed under sterile, controlled temperature and atmospheric conditions. A bioreactor is

a chamber used to culture cells in which environmental conditions such as temperature, atmosphere, agitation and/or pH can be monitored. The bioreactor can be a stainless steel chamber or a pre-sterilised plastic bag (e.g., Cellbag.RTM., Wave Biotech, Bridgewater, N.J.). The bioreactor may be dimensioned for cultures of about 20 L to about 50,000 L.

5

In some embodiments, the method further comprises recovering the VLP from the cell culture. E proteins and VLPs are isolated from the cell and/or media fraction using methods that preserve their integrity, such as by gradient centrifugation, e.g., caesium chloride, sucrose and iodixanol, as well as standard purification techniques including, e.g., ion

10

An exemplary method for the production and purification of dengue VLPs is provided in Galula et al. (*Bio Protoc* 2019 Jun 20;9(12):e3280. doi: 10.21769/BioProtoc.3280), which is incorporated by reference herein in its entirety. A person of skill in the art appreciates that

15

there are additional methods that can be used to make and purify VLPs. Accordingly, the invention is not limited to the methods described herein.

Immunogenic compositions and kits

20 The flavivirus E proteins and VLPs of this disclosure are useful for preparing vaccines and immunogenic compositions for inducing an immune response in a subject.

Disclosed herein is an immunogenic composition comprising a flavivirus E polypeptide, VLP, isolated polynucleotide or vector as defined herein.

25

The immunogenic composition may be mono- or multivalent. A multivalent immunogenic composition may comprise E proteins with different ectodomains, such as E proteins with ectodomains from one, two, three, four or more different flaviviruses, or alternatively it may comprise multivalent VLPs (i.e., VLPs containing E proteins with different ectodomains) or

30

a combination of monovalent VLPs. The immunogenic composition may be a formulation or mixture of isolated E polypeptides or VLPs. For example, the immunogenic composition may comprise a formulation or mixture of four isolated dengue VLPs (such as DENV1, DENV2, DENV3 and DENV4 VLPs).

The preparation of immunogenic compositions and vaccines is known to one skilled in the art. The immunogenic composition may comprise a pharmaceutically acceptable carrier or excipient. A pharmaceutically acceptable carrier or excipient according to the present invention can be any solvent or dispersing medium etc., commonly used in the formulation of pharmaceuticals and immunogenic compositions to enhance stability, sterility and deliverability of the active agent and which does not produce any secondary reaction, for example an allergic reaction, in a subject. The excipient is selected on the basis of the pharmaceutical form chosen, the method and the route of administration. Appropriate excipients, and requirements in relation to pharmaceutical formulation, are described in *Remington's Pharmaceutical Sciences* (19th Edition, A. R. Gennaro, Ed., Mack Publishing Co., Easton, Pa. (1995)). For example, E proteins may be encapsulated in liposomes for administration.

An immunogenic composition of the present invention may optionally contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, human serum albumin, essential amino acids, nonessential amino acids, L-arginine hydrochlorate, saccharose, D-trehalose dehydrate, sorbitol, tris-(hydroxymethyl) aminomethane and/or urea. In addition, the immunogenic composition may optionally comprise pharmaceutically acceptable additives including, for example, diluents, binders, stabilizers, and preservatives.

In some embodiments, the immunogenic composition of the present disclosure comprises one or more adjuvants. The adjuvant can enhance the immunogenicity of the polypeptide, VLP or polynucleotide of the immunogenic composition. The inclusion of any adjuvant described in Vogel et al., *A Compendium of Vaccine Adjuvants and Excipients* (2nd Edition), herein incorporated by reference in its entirety, is envisioned within the scope of this invention. Exemplary adjuvants include complete Freund's adjuvant (a non-specific stimulator of the immune response containing killed *Mycobacterium tuberculosis*), incomplete Freund's adjuvant and aluminum hydroxide. Other adjuvants may include GMCSF, BCG, muramyl dipeptide (MDP) compounds, such as thur-MDP and nor-MDP, muramyl tripeptide phosphatidylethanolamine (MTP-PE), lipid A, and monophosphoryl lipid A (MPL). RIBI, which contains three components extracted from bacteria, MPL,

trehalose dimycolate (TDM) and cell wall skeleton (CWS) in a 2% squalene/Tween-80 emulsion also is contemplated. MF-59, novasomes, MHC antigens may also be used. Those skilled in the art will be able to select an adjuvant which is appropriate for the active immunological substance.

5

As appreciated by skilled persons, an immunogenic composition of the present invention is suitably formulated to be compatible with the intended route of administration. Examples of suitable routes of administration include for instance intramuscular, transcutaneous, subcutaneous, intranasal, oral or intradermal. The immunogenic composition may be formulated as a sterile liquid for administration, or as a solid (such as a lyophilised powder or water-free concentrate) for reconstitution in a suitable solvent or dispersing medium.

10

The immunogenic compositions of the present disclosure may be administered using conventional hypodermic syringes or safety syringes such as those commercially available from Becton Dickinson Corporation (Franklin Lakes, N.J., USA) or jet injectors. For intradermal administration, conventional hypodermic syringes may be employed using the Mantoux technique or specialized intradermal delivery devices such as the BD Soluvia™ microinjection system (Becton Dickinson Corporation, Franklin Lakes, N.J., USA), may be used.

20

The volume of an immunogenic composition of the present invention administered will depend on the method of administration. In the case of subcutaneous injections, the volume is generally between 0.1 and 1.0 ml, preferably approximately 0.5 ml.

25

This disclosure also provides for pharmaceutical kits comprising one or more components of the immunogenic compositions or vaccine formulations defined herein. The kit may comprise a flavivirus E polypeptide, VLP or polynucleotide as defined herein in a first container, and may optionally contain a suitable dispersing medium and/or an adjuvant in separate containers.

30

Methods of treatment and prevention

The modified flavivirus E proteins of this disclosure may be used to stimulate an immunological response in a subject, such as in a human or an animal. The immune response

can protect the subject from a flavivirus infection, such as a DENV, ZIKV, WNV, YFV and/or JEV infection. The E proteins may be used either therapeutically in subjects already experiencing a flavivirus infection, or may be used prophylactically to prevent a flavivirus infection.

5

Disclosed herein is a flavivirus E polypeptide, VLP, isolated polynucleotide, vector, or immunogenic composition as defined herein, for use as a medicament or vaccine.

Disclosed herein is a method of modulating an immune response in a subject, the method comprising administering a therapeutically effective amount of an immunogenic composition as defined herein to the subject.

10

Disclosed herein is a method of preventing or treating a flaviviral infection in a subject, the method comprising administering to the subject a therapeutically effective amount of an immunogenic composition as defined herein.

15

Disclosed herein is an immunogenic composition as defined herein, for use in preventing or treating a flaviviral infection in a subject.

Disclosed herein is the use of an immunogenic composition in the manufacture of a medicament for preventing or treating a flaviviral infection in a subject.

20

Methods herein may be used to prevent or treat an infection caused by any flavivirus, or any disease or disorder arising from a flaviviral infection, preferably infections and diseases caused by arthropod-borne flaviviruses, e.g., mosquito- or tick-borne flaviviruses. Non-limiting examples of flaviviral infections and associated diseases include dengue fever, Zika fever, Japanese encephalitis, yellow fever, West Nile fever, Kunjin infection, Murray Valley encephalitis, tick-borne encephalitis, Siberia fever, louping-ill disease, Kyasanur Forest disease, Alkurma haemorrhagic fever, Omsk haemorrhagic fever and Powassan encephalitis.

25

30

In some embodiments, the immunogenic composition is administered in two or more doses. The method may comprise administering a first or prime dose of an immunogenic composition of this disclosure, followed by one or more boost doses of the same immunogenic composition after a suitable period of time. The first of the one or more boost

doses may be administered, for example, between one week and ten years after the prime dose, for example, about 1 week, about 2 weeks, about 3 weeks, about 1 month, about 2 months, about 3 months, about 4 months, about 5 months, about 6 months, about 7 months, about 8 months, about 9 months, about 10 months, about 11 months, about 12 months, about 13 months, about 14 months, about 15 months, about 16 months, about 17 months, about 18 months, about 19 months, about 20 months, about 21 months, about 22 months, about 23 months, about two years, about three years, about four years, about five years, about six years, about seven years, about eight years, about nine years, or about ten years after the prime dose. As a non-limiting example, a first boost dose can be administered about 1 month after the prime dose, and about six months after the first boost dose, a second boost dose can be administered, about 12 months after the second boost dose, a third boost dose can be administered. Additional boost doses can be periodically administered; for example, every 5 years, every 10 years, etc. The timing between the prime and boost doses can vary, for example, depending on the stage of infection or disease (e.g., non-infected, infected, number of days post-infection), and the subject's health. The dosage amount can also vary between the prime and boost dosages. As a non-limiting example, the prime dose can contain fewer copies of the active immunological substance compared to the boost dose.

As used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (or).

As used in this application, the singular form "a," "an," and "the" include plural references unless the context clearly dictates otherwise. For example, the term "an agent" includes a plurality of agents, including mixtures thereof.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

Throughout this specification and the claims which follow, unless the context requires otherwise, the phrase "consisting essentially of", and variations such as "consists essentially of" will be understood to indicate that the recited element(s) is/are essential i.e. necessary

elements of the invention. The phrase allows for the presence of other non-recited elements which do not materially affect the characteristics of the invention but excludes additional unspecified elements which would affect the basic and novel characteristics of the method defined.

5

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour

10 to which this specification relates.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications, which fall within the spirit and scope. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

15

Unless otherwise defined, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

20

Certain embodiments of the invention will now be described with reference to the following examples which are intended for the purpose of illustration only and are not intended to limit the scope of the generality hereinbefore described.

25

EXAMPLES

Materials and methods

30

Virus-like particles (VLPs) and mutants

Three different previously constructed and characterized recombinant DNA plasmids expressing prM-E of flavivirus were used to generate VLPs from COS-1 cells (monkey fibroblast kidney tissue, ATCC #CRL-1650), including D2VLP (mD2VLP and imD2VLP),

JEV and ZIKV. D2VLP-expressing plasmid pVD2 encodes a chimeric envelope protein containing a portion of the envelope protein of DENV-2 (Asian onc genotype, strain 16681) at the N-terminus (80%) and a portion of the envelope protein of JEV (Nakayama strain) at the C-terminus (20%), and a prM protein with a modified furin cleavage site to generate mature and immature VLPs. The amino acid sequence of the chimeric DENV2-JEV E protein with wild-type JEV sequence is given in SEQ ID NO: 17. Additionally, the plasmid was used for site-directed mutagenesis following the manufacturer's protocol (Stratagene, La Jolla, CA). The primers used for site-directed mutagenesis are provided in Table 2. Nucleotide sequencing confirmed that all the plasmids contained no other mutations other than those indicated.

DENV2-JEV E protein (JEV stem-transmembrane sequence underlined)

MRCIGMSNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEAKQPAT
LRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFVCKHSMVDRGWNGCGGLFGK
GGIVTCAMFRCKKNMEGKVVPENLEYTIVITPHSGEEHAVGNDTGKHGKEIKITP
QSSITEAELTGYGTVTMECSPTGLDFNEMVLLQMENKAWLVHRQWFLDLPLPW
LPGADTQGSNWIQKETLVTFKNPHAKKQDVVVLGSQEGAMHTALTGATEIQMSS
GNLLFTGHLKRLRMDKLQLKGMSYSMTGKFKVVKEIAETQHGTIVIRVQYEG
DGSPCKIPFEIMDLEKRHVLGRLITVNPVTEKDSPVNIEAEPFPGDSYIIIGVEPGQL
KLNWFKKGSTLGKAFSTTLKGAQRLAALGDTAWDFGSIGGVFNSIGKAVHQVFG
GAFRTLFGGMSWITQGLMGALLWMGVNARDRSIALAFLATGGVLVFLATNVHA
(SEQ ID NO: 25)

Equal quantity (4 µg) of plasmids expressing wild-type or mutant D2VLPs were also transfected into HEK-293T seeded into 24-well plate (human embryonic kidney, ATCC#CRL-3216), using branched polyethylenimine (Sigma-Aldrich Inc., St. Louis, MO, USA) with the PEI: DNA ratio of 1:1. The culture supernatant after transfection was harvested on a daily base for ELISA up to day 5. At 5 days post-transfection, the cells were fixed with 70% acetone for immunofluorescent assay (IFA).

Table 4. Primers for site-directed mutagenesis of mD2VLP.

Amino acid mutant	Forward primer	Reverse primer
L398A	5'- AAGAAAGGAAGCAC <u>GGC</u> GGGCAAGGCCTTTTC-3' (SEQ ID NO: 26)	5'- GAAAAGGCCTTGCCC <u>GCC</u> GTGC TTCCTTTCTT-3' (SEQ ID NO: 27)
L398E	5'- AAGAAAGGAAGCAC <u>GGA</u> GGGCAAGGCCTTTTC-3' (SEQ ID NO: 28)	5'- GAAAAGGCCTTGCCC <u>TCC</u> GTGCT TCCTTTCTT-3' (SEQ ID NO: 29)
F402A	5'- ACGCTGGGCAAGGCC <u>CGC</u> TTCAACAACCTTTGAAG-3' (SEQ ID NO: 30)	5'- CTTCAAAGTTGTTGAAGcGGCCTT GCCCAGCGT-3' (SEQ ID NO: 31)
F402E	5'- ACGCTGGGCAAGGCC <u>GAA</u> <u>A</u> TTCAACAACCTTTGAAG-3' (SEQ ID NO: 32)	5'- CTTCAAAGTTGTTGAT <u>TTC</u> GGCCT TGCCCAGCGT-3' (SEQ ID NO: 33)
T405A	5'- AAGGCCTTTTCAAC <u>AGCT</u> TTGAAGGGAGGTAAG-3' (SEQ ID NO: 34)	5'- CTTACCTCCCTTCAA <u>AGCT</u> GTTG AAAAGGCCTTG-3' (SEQ ID NO: 35)
T405E	5'- AAGGCCTTTTCAAC <u>GAA</u> TTGAAGGGAGGTAAG-3' (SEQ ID NO: 36)	5'- CTTACCTCCCTTCAA <u>TTC</u> GTTG AAAAGGCCTTG-3' (SEQ ID NO: 37)
A409E	5'- CCTTTCTCTCCACAG <u>AAC</u> AAAGACTGGCAGCG-3' (SEQ ID NO: 38)	5'- CGCTGCCAGTCTTTG <u>TTC</u> TGTGG AGAGAAAGG-3' (SEQ ID NO: 39)

L412A	5'- CTCCACAGCTCAAAGAG CGGCAGCGTTGGGCGAC- 3' (SEQ ID NO: 40)	5'- GTCGCCCCAACGCTGCCGCTCTTT GAGCTGTGGAG-3' (SEQ ID NO: 41)
L412E	5'- CTCCACAGCTCAAAGAG AGGCAGCGTTGGGCGAC- 3' (SEQ ID NO: 42)	5'- GTCGCCCCAACGCTGCCCTCTCTTT GAGCTGTGGAG-3' (SEQ ID NO: 43)
L415A	5'- CAAAGACTGGCAGCGGC GGGCGACACAGCCTG-3' (SEQ ID NO: 44)	5'- CAGGCTGTGTCGCCCGCGCTGC CAGTCTTTG-3' (SEQ ID NO: 45)
L415E	5'- CAAAGACTGGCAGCGGA GGGCGACACAGCCTG-3' (SEQ ID NO: 46)	5'- CAGGCTGTGTCGCCCTCCGCTGC CAGTCTTTG-3' (SEQ ID NO: 47)

VLP secretion level determined by enzyme-linked immunosorbent assay (ELISA)

ELISA was performed on VLP-containing supernatant harvested from the plasmid-transfected cells to determine secretion levels. In short, 96-well ELISA plates (Thermo, #442404) were coated with in-house prepared rabbit sera against D2-VLP, ZIKV-VLP or JEV-VLP overnight at 4°C. 50 uL of each transfected supernatant were added to each well and incubated at 37°C for one hour. The pan-flavivirus anti-E monoclonal antibody FL0231 and secondary goat anti-mouse-IgG-HRP antibody were added sequentially for detection.

10 Immunofluorescent assay (IFA)

HEK-293T cells were transfected with plasmids expressing D2VLP wild-type or mutants to evaluate intracellular protein expression. Five days post-transfection, the cells were fixed by 70% acetone at room temperature for 10 minutes. Cells were then blocked with 3% ovine serum albumin (BSA) at 37°C for 1 hour. The same pan-flavivirus anti-E monoclonal antibody FL0231 was added to the wells and incubated at 4°C overnight. The cells were washed three times and incubated with a mixture of FITC-conjugated goat anti-mouse IgG

(1:200; Jackson ImmunoResearch, West Grove, PA) and DAPI (4',6-diamidino-2-phenylindole dihydrochloride; 1:500; Invitrogen, Molecular Probes, Inc., Eugene, OR) for 1 h at 37°C. After washing 5 times, cells were visualised under a fluorescence microscope (CKX41; Olympus).

5

Optimisation of lipid extraction methods

Three methods were used for lipid extraction of mD2VLP: chloroform extraction, alcohol extraction, and methyl tertiary butyl ether (MTBE) extraction. These lipid extraction methods were previously used to explore the lipid content in virus particles or VLPs. The chloroform lipid extraction method is commonly used on virus particles and is a modified form of the Bligh and Dyer method of Folch lipid extraction. Considering that the lipid composition of mD2VLP has not been previously examined, all three approaches were applied to the same batch of purified mD2VLP and yields from each method were compared. 10 nL of purified mD2VLP from the HEK-293T transfected group (mD2VLP) were subjected to lipid extraction: (1) Chloroform extraction: 800 µL 0.1N HCl:CH₃OH (1:1) were added to the sample, followed by adding 400 µL of chloroform and shaking at 4°C for 1 minute. The mixture was centrifuged at 16,000xg for 7 minutes to separate the organic from the water phase. The lipid-containing organic layer was collected and air-dried; (2) EtOH extraction: 900 µL of acidic ethanol (0.18 M HCl:ethanol = 1:3) was added to the sample and subjected to low-temperature ultrasonication for 3 minutes. The supernatant was collected after the centrifugation at 14,000xg and 4°C and dried under vacuum; (3) Methyl tertiary butyl ether (MTBE) extraction: 225 µL of methanol at 4°C was added to the sample and vortexed for 10 seconds. Then 750 µL MTBE at 4°C was added to the sample and shaken for 6 minutes. The sample was centrifuged at 14,000xg for 2 minutes. An additional 188 µL of double-distilled water was added and the mixture shaken for 20 seconds. The upper organic layer was collected and dried under vacuum. All dried samples were stored at -20°C and re-dissolved in 0.1% formic acid buffer before for LC/MS/MS analysis. The mock-transfection group (293T) were also subjected to lipid extraction as the negative control.

30 LC-TOF mass spectrometry and lipidomic analysis

Liquid chromatographic time of flight (TOF) mass spectrometry was performed on a Waters ACQUITY UPLC I-Class system with Xevo G2XS ToF mass spectrometry (MS) operated in positive or negative ion mode. The LC system consists of a Waters Acquity UPLC system. Lipid molecules were separated on an Acquity UPLC BEH Amide column (150 × 2.1 mm;

1.7 μm) (Waters, Milford, MA, USA). The column was maintained at 45°C at a flow rate of 0.4 mL/min. The mobile phases consisted of (A) water with 10 mM ammonium formate; 0.125% formic acid and (B) acetonitrile:water (95:5, v:v) with 10mM ammonium formate 0.125% formic acid. A sample volume of 0.5–3 μL was used for the injection. The separation
5 was conducted under the following gradient: 0 min 100% (B); 2 min 100% (B); 7.7 min 70% (B); 9.5 min 40% (B); 10.3 min 30% (B); 12.8 min 100% (B); 17 min 100% (B). The sample temperature was maintained at 10°C. The MS scan range varied from 50 to 1200 m/z. The desolvation gas was set to 800 l/h at a temperature of 500°C, and the source temperature was set at 120°C. The capillary voltage and cone voltage were set to +2.5/-2 kV and 25 V,
10 respectively. Data-dependent analysis (DDA) was performed.

Subsequently, the extracted lipids were subjected to LC/MS/MS for lipidomics analysis. It was found that the group extracted with chloroform has the larger sum of the total 2 spectra obtained in the positive ion scanning mode. The negative ion scanning mode was also
15 compared with the group extracted by the other two extraction methods (Figure 12). The lipid extraction approach using chloroform was the most suitable lipid extraction method.

Molecular dynamics simulations

Parameters for molecular dynamics simulations are as provided in Tables 5 and 6.

Table 5. Simulation system setup of all atom molecular dynamics simulations. Simulations were run for 500 ns using CHARMM36m force field with 120mM ion concentration in a ~16.6×16.6×12.2 nm³ box size with ~66,000 TIP3P water molecules.

	E protein mutation	Lipid membrane composition
DENV (E _{chr-M} protein)	WT	POPC:POPE:POPS (60:30:10)
	L398A	
	L398E	
	F402A	
	F402E	
	T405A	
	T405E	
	A409E	
	L412A	
	L412E	
DENV (E _{chr-M} protein)	WT	DAG:FA:SAPC (56:26:8)
	L398A	
	L398E	
	F402A	
	F402E	
	T405A	
	T405E	
	A409E	
	L412A	
	L412E	
DENV	WT	POPC:POPE:POPS (60:30:10)

5

Table 6. Simulation system setup of coarse-grained molecular dynamics (CG-MD) simulations.

	E protein mutation	Box Size (nm)	No. of water molecules	Lipid membrane composition	NaCl concentration	Production run (ns)		
DENV (E _{chr-M} protein)	VLP-1	40.5×40.5×40.5	507491	POPC:POPE:POPS (60:30:10)	120mM	3 × 2000 ns		
	VLP-2	32.5×32.5×32.5	237556					
	VLP-3	32.5×32.5×32.5	239474					
DENV (E _{chr-M} protein)	VLP-1	32.5×32.5×32.5	235532	DAG:SAPC:FA (56:26:8)				
	VLP-2	32.5×32.5×32.5	237188					
	VLP-3	32.5×32.5×32.5	238864					

10 **Simulation system setup**

E protein dimers from dengue (gene ID: KU725663.1) and Chimeric dengue [(80% DENV (residues 1–394) + 20% JEV (residues 395–495))] were modelled using Modeller version 9.1024 using a PDB templates 3J2726 and 5WSN, respectively. Out of 100 models generated for each E protein dimer, the lowest energy structure was chosen to perform MD simulations.

The native and chimeric E protein along with DENV M protein dimeric structures were assigned with charges according to neutral pH using CHARMM36m force field. The proteins were embedded into two different lipid membrane systems composed of either POPC:POPE:POPS or DAG:POPC:FA [POPC: Phosphatidylcholine (16:0/18:1), POPE: Phosphatidylethanolamine (16:0/18:1), POPS: Phosphatidylserine (16:0/18:1), DAG: Diacylglycerol (18:0/20:4), SAPC: Phosphatidylcholine (18:0/20:4), FA: Fatty acid (18:1)].

Lipid bilayers were built using CHARMM-GUI Membrane builder. The phospholipid composition of 60:30:10 was used and for DAG:SAPC:FA a ratio of 56:26:8 used as per lipidomic data from the current study. Orientation of E-M protein dimers with respect to the lipid bilayers was predicted using orientation of membrane protein (OMP) reference server. E protein-lipid bilayer systems were solvated in TIP3P33 water box and 150 mM NaCl (Table 5) on top of neutralising the overall system charge. Mutations of transmembrane region (residues 450–495) to JEV as well as point mutations on the chimera were done using CHARMM_GUI PDB Reader & Manipulator tool. Each E-M protein-lipid membrane system was subjected to 5000 steps of steepest descent minimisation. The final converged structures from energy minimisation step were subjected to multiple steps of equilibration consisting of position restraints, two NVT and four NPT equilibrations with gradually decreasing force constants (4000 to 0 kJ/mol/nm²) of protein backbone and lipid head groups (2000 to 0 kJ/mol/nm²). A total of 1.5 ns of equilibration simulation was performed in NVT with a 1 fs time step, while NPT equilibrations steps used 2 fs time step using leap-frog algorithm. Semi-isotropic Berendsen pressure coupling of 1 bar was used during equilibrations, while temperature was controlled using Berendsen thermostat. Each system production run was forest to 500 ns in GROMACS 2018. Particle mesh Ewald (PME) method was used for long range electrostatic interaction with a cut-off distance of 1.2 nm and switching function after 1 nm. Production runs included Nose-Hoover thermostat (303K) and semi-isotropic Parrinello-Rahman barostat (1bar) with 1 ps and 5 ps coupling respectively. Trajectory visualisation and analysis such RMSD, solvent accessibility, principal component analysis (essential dynamics), secondary structure content, etc. was performed used VMD version 1.9.3 and gromacs tools.

Rosetta membrane protein (RosettaMP) module to predict $\Delta\Delta G$ of chimeric E protein (E_{chi}) mutants

E_{chi}-M dimer protein structures oriented as per orientation of proteins in membrane (OPM) databases was used. The input pdb file was cleaned and remembered using clean_pdb.py (from ROSETTA scripts) and span files were generated using span_from_pdb application in RosettaMP module. To access the effect of point mutations on stem helix-1 of E protein

5 MPddG module was used. MPddG used RosettaMP framework with fixed backbone prediction protocol and franklin2019 scoring function. Also, side chain conformations were sampled for residues within 8Å of mutant position and the ΔΔG was calculated using Rosetta Energy Units (REU) between mutant and native structures. The mutants showing favourable ΔΔG values (i.e., negative) were used to perform molecular dynamics as described above

10 and experimentally tested for secretion efficiency.

Coarse-grained modelling and simulation setup

ChimeraX was used to build atomistic model of the full VLP-based cryo-EM density map (EMDB:6926). VLP structure was consists of 60 E/M protein dimers embedded in lipid

15 vesicle as described previously. All atom coordinates of VLP were converted to MARTINI3 coarse-grained (CG) model. The CG coordinates were embedded into the lipid vesicle with the radius of ~90Å. Lipid vesicle was generated using CHARMM-GUI Martini maker. Two different lipid composition were used: i) POPC:POPE:POPS (60:30:10) or ii) DAG:SAPC:FA (56:26:8) lipid membranes respectively [(POPC: Phosphatidylcholine

20 (16:0/18:1), POPE: Phosphatidylethanolamine (16:0/18:1), POPS: Phosphatidylserine (16:0/18:1), DAG: Diacylglycerol (18:0/20:4), SAPC: Phosphatidylcholine (18:0/20:4), FA: Fatty acid (18:1)]. After embedding E/M proteins into the lipid vesicle, lipid molecules overlapping with protein CG molecules were removed. Multiple initial CG-VLP models were generated with different number of lipid molecules. Each VLP model was centred in a

25 cubic box and solvated with MARTINI water and NaCl beads corresponding to 150mM salt concentration. All systems were subjected to energy minimisation of 50,000 steps using steepest descent algorithm. All simulations were performed using the GROMACS 2018 version.

30 To retain native folded structure of E protein, elastic network (EN) was introduced within each E and M protein monomer between protein backbone beads. EN was used for beads with a distance ranging from 0.5 to 0.9 nm and a force constant of 1000 kJ/mol/nm². The elastic network was used only within each of the DI, DII and DIII of the ectodomain while E and M transmembrane regions maintained their secondary structure based on angles and

dihedrals alone. In all the simulations bonds lengths were constrained using LINCS algorithm and equations of motion were integrated using leapfrog algorithm and a 10fs time step. Pressure and temperature were maintained at 1 bar (isotropic) and 310K using velocity rescaling thermostat and Berendsen barostat respectively. Six sets of equilibration steps in NVT and NPT ensembles were performed with lowering restraint on protein back bone beads, protein side chain beads and lipid beads and progressively increasing the time step. The final NPT equilibration step involved position restraints on protein beads for 500 ns to allow equilibration of lipids. The unrestrained production run of 2000 ns in three independent replicas was run for each VLP CG systems in both lipid compositions (Table 5).

EXAMPLE 1: E-H1 helix interactions with lipid bilayer govern VLP secretion efficiency

To investigate the lipid composition of flavivirus VLPs produced in different mammalian cell lines, three different purified VLPs were used: i) dengue mD2VLP prepared from both HEK-293T and COS-1 cells, ii) Japanese encephalitis VLP (JEVLP) prepared from the COS-1 cells and iii) Zika virus VLP (ZIKAVLP) prepared from the COS-1 cells. These were subjected to lipid extraction followed by LC/MS/MS for lipidomic data after confirmation of particles formation by electron microscope (EM) (Figure 1). The purpose of choosing various flaviviruses was to assess whether various flavivirus VLPs produced from different mammalian cells have different lipid compositions. The LC-MS/MS analysis showed similar lipid compositions across different flavivirus VLPs from the same cell lines or the same VLPs from different cell lines. The two most abundant lipids extracted from purified VLPs were diacylglycerols (DG) of the glycerolipids category and fatty acid (FA) of the fatty acyls category. These are mainly unsaturated fatty acid chains with oxidation at C34 and C18, respectively. Phosphatidylcholine (PC) of the glycerophospholipids category, oxidised fatty acid (OxFA) of the fatty acyls category or other lipids like ceramide alpha-hydroxy fatty acid-phytosphingosine (cerAP) of the sphingolipids category have relatively low abundance (Table 5).

Additionally, to confirm if VLPs of different maturities have different lipid compositions, immature VLP of DENV-2 (imD2VLP) prepared from CHO-K1 stable cell line were also subjected to lipid extraction by chloroform and composition analysis by LC/MS/MS. Similar

lipid composition of VLPs is DG, FA, PC, and OxFA was found (Table 5). By summing up all the m/z spectra in positive or negative lipids, no differences on the total amount of lipids were observed among different types of VLP (Figure 2).

5 **Table 7.** Lipid composition of flavivirus VLPs produced from different host cells. DG: diacylglycerol; FA: fatty acid, PC: phosphatidylcholine, and OxFA: oxidised fatty acid.

mD2VLP(HEK293T)					mE2VLP(COS-1)				
major lipids	DG	FA	PC	OxFA	major lipids	DG	FA	PC	OxFA
major carbon-number	34	18	34	18	major carbon-number	34	18	30	18
lipids ratio	18	8	3	1	lipids ratio	18	9	3	1
Saturated fat acid	24%	0%	0%	15%	Saturated fat acid	28%	0%	52%	16%
Unsaturated fatty acid	76%	100%	100%	85%	Unsaturated fatty acid	72%	100%	48%	84%

JEVLP(COS-1)					ZIKAVLP(COS-1)				
major lipids	DG	FA	OxFA	PC	major lipids	DG	FA	OxFA	PC
major carbon-number	34	18	18	34	major carbon-number	34	18	18	34
lipids ratio	64	27	4	1	lipids ratio	31	15	2	1
Saturated fat acid	26%	0%	18%	2%	Saturated fat acid	25%	0%	15%	0%
Unsaturated fatty acid	74%	100%	84%	98%	Unsaturated fatty acid	75%	100%	85%	100%

Previous studies have shown that secretion of DENV-2 and SLEV VLPs from mammalian cells is greatly improved if the VLPs contained chimeric E proteins in which the C-terminal 20% of the native gene is replaced with the corresponding region from JEV. This C-terminal region contains the E-H1, E-H2 and E-H3 helices which interact with the lipid bilayer of the VLP. By examining the arrangement of residues in the C-terminal helices (Fig. 3), it is hypothesised that the L398, F402, T405, A409, L412 and L415 residues may interact with the head groups of the lipid bilayer. Site-directed mutagenesis (SDM) was performed to substitute these residues with an alanine (A) residue (to reduce hydrophobicity) or glutamate (E) residue (for charge inversion). pVD2 plasmid encoding the E protein mutants were mixed with PEI at a ratio of 1:30 and transfected into HEK-293T cells. To determine VLP assembly and secretion, E protein ELISA was performed on culture supernatant, and immunofluorescence was performed on the cells.

The ELISA results suggested that the mutations had variable effects on VLP secretion. However, mutations at A409 and L412 had the greatest effect on VLP secretion, with positive/negative (P/N) ratio close to 1, indicating minimal secretion (Figure 4). Immunofluorescent staining confirmed the intracellular accumulation of the A409E, L412A, and L412E E protein mutants (Figure 5). Thus, amino acid residues at critical sites within the stem helix-1 (E-H1) may have a stabilising effect on VLP assembly through their interactions with the lipid bilayer.

To gain mechanistic insights into the effect of the Glu and Ala mutations at the L398, F402, T405, A409 and L412 sites within E-H1, all-atom molecular dynamics (MD) simulation was performed on dimeric DENV with wild-type DENV-2 C-terminal (E-M), JEV C-terminal (E_{chi}-M) and JEV C-terminal containing point substitutions (EH1-mutants) embedded in lipid membrane with two different compositions: POPC:POPE:POPS based on the virion lipid composition, and DAG:POPC:FA based on the lipid composition in this study (Table 5). Each system of lipid and E-M protein dimer was simulated for 500 ns to investigate the effect of site-specific mutations on lipid-protein interactions (Figs. 6 and 7).

10

The simulations revealed that the E_{chi}-M protein dimers had a higher number of lipid contacts compared to the E-H1 Ala and Glu mutants (Figs. 7A) for both lipid membrane systems. E-H1 helix insertion depth measurements revealed that the E-H1 helix of E_{chi} is inserted more deeply into the lipid membrane than other E-H1 mutants (Fig. 7C and Fig. 12) in both lipid membrane systems. Also, insertion depth plots with mean E-H1 helix insertion highlights the average orientation of the E-H1 helix with respect to lipid membrane. Interestingly, it was observed that E-H1 helices from mutants F402A/E, T405A/E, A409E and L412A/E point away from the membrane, indicating reduced interactions with lipids. The mutations on E-H1 sites show similar trend in POPC dominant lipid E-M dimer simulations. Along with E-H1 helix lipid interactions, secondary structural changes in E-H1 helix were also tracked. Interestingly, it was observed that the E-H1 helix from E_{chi}-M dimer embedded in both DAG- and POPC- dominant lipid systems underwent local unfolding within the 500 ns trajectory time. This local unfolding of the helix could potentially destabilise the E-M protein dimers in fully assembled VLPs. In particular, F402E, A409E and L412E mutants showed a high degree of unfolding in both the E protein monomers (Fig. 13). Unlike Glu mutants, Ala substitution in E-H1 helix showed minimal effect on the local secondary structure of the E-H1 helix (Figure 14).

The interactions between native dengue E protein dimers or E_{chi} protein dimers with DAG- or POPC-dominant lipid membrane systems were further compared. In general, it was observed that the measured parameters i.e., E-H1-lipid contacts, E-H1 insertion depth and secondary structure of E_{ch}, exhibited minor differences (Fig. 8). However, lipid-E-H1 helix interaction and E-H1 helix insertion into the lipid membrane were consistently higher for the E_{chi}-M dimeric system compared to native E protein in both the lipid systems (Figure 8B

30

and 8C). Also, RMSF of stem helix was observed to be higher in the case of E_{chi} (Fig. 14). Therefore, increased dynamics and LEU mutations at critical sites L398 and L412 possibly contribute to higher number of interactions with the lipid membrane and better insertion of the E-H1 helix into the membrane.

5

EXAMPLE 2: E-H1 helix engineering enhances VLP secretion

MD simulations and secretion assay highlighted the presence of particular residues in the E-H1 helix which are critical for E-H1-lipid interactions. To enhance VLP stability and secretion efficiency, the E-H1 helix was engineered to enhance its interaction with the lipid envelope without drastically perturbing E protein dimer formation. The RosettaMP module was utilised to pinpoint amino acids that are likely to enhance E-H1 helix-lipid interactions. RosettaMP predicted that hydrophobic residues (especially Leu) showed negative free energy values at all the sites of interest i.e., 402, 405, 409 and 412 (Fig. 9A, B). Site-directed mutagenesis was then performed to introduce single, double, and triple Leu substitutions at the identified sites in the E-H1 helix. Interestingly, di- and tri-Leu substituted VLPs showed 2-fold higher secretion compared to unmodified chimeric VLPs and >150-fold high secretion compared to native DENV VLPs. Furthermore, the MD simulations showed that di- and tri-Leu substitutions induced increased helicity of the E-H1 helix and more helix-lipid interactions. Also, the Leu-substituted E-H1 helix showed greater depth of insertion into the lipid envelope (Fig. 9). Collectively, MD simulations and the *in vitro* assays clearly showed that the interaction between the lipid membrane and the H1 helix of the E protein are critical for VLP assembly.

25 Additional amino acid residues at positions within the stem-helix-transmembrane region (i.e., residues spanning 395–495) were also mutated. Over the course of 500 ns long MD simulations of each mutant, increased insertion depth and lipid interactions were observed (Fig. 17), indicating enhanced interaction of the E protein with the lipid membrane.

30 Fig. 18 is a heat map of free energy changes upon mutating each amino acid residue in the stem-transmembrane region of JEV (residues 395 to 495) to any one the 19 amino acids (except proline). More negative values (black or darker shades) correspond to favourable mutations which enhance protein stability and VLP secretion, whereas more positive values (white or lighter shades) correspond to unfavourable mutations which decrease protein

stability and diminish secretion. It can be seen that mutations to non-polar residues along the stem-transmembrane region, especially mutations to F, I or L, can enhance VLP secretion as these mutations have negative free energy difference values ($\Delta\Delta G$) as shown in Fig. 18.

5 EXAMPLE 3: Coarse-grained (CG) molecular dynamics simulations of DENV VLPs

Two different lipid compositions were used to model POPC:POPE:POPS (60:30:10) and DAG:SAPC:FA (56:26:8) lipid vesicles with varying lipid number (Fig. 11). Each VLP model was subjected to 500 ns long equilibration simulation with restrained protein and lipid
10 beads were allowed to freely rearrange to adopt a stable configuration around the E-M protein shell. For POPC-dominant VLPs, the simulation showed that lipid molecules were uniformly distributed under the protein shell, whereas DAG-dominant VLPs formed lipid bulges as shown in head group density distribution along coordinate axes (Fig. 16A-16B). Also, lipid vesicle-only simulations were performed with lipid composition identical to VLP
15 models and similar behaviour was observed to that of lipid vesicle in VLP models, i.e., POPC-dominant vesicles had a spherical shape while DAG-dominant vesicles showed lipid bulge formation (Fig. 16C-16D).

The last frame from each VLP model equilibration run was used to perform unrestrained 2
20 μ s-long simulation and each VLP model system was run in triplicates. Interestingly, POPC-dominant and DAG-dominant lipid systems had distinct effects on VLP spherical morphology during unrestrained production runs (Fig. 10). The sphericity of the VLPs in the case of POPC-dominant CG-models showed dependency on the lipid number (Fig. 10B, C). As shown in Fig. 10B, the CG-model VLP-2 and 3 showed a stable spherical morphology
25 throughout the simulation time in POPC-dominant VLPs. VLP-1 (POPC-dominant lipid system) CG-model with highest number of lipid molecules adopted an ellipsoid shape which have been reported to be present in cryo-EM micrographs but not considered for image reconstruction. This suggested that in the VLP or native virus assembly process, the incorporation of lipid molecules is random, i.e., the number of lipid molecules in a viral or
30 VLP envelope is not precisely regulated. This explains the significant number of non-spherical particles observed in negative stain images or cryo-EM electron micrographs of VLP or native viral particles.

Among the DAG-dominant VLP CG-model trajectories, it was observed that the bilayer structure was lost resulting in the formation of bulges in the lipid vesicle below the E_{chl}-M protein dimers (Fig. 10C). The lipid bulge formed during equilibration run further expanded to cover a large part of the interior of the VLPs in the DAG-dominant VLP systems. Previous studies have reported a similar trans-bilayer activity of DAG leading to a lipid bulge or lipid lens formation. DAGs were previously shown to be critical in forming lipid droplets (LD). As viral assembly and LD generation occurs at the ER membrane the VLPs incorporate DAGs into the vesicle in a host cell-specific manner. Lipidomic analysis in this work showed that the DAG is the dominant lipid species in VLPs among the cell lines tested. The asymmetric lipid bulges occurring in one section of the VLPs results in modulating the E protein ectodomain solvent exposure along with global morphological heterogeneity. Therefore, lipid lens formation and resulting non-uniform E protein-lipid interactions led to exposure of E protein regions which are occluded in the assembled VLPs. This may result in better exposure of E protein epitopes which are otherwise cryptic on virions of flaviviruses.

The conformational dynamics of VLPs was further characterised to assess the contribution of protein-lipid interactions to overall VLP stability and to capture the heterogeneity in epitope exposure due to VLP dynamics as a result of varying lipid composition and number. Previous studies have shown that antibodies target conformationally dynamic epitope sites contained in E protein monomer or dimer or across more complex quaternary sites on the flavivirus virion surface. Also, epitope exposure is a crucial variable among vaccine candidates that governs the magnitude of the host immune response. Therefore, the solvent accessible surface area (SASA) was used to assess the extent of epitope exposure at domain-I (DI), domain-II (DII) and domain-III (DIII) of VLPs from the CG simulation trajectories as E protein ectodomain major targets of neutralising antibodies. Fig. 11A-B shows the distribution of SASA of domains from each VLP CG simulation. It was observed that VLPs with different lipid composition and varying lipid number show distinct propensity of domain exposure in line with the morphological heterogeneity as discussed in previous section. VLPs with maximum number of lipid molecules in POPC-dominant lipid vesicle showed highest exposure of all three domains. This indicated that the over-packing of lipid molecules results in asymmetric shape which disrupts native contacts between E dimers. The disruption of interactions at the dimer interface enhances overall VLP dynamics and results in exposure of inter and intra-dimeric regions in VLPs. Also, comparing number of protein

(E protein shell)-lipid (vesicle) contacts over the simulation time clearly shows that VLP-1 maintains more interaction points compared to VLP-2 and 3.

In the case of DAG-dominant VLPs, due to the lipid lens formation, the VLPs shrank in diameter and SASA remained constant across all DAG VLP models along with drastic reduction in protein-lipid contacts. Comparing SASA of domains between DAG- and POPC-dominant VLPs showed that in the latter system VLPs have higher magnitude of domain which can potentially present more epitopes to host immune system. Our analysis showed that the lipid molecules incorporated into the VLPs dictate the overall morphology and size of the particle. This explains the structural heterogeneity observed among VLPs in multiple cryo-EM studies, i.e., the incorporation of varying number of lipid species at the ER membrane produces heterogeneous VLP sizes. In the case of virion, the nucleocapsid and capsid protein may fine-tune the lipid number to achieve spherical morphology. Additionally, the distance between dimers of each VLP was calculated with respect to time and the distribution plot of minimum distances between each dimer showed that the dimers in POPC-dominant VLPs fluctuated more than in DAG-dominant VLPs.

Principal component analysis was performed on every third protein backbone bead of VLP from a concatenated VLP trajectory of 6 μ s. The dominant motion in POPC-dominant VLPs were small scale coordinated convergent movements of E dimers along the circumference of the lipid vesicle in VLP2 and VLP-3 which retains the spherical morphology. For VLP-1 with an ellipsoid morphology and large scale motion, it was observed that both E protein dimers underwent convergent and divergent motions. In the case of DAG-dominant VLPs, due to the formation of lipid lens and non-uniform distribution of lipid molecules, large scale E monomer motion resulted in shrinkage of VLPs. Overall, the VLP CG simulations revealed the role of E-M protein shell-lipid interactions in influencing the morphology and modulating the VLP dynamic motions and epitope exposure.

EXAMPLE 4: Discussion

Site-directed mutagenesis and multi-scale simulations were used to investigate biochemical and structural aspects of flavivirus VLPs to guide the development of vaccines based on enveloped VLPs. According to cryoEM data, three structural motifs of flavivirus VLPs predominantly interact with the vesicle envelope lipids: i) the envelope protein H1 helix (E-

H1, residues 396-420); ii) the envelope protein H2 helix (E-H2, residues 428-447); and iii) the TM helices (E-TM, residues 450-495). The E-H1 helix is positioned on the membrane surface, interacting with the outer leaflet. The E-H1 helix of the stem region was shown to be positioned with a slight 'tilt' with respect to the membrane. Considering the structural features and amino acid composition of the E-H1 helix, *in silico* mutations were introduced to residues pointing towards the lipids: residues 398, 402, 405, and 409 (Fig. 3). The aim was to strengthen protein-lipid interactions to give a more robust VLP envelope with improved secretion levels.

To assess key hot-spots for determining lipid-protein binding, alanine (A) or glutamate (E) mutations were first introduced, followed by leucine (L) and isoleucine (I) mutations at these specific sites as single, double, and triple mutations. Extensive molecular dynamics simulations were performed with the native DENV E/M protein dimer, the JEV hybrid E-M construct, and corresponding E-H1 helix mutants, embedded in a bilayer model of experimentally determined lipid composition. The simulations revealed that disruptive mutants (E/A) decreased the number of protein-lipid contacts and can affect the secondary structure of E-H1. This was in agreement with transfection and secretion experiments measured by ELISA, where relative secretion levels of VLPs were significantly affected by those mutations (Fig. 4).

20

Simulation of the envelope and membrane protein dimers in which the selected residues were mutated to non-polar Leu/Ile showed increased protein-lipid interactions and greater stability of the protein within the lipid bilayer. Subsequent experimental secretion assays of single Leu mutants showed a ~10–20% increase in secretion when compared to native VLPs carrying chimeric DENV proteins or JEV VLP hybrid constructs (Fig. 9B). Selected double mutants were then shown to have a ~2-fold increase (A409L and F409L) in secretion (Fig. 9B).

Overall, site-directed mutations in the E protein have identified critical sites for engineering to enhance VLP secretion. The all-atom simulated system provides a narrow and focused view of dynamics and interactions at the E dimer and lipid interface, and it is plausible that the E-H1 mutants may exhibit a higher degree of difference compared to wild-type E protein during VLP maturation. Other biophysical features like quaternary contacts and membrane curvature may also explain the sensitivity of E-H1 helix to mutations. An additive effect

30

composed of amino acid polarity, local dynamics and peptide-lipid interaction dictate the secretion efficiency and VLP assembly, and VLP engineering should thus focus on enhancing E-H1 helix-lipid interactions.

- 5 CG simulations with POPC-dominant lipid system revealed the number of lipid molecules needed to attain a spherical shape, which can be used to reduce the heterogeneity in VLP sizes. Non-spherical and heterogenous VLPs, which constitute majority of VLPs during production, display a diverse set of epitopes. The use of these VLPs may reduce ADE caused by epitope-like fusion loop epitopes (FLEs).

10

It will be appreciated that many further modifications and permutations of various aspects of the described embodiments are possible. Accordingly, the described aspects are intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims.

15

CLAIMS

1. An isolated flavivirus envelope (E) polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain is distinguished from a wild-type stem-transmembrane domain by two or more amino acid substitutions in the wild-type stem-transmembrane domain, wherein the two or more amino acid substitutions are at positions corresponding to positions 8, 11 or 15 of SEQ ID NO: 1.
2. The isolated flavivirus E polypeptide of claim 1, wherein each of the two or more amino acid substitutions is a substitution to a non-polar amino acid residue.
3. The isolated flavivirus E polypeptide of claim 2, wherein the non-polar amino acid residue is alanine (A), phenylalanine (F), isoleucine (I), leucine (L), methionine (M), proline (P), valine (V) or tryptophan (W).
4. The isolated flavivirus E polypeptide of claim 3, wherein the non-polar amino acid residue is isoleucine (I), leucine (L) or phenylalanine (F).
5. The isolated flavivirus E polypeptide of any one of claims 1 to 4, wherein the ectodomain is from dengue virus (DENV), Zika virus (ZIKV), West Nile virus (WNV), yellow fever virus (YFV) or Japanese encephalitis virus (JEV).
6. The isolated flavivirus E polypeptide of claim 5, wherein the ectodomain is from DENV.
7. The isolated flavivirus E polypeptide of any one of claims 1 to 6, wherein the stem-transmembrane domain is from dengue virus (DENV), Zika virus (ZIKV), West Nile virus (WNV), yellow fever virus (YFV), or Japanese encephalitis virus (JEV).
8. The isolated flavivirus E polypeptide of claim 7, wherein the stem-transmembrane domain is from DENV or JEV.
9. The isolated flavivirus E polypeptide of any one of claims 1 to 8, wherein the stem-

transmembrane domain further comprises one or more amino acid substitutions at positions corresponding to position 1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or 101 of SEQ ID NO: 1.

10. The isolated flavivirus E polypeptide of claim 9, wherein the stem-transmembrane domain further comprises one or more amino acid substitutions at positions corresponding to position 14, 19, 26, 28, 35, 37, 39, 40, 49, 51, 52, 58, 59, 60, 61, 62 or 63 of SEQ ID NO: 1.
11. The isolated flavivirus E polypeptide of claim 9 or 10, wherein each of the one or more amino acid substitutions is a substitution to a non-polar amino acid residue.
12. The isolated flavivirus E polypeptide of claim 11, wherein the non-polar amino acid residue is alanine (A), phenylalanine (F), isoleucine (I), leucine (L), methionine (M), proline (P), valine (V) or tryptophan (W).
13. The isolated flavivirus E polypeptide of claim 12, wherein the non-polar amino acid residue is isoleucine (I), leucine (L) or phenylalanine (F).
14. The isolated flavivirus E polypeptide of any one of claims 9 to 13, wherein the polypeptide comprises an amino acid substitution at positions corresponding to one of the following sets of positions of SEQ ID NO: 1:
 - (a) 8, 14 and 15;
 - (b) 8, 11, 14 and 15;
 - (c) 8, 15, 26 and 28;
 - (d) 8, 15, 35, 37 and 39;
 - (e) 8, 15, 40, 49, 51 and 52; and
 - (f) 8, 15, 58, 59, 61, 62 and 63.
15. The isolated flavivirus E polypeptide of any one of claims 9 to 14, wherein the stem-transmembrane domain comprises at least two of the following:

- (a) a substitution to F, I or L at a position corresponding to position 8 of SEQ ID NO: 1;
- (b) a substitution to F, I or L at a position corresponding to position 11 of SEQ ID NO: 1;
- (c) a substitution to F, I or L at a position corresponding to position 14 of SEQ ID NO: 1;
- (d) a substitution to F, I or L at a position corresponding to position 15 of SEQ ID NO: 1;
- (e) a substitution to F, I or L at a position corresponding to position 26 of SEQ ID NO: 1;
- (f) a substitution to F, I or L at a position corresponding to position 28 of SEQ ID NO: 1;
- (g) a substitution to F, I or L at a position corresponding to position 35 of SEQ ID NO: 1;
- (h) a substitution to F, I or L at a position corresponding to position 37 of SEQ ID NO: 1;
- (i) a substitution to F, I or L at a position corresponding to position 39 of SEQ ID NO: 1;
- (j) a substitution to F, I or L at a position corresponding to position 40 of SEQ ID NO: 1;
- (k) a substitution to F, I or L at a position corresponding to position 49 of SEQ ID NO: 1;
- (l) a substitution to F, I or L at a position corresponding to position 51 of SEQ ID NO: 1;
- (m) a substitution to F, I or L at a position corresponding to position 52 of SEQ ID NO: 1;
- (n) a substitution to F, I or L at a position corresponding to position 58 of SEQ ID NO: 1;
- (o) a substitution to F, I or L at a position corresponding to position 59 of SEQ ID NO: 1;
- (p) a substitution to F, I or L at a position corresponding to position 61 of SEQ ID NO: 1;
- (q) a substitution to F, I or L at a position corresponding to position 62 of SEQ ID NO: 1; or

- (r) a substitution to F, I or L at a position corresponding to position 63 of SEQ ID NO: 1;
16. The isolated flavivirus E polypeptide of any one of claim 15, wherein the stem-transmembrane domain comprises at least two of the following:
- (a) a substitution to L at a position corresponding to position 8 of SEQ ID NO: 1;
 - (b) a substitution to L at a position corresponding to position 11 of SEQ ID NO: 1;
 - (c) a substitution to L at a position corresponding to position 14 of SEQ ID NO: 1;
 - (d) a substitution to L at a position corresponding to position 15 of SEQ ID NO: 1;
 - (e) a substitution to F at a position corresponding to position 26 of SEQ ID NO: 1;
 - (f) a substitution to L at a position corresponding to position 28 of SEQ ID NO: 1;
 - (g) a substitution to L at a position corresponding to position 35 of SEQ ID NO: 1;
 - (h) a substitution to L at a position corresponding to position 37 of SEQ ID NO: 1;
 - (i) a substitution to L at a position corresponding to position 39 of SEQ ID NO: 1;
 - (j) a substitution to L at a position corresponding to position 40 of SEQ ID NO: 1;
 - (k) a substitution to L at a position corresponding to position 49 of SEQ ID NO: 1;
 - (l) a substitution to L at a position corresponding to position 51 of SEQ ID NO: 1;
 - (m) a substitution to L at a position corresponding to position 52 of SEQ ID NO: 1;
 - (n) a substitution to L at a position corresponding to position 58 of SEQ ID NO: 1;
 - (o) a substitution to L at a position corresponding to position 59 of SEQ ID NO: 1;
 - (p) a substitution to L at a position corresponding to position 61 of SEQ ID NO: 1;
 - (q) a substitution to L at a position corresponding to position 62 of SEQ ID NO: 1; or
 - (r) a substitution to I at a position corresponding to position 63 of SEQ ID NO: 1;
17. The isolated flavivirus E polypeptide of claim 16, wherein the stem-transmembrane domain comprises or consists of an amino acid sequence having at least 80% sequence identity to an amino acid sequence set forth in SEQ ID NO: 7–16.
18. An isolated flavivirus envelope (E) polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain comprises at least two of the following:
- (a) a non-polar amino acid residue at a position corresponding to position 8 of SEQ ID NO: 1;
 - (b) a non-polar amino acid residue at a position corresponding to position 11 of SEQ

ID NO: 1; and

- (c) a non-polar amino acid residue at a position corresponding to position 15 of SEQ ID NO: 1.

19. A chimeric flavivirus E polypeptide comprising an ectodomain from a first flavivirus and a stem-transmembrane domain from a second flavivirus, wherein the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to positions 8, 11 or 15 of SEQ ID NO: 1.
20. The chimeric flavivirus E polypeptide of claim 19, wherein the first and second flavivirus are different flaviviruses.
21. A virus-like particle (VLP) comprising a flavivirus E polypeptide comprising an ectodomain and a stem-transmembrane domain, wherein the stem-transmembrane domain comprises two or more amino acid substitutions at positions corresponding to positions 8, 11 or 15 of SEQ ID NO: 1.
22. The VLP of claim 21, wherein the E polypeptide is an E polypeptide of any one of claims 1 to 20.
23. An isolated polynucleotide comprising a nucleic acid sequence encoding a flavivirus E polypeptide of any one of claims 1 to 20.
24. A vector comprising an isolated polynucleotide of claim 21.
25. A host cell comprising an isolated polynucleotide of claim 23 or a vector of claim 24.
26. A method for producing a VLP, comprising culturing a host cell of claim 25 under conditions suitable for expressing the encoded flavivirus E polypeptide to assemble the VLP.
27. The method of claim 26, wherein the host cell is a mammalian cell.
28. The method of claim 26 or 27, further comprising recovering the VLP from the cell

culture.

29. An immunogenic composition comprising a flavivirus E polypeptide of any one of claims 1 to 20, a VLP of claim 19 or 20, an isolated polynucleotide of claim 21, or a vector of claim 22.
30. A flavivirus E polypeptide of any one of claims 1 to 18, a VLP of claim 21 or 22, an isolated polynucleotide of claim 23, a vector of claim 24, or an immunogenic composition of claim 29, for use as a medicament or vaccine.
31. A method of modulating an immune response in a subject, the method comprising administering a therapeutically effective amount of an immunogenic composition of claim 29 to the subject.
32. A method of preventing or treating a flaviviral infection in a subject, the method comprising administering a therapeutically effective amount of an immunogenic composition of claim 29 to the subject.
33. An immunogenic composition of claim 29, for use in preventing or treating a flaviviral infection in a subject.
34. Use of an immunogenic composition of claim 29 in the manufacture of a medicament for preventing or treating a flaviviral infection in a subject.
35. The method of claim 31 or 32, immunogenic composition for use of claim 33, or use of claim 34, wherein the immunogenic composition is administered in two or more doses.