Supporting Information

Dessau and Modis 10.1073/pnas.1217780110

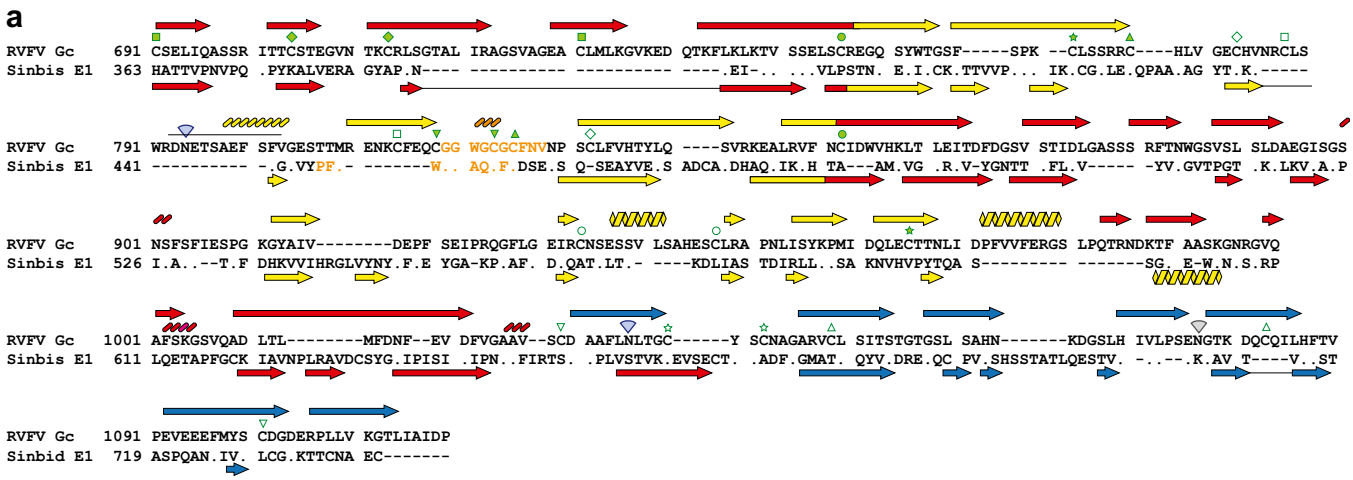

Fig. S1. Amino acid sequence alignments of Rift Valley fever virus (RVFV) glycoprotein G_C and related proteins. (*A*) Structure-based protein sequence alignment of RVFV G_C and Sindbis E1. The sequences were aligned based on a pairwise comparison of the corresponding crystal structures (PDB entries 4HJ1 and 3MUU). Conserved residues were replaced with a period. Domain colors are colored as in Fig. 1. Arrows denote β-strands; zigzags, α-helices; angled ellipsoids, 3_{10} helices. Disulfide bonds are indicated by pairs of matching green symbols. Light blue and gray sectors represent ordered and disordered glycans, respectively. (*B*) Multiple amino acid sequence alignment of G_C from selected phleboviruses. Database sequence accession nos.: RVFV, gb:ABD38819.1; Uukuniemi virus, gi:38371706; Punta

Legend continued on following page

toro virus, gb:AAA47110.1; Toscana virus, gb:ABS85172.1; Massilia virus, gb:ACI24011.1; Sand fly fever virus, gb:AAA75043.1; Candiru virus, gb:AEA30045.1; Heartland virus, gb:AFP33393.1; severe fever with thrombocytopenia syndrome (SFTS) virus, gb:ADZ04471.1. Periods indicate conserved residues. Colors correspond to the domains of RVFV G_C as defined in Fig. 1*A*. The stem region (light gray shading), transmembrane anchor (medium gray shading), and cytoplasmic tail (dark gray shading) are missing in the RVFV G_C crystal structure.

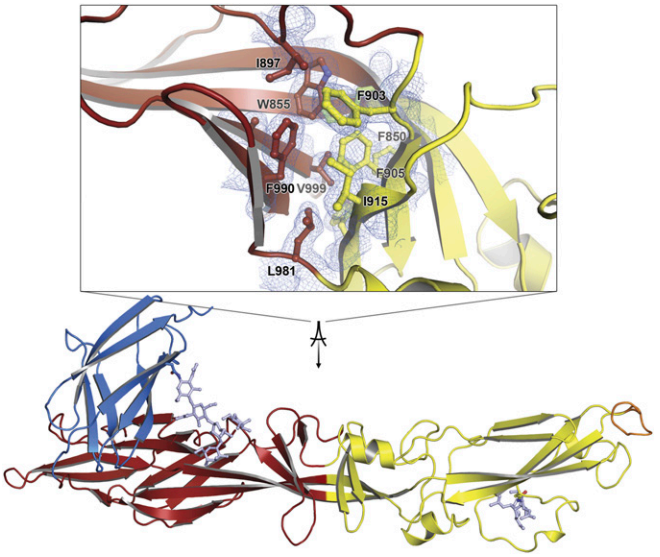

Fig. S2. The β-barrel between domains I and II of RVFV G_C. A close-up (*Inset*; viewed along the direction of the arrow) shows the hydrophobic core of the β-barrel. The environment is similar to the domain I–II interface of dengue E and could conceivably also accommodate a hydrophobic ligand. A $2F_o$–F_c electron density map contoured at 1 σ (blue mesh) is shown around the residues in the hydrophobic core (in ball-and-stick representation).

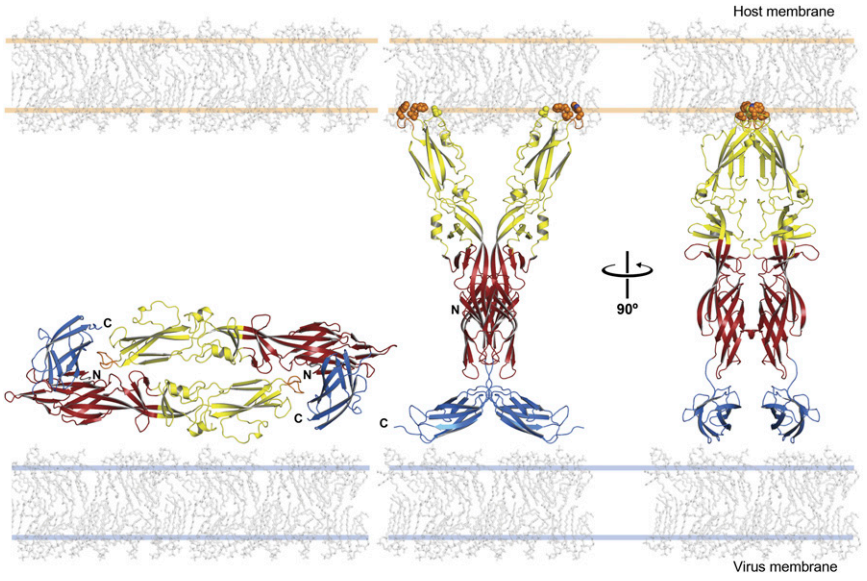

Fig. S3. Nonglycosylated G_C has an extended conformation that is likely to correspond to the so-called prefusion or prehairpin intermediate postulated for all fusion proteins. *Left*, the glycosylated G_C dimer structure shown with its dyad axis perpendicular to the membrane, as in the outer protein shell assembly proposed in Fig. 4. *Center* and *Right*, the nonglycosylated G_C structure. Hydrophobic residues in the fusion loop are shown as orange space-filling spheres. Extensive packing contacts around a crystallographic twofold symmetry axis result in a crystallographic dimer with a buried surface area at the dimer interface of 1,180 $Å^2$. In the orientation imposed by the crystallographic dimer, Leu779 (yellow spheres), on a loop adjacent to the fusion loop, is positioned to insert into the target membrane along with the fusion loop. The positions of the N- and C termini are labeled *Left* and *Center*.

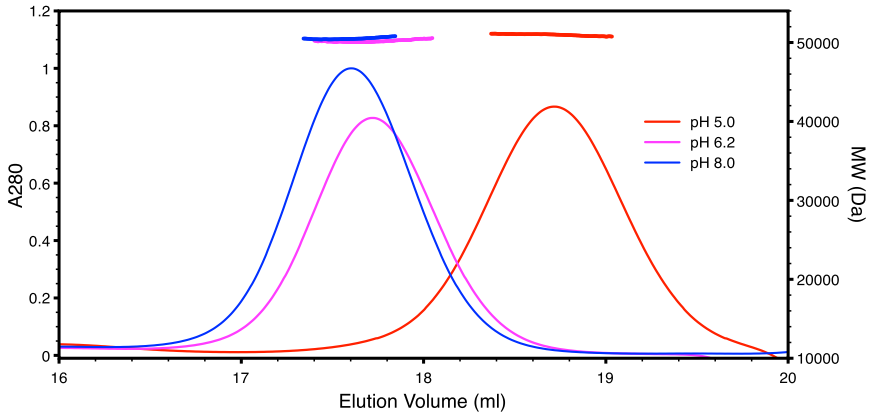

Fig. S4. RVFV G_C is a monomer in solution at different pH conditions. A total of 0.1 mL of G_C (1 g/L) was loaded onto a Superdex 200 (30/100) size-exclusion column preequilibrated with different buffers: red, 50 mM NaOAc pH 5.0; magenta, 50 mM MES pH 6.2; blue, 50 mM Tris·HCl pH 8.0. The eluate was analyzed for absorbance at 280 nm (*Left y* axis) and for multiangle light scattering, which was converted into molecular mass (*Right y* axis; *Materials and Methods*). Elution of G_C was significantly retarded due to nonspecific interactions of the protein with the dextran resin in all three buffers, although the effect was more pronounced in the pH 5 buffer. A similar effect was reported for flavivirus E proteins and attributed to exposure of the hydrophobic fusion loop (1).

1. Kanai R, et al. (2006) Crystal structure of West Nile virus envelope glycoprotein reveals viral surface epitopes. *J Virol* 80(22):11000–11008.

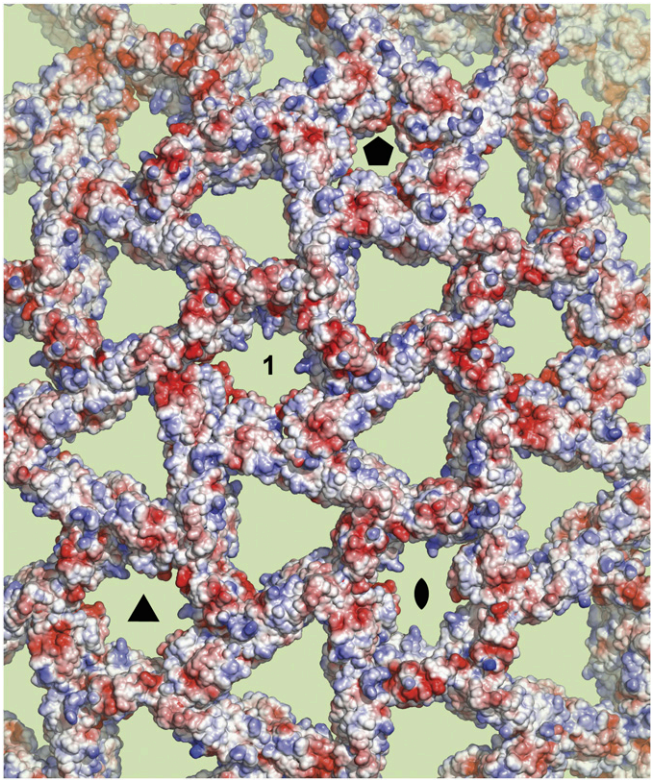

Fig. S5. Electrostatic surface potential of an icosahedral lattice of G_C assembled as shown in Fig. 4. The electrostatic potential of the outer surface of the G_C lattice is similar and slightly negative (red) around the (pseudo)-symmetry axes of each of the four types of capsomers: two-, three-, fivefold (labeled with standard symbols) and pseudo-sixfold, labeled "1." This was prepared with PyMOL using vacuum electrostatic potentials.

Table S1. Quality-of-fit statistics for fitting the crystal structure of the RVFV G_C dimer (PDB ID code 4HJ1) into the RVFV EM structure (EMDataBank ID code EMD-1550)

	Manually fitted coordinates	Coordinates after fitting with UCSF Chimera
Nominal resolution of EM map (Å)	22	22
Contour level of EM map* (σ)	1.3	1.3
Resolution of map calculated from atomic coordinates for fitting/correlation measurement† (Å)	20	20
CC between EM and calculated maps‡	0.64	0.77
CAM between EM and calculated maps§	0.26	0.34
CC with masked EM map used for fitting‡	0.66	0.80
CAM with masked EM map used for fitting§	0.27	0.44
CC with masked EM map, and icosahedral symmetry applied to calculated map during fitting	NA¶	0.88
Total atoms per ASU‖	40,074	40,074
Atoms per ASU outside the unmasked EM density	14,020	8,949
Intra- and intermolecular contacts within the ASU**	35,038	35,964
Intra- and intermolecular clashes within the ASU††	2,055	2,394
Intermolecular contacts with other ASUs**	8,080	6,968
Intermolecular clashes with other ASUs††	3,080	2,658

*The map sampling density was set to 2 in UCSF Chimera.

†The map sampling density was set to 1.

‡CC, correlation coefficient for map-in-map fitting, defined as follows: $CC = \langle \mathbf{u},\mathbf{v} \rangle / |\mathbf{u}||\mathbf{v}|$, where **u** and **v** are vectors expressing the values of the fit and reference maps, respectively, at a certain point in space, and $\langle \mathbf{u},\mathbf{v} \rangle$ (the "overlap", or the inner product of vectors **u** and **v**) is the sum over fit map grid points of the product of the fit map value and the reference map value at that point, determined by trilinear interpolation.

§CAM, correlation about mean map values, defined as $CAM = \langle \mathbf{u} - \mathbf{u}_{ave}, \mathbf{v} - \mathbf{v}_{ave} \rangle / |\mathbf{u} - \mathbf{u}_{ave}||\mathbf{v} - \mathbf{v}_{ave}|$, where $\mathbf{u}_{ave}$ is a vector with all components equal to the average of the components of **u** and $\mathbf{v}_{ave}$ is defined analogously. The CAM equals the cosine of the angle between the vectors (after subtraction of averages) and can range from –1 to 1, whereas the range of overlap (and hence CC) values depends on the scaling of the maps.

¶NA, not applicable.

‖ASU, icosahedral asymmetric unit.

**Contact defined as atom pairs more than four bonds apart with van der Waals overlap ≥ −0.4 Å.

††Clash defined as non–hydrogen-bonding atom pairs more than four bonds apart with van der Waals overlap ≥ 0.6 Å, or potentially hydrogen-bonding atom pairs more than four bonds apart with van der Waals overlap ≥ 1.0 Å.