SUPPLEMENTARY INFORMATION

CryoEM Structure of the Mature Dengue Virus at 3.5-Å Resolution

Xiaokang Zhang [1–5,†], Peng Ge [1–3,†], Xuekui Yu [1–3], Jennifer M. Brannan [3, ‡], Guoqiang Bi [4,5], Qinfen Zhang [6], Stan Schein [2,7], Z. Hong Zhou [1–5]

[1] Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-7364, USA.
[2] California NanoSystems Institute, UCLA, Los Angeles, CA 90095-7227, USA.
[3] Department of Pathology and Laboratory Medicine, University of Texas Medical School at Houston, Houston, TX 77030, USA
[4] Hefei National Laboratory for Physical Sciences at the Microscale, and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China.
[5] School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China.
[6] School of Life Sciences, State Key lab for Biocontrol, Sun Yat-Sen University, Guangzhou, Guangdong 510275, China

[7] Department of Psychology, and the Brain Research Institute, UCLA, Los Angeles, CA 90095-1563, USA

‡ Current address: US Army Medical Research Institute, 1425 Porter Street, Frederick, MD 21702-9211, USA.

† These authors contributed equally to this work.

Correspondence should be addressed to Z.H.Z. (Hong.Zhou@UCLA.edu).

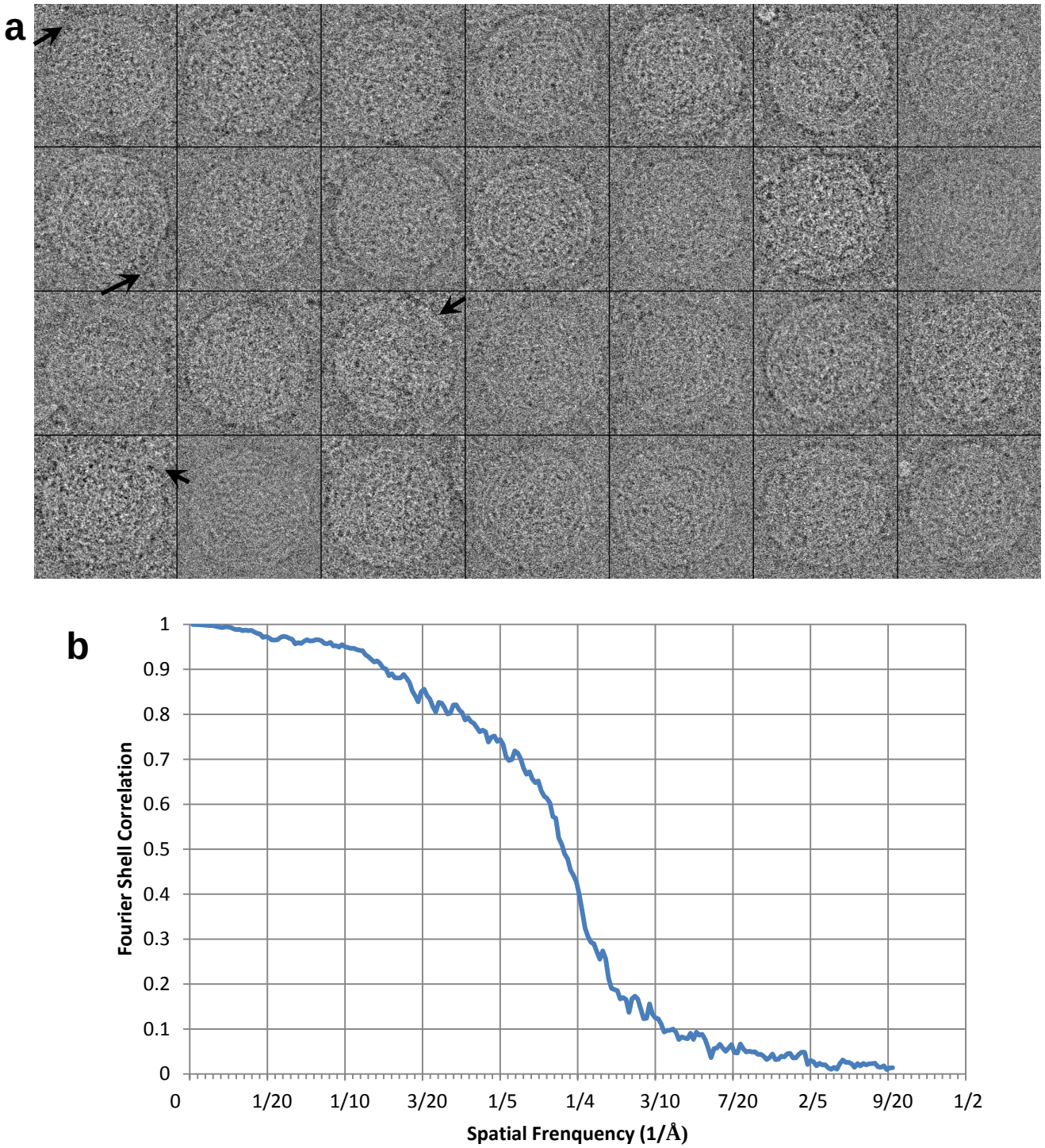
SUPPLEMENTARY INFORMATION

Supplementary Information includes ten supplementary figures, two supplementary tables and two movies.

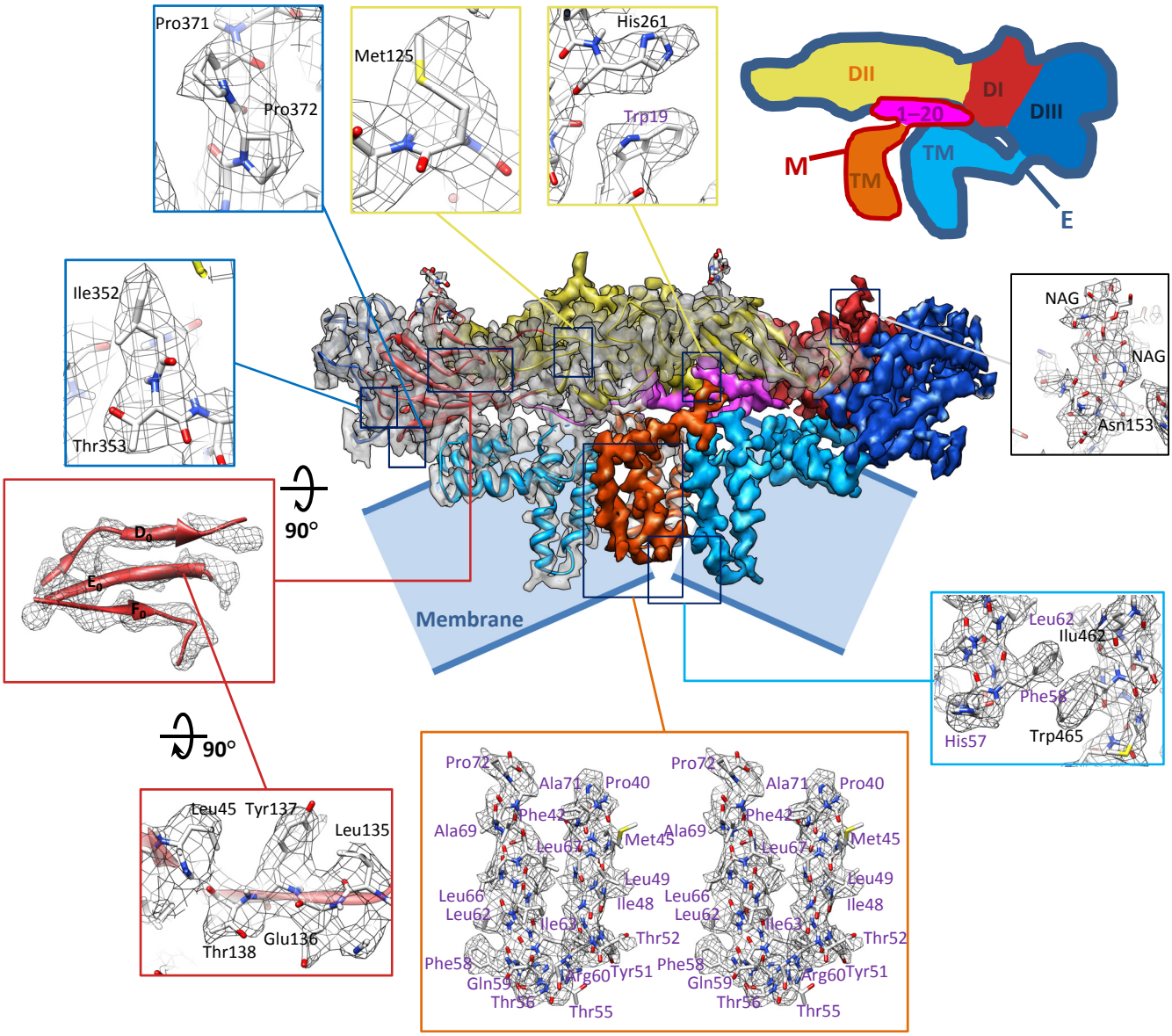
Supplementary Figures 1–10

Supplementary Tables 1–2

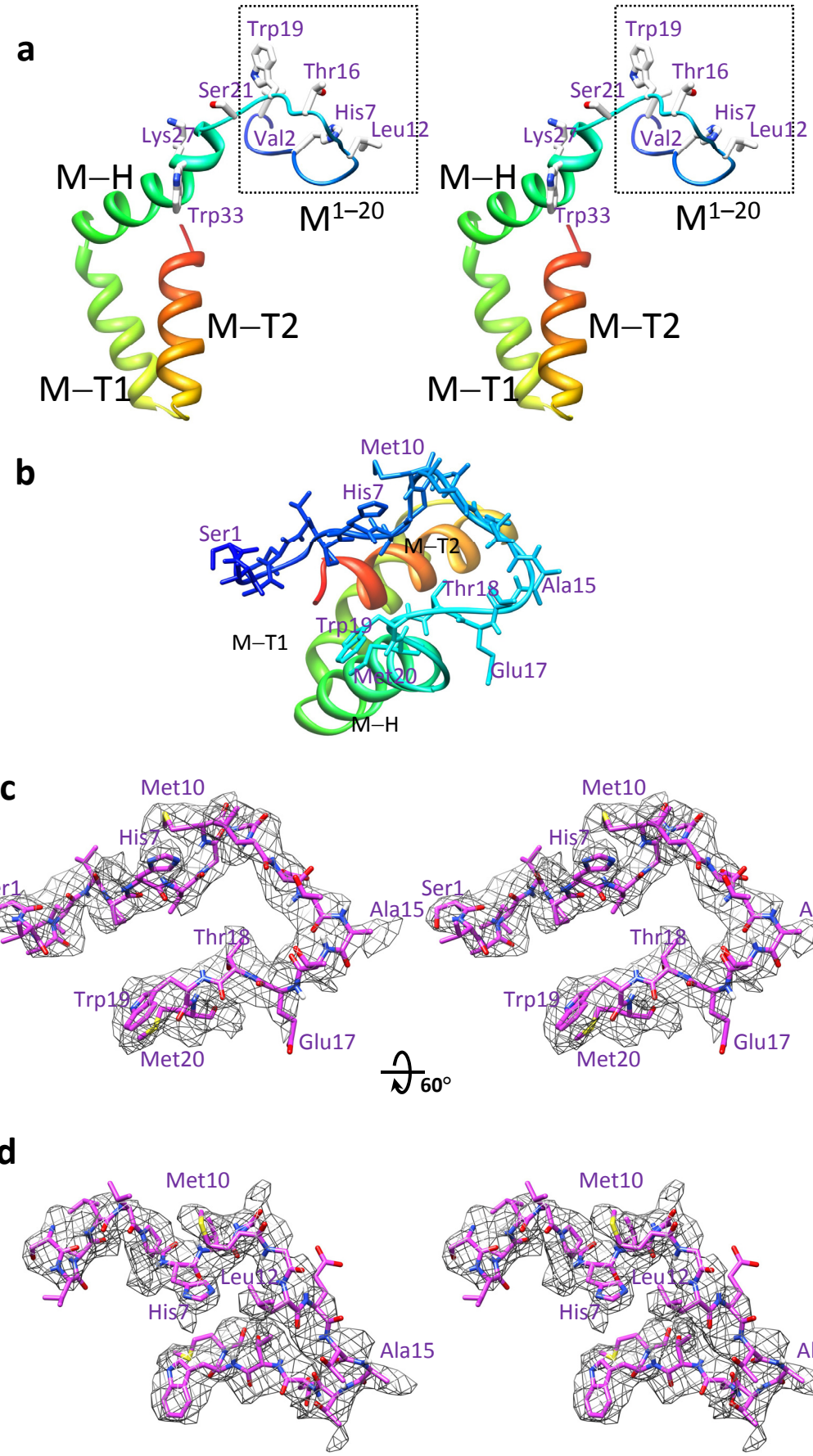
Supplementary Movies 1–2



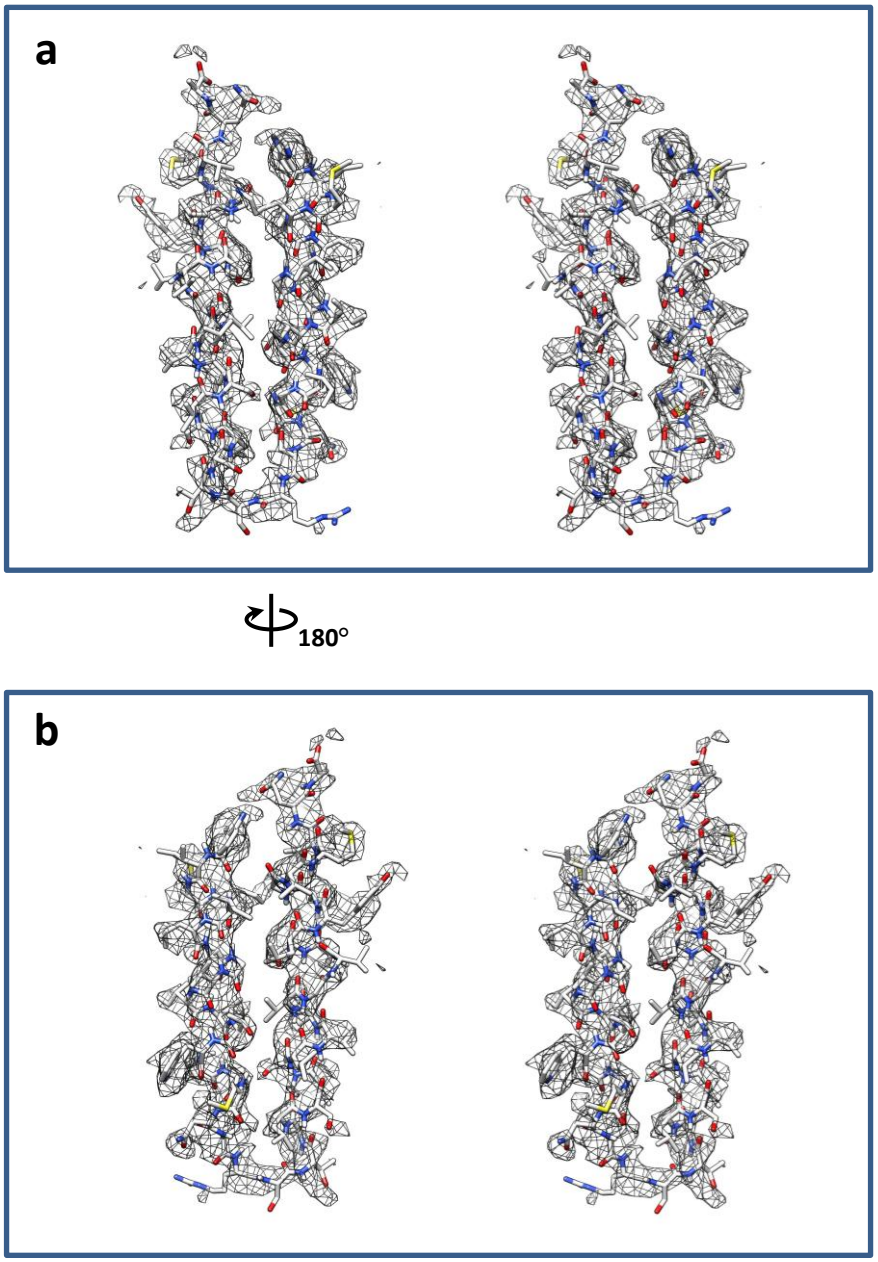
Supplementary Figure 1 Examples of 'good' particles and assessment of 3D reconstruction. (**a**) Top 28 scoring particles from the 9288 particles used for reconstruction. Arrows indicate local imperfections in some of these particles. Box side: 530 Å. (**b)** Fourier shell correlation (FSC) coefficients between the 3D reconstructions from two randomly separated half datasets, plotted against spatial frequency. Using the Rosenthal and Henderson criterion of 0.143 FSC cutoff [14], the effective resolution of the reconstruction is 3.5 Å. If the 0.5 FSC cutoff criterion were used, the effective resolution of the reconstruction would be 4.2 Å.



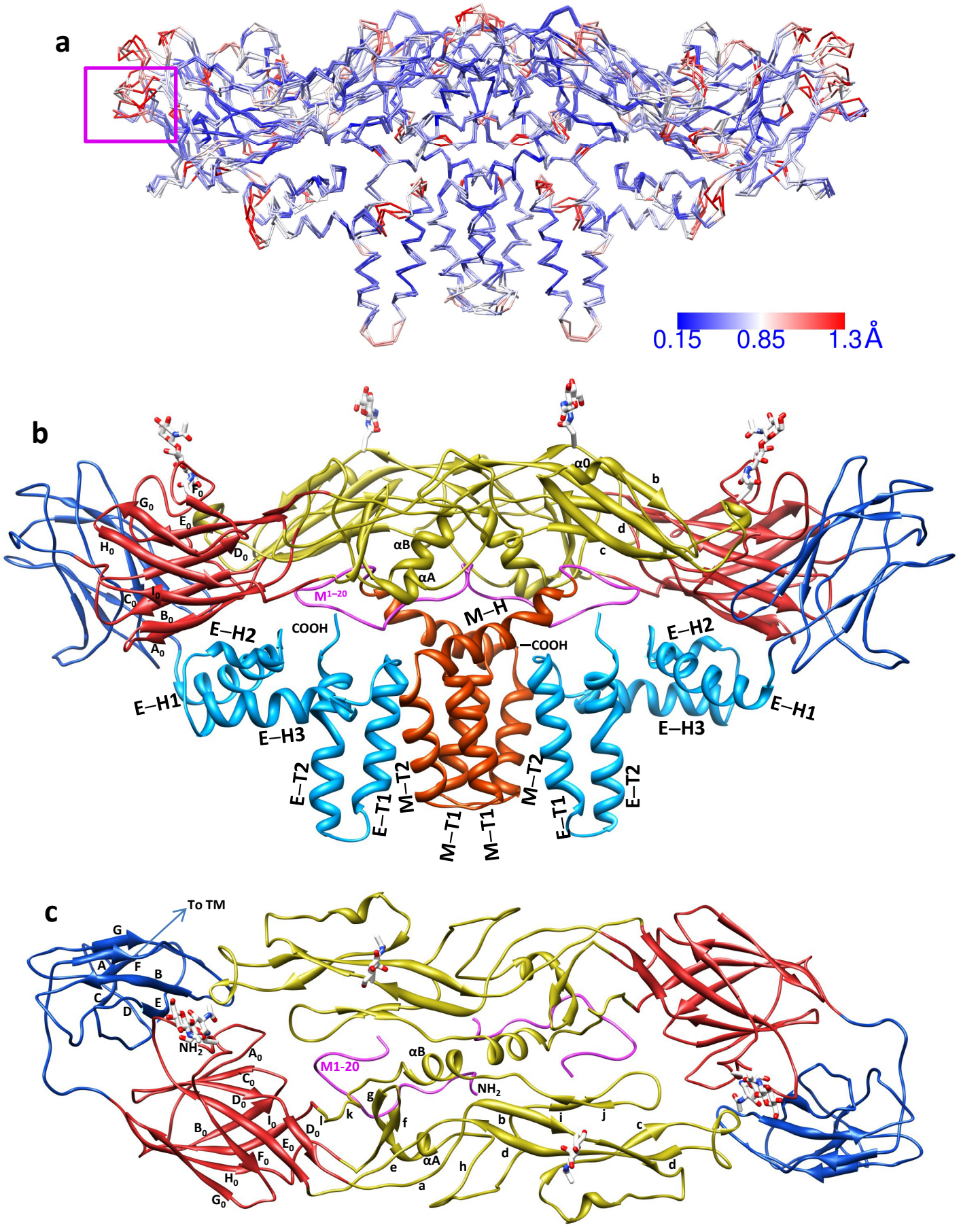
Supplementary Figure 2 Density map for E:M:M:E heterotetramer. Center: Shaded surface representation of the density map of the heterotetramer, composed of two E:M dimers, one at right and also behind in the center, the other at left and also in front in the center. Illustrated in the upper right corner, the color scheme of the domains of E:M is similar to previous work on dengue virus proteins [8–11]: for E, red for domain I (dI), yellow for domain II (dII) and blue for domain III (dIII), comprising the three parts of the ectodomain; cyan for the transmembrane domain (TM); for M, magenta for the first 20 amino acids ($M^{1–20}$) (ectodomain) and orange for the transmembrane domain (TM). The right copy of the E:M dimer is shown in solid surface, the left copy in semi-transparent gray with ribbon models of E and M superimposed. The membrane bilayer is indicated. Surrounding boxes: Representative areas are shown in wireframe superimposed on their atomic models either as sticks or ribbons. In the stick models, residue numbers are indicated (black labels for those in E, purple for those in M).

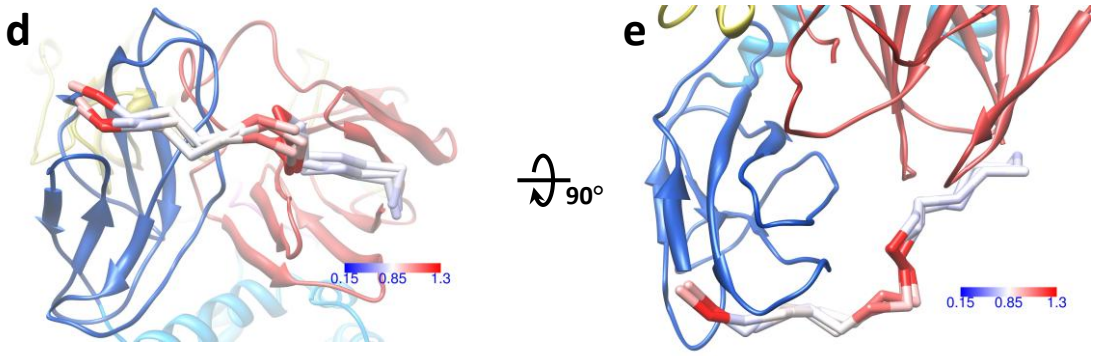

a
Trp19
Thr16
Ser21
His7
Lys27
Val2
Leu12
M–H
Trp33
M1–20
M–T2
M–T1
b
Met10
His7
Ser1
M–T2
Thr18
Ala15
Trp19
M–T1
Met20
Glu17
M–H
c
Met10
His7
Ser1
Ala15
Thr18
Trp19
Glu17
Met20
60°
d
Met10
Leu12
His7
Ala15

Supplementary Figure 3 Stereo views of the atomic model of M. (**a**) Stereo view of Figure 2e. (**b**) Overview of the structure of M, colored and labeled as in panel **a,** to show the orientation of (**c**). (**c** and **d**) Stereo view of the density map of $M^{1–20}$ (mesh) superimposed on its atomic model (sticks) colored by atom types, red O, blue N, yellow S and purple C. (**d**) is rotated horizontally 60° from (**c**).

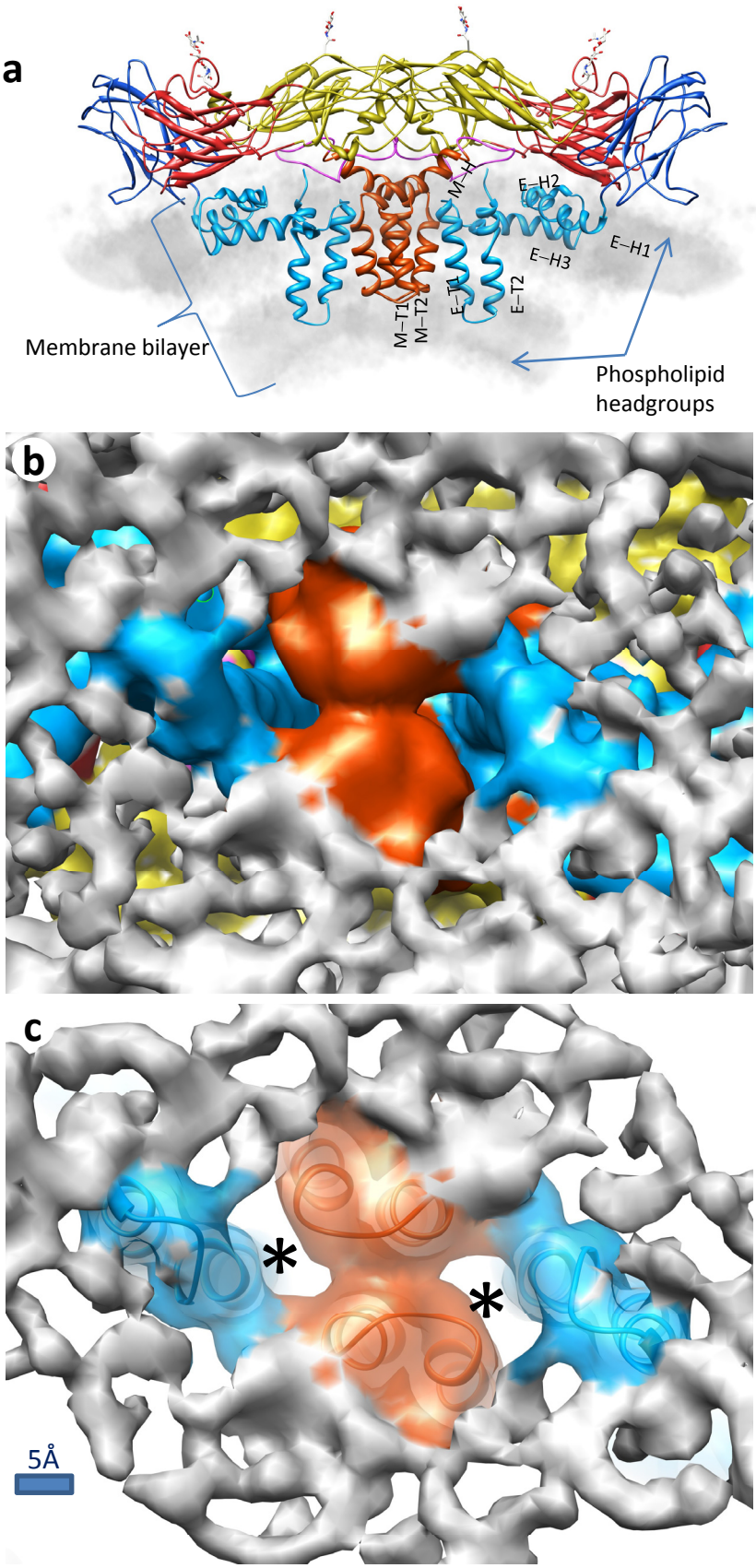


Supplementary Figure 4 Stereo views of the two transmembrane helices of E. (**a**) Stereo view of the density map of E–T1 and E–T2 (mesh) superimposed on its atomic model (sticks) colored by atom types, red O, blue N, yellow S and white C. (**b**) Rotated horizontally 180° from (**a**).

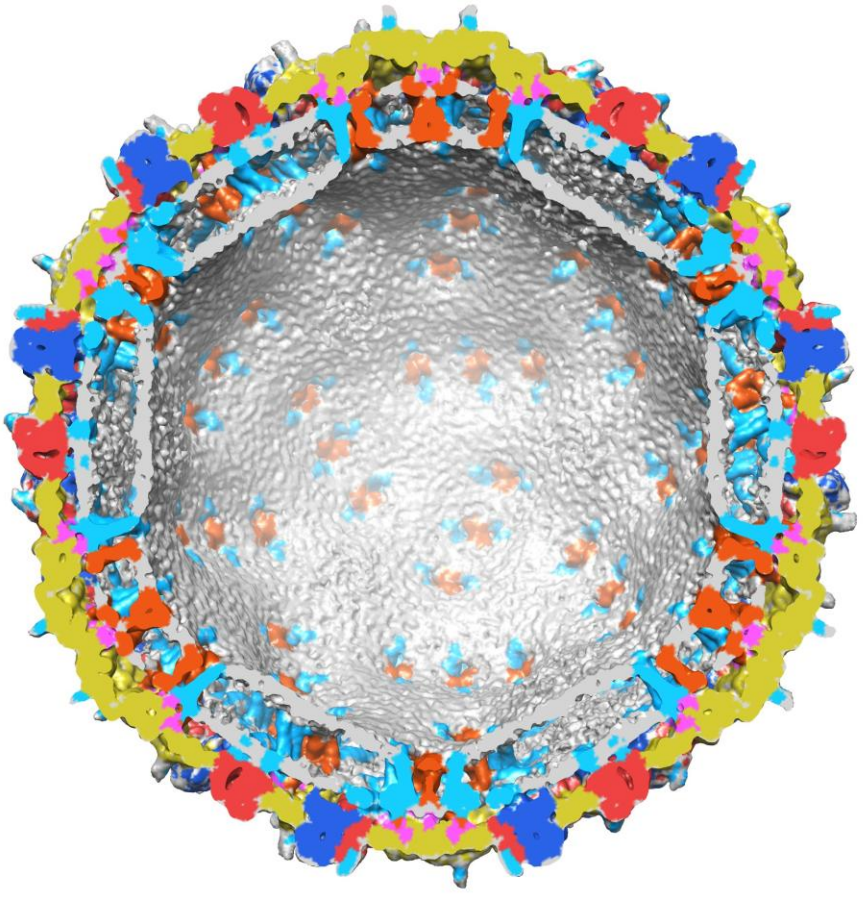

a
0.15
0.85
1.3Å
b
G_0
F_0
E_0
H_0
D_0
C_0
I_0
B_0
A_0
αB
αA
α0
b
d
c
M^{1-20}
M–H
COOH
–COOH
E–H2
E–H1
E–H3
E–T2
E–T1
M–T2
M–T1
c
To TM
G
A
F
B
C
D
E
NH_2
A_0
C_0
D_0
I_0
B_0
E_0
F_0
H_0
G_0
M1-20
αB
NH_2
αA
a
b
c
d
e
f
g
h
i
j
k
l



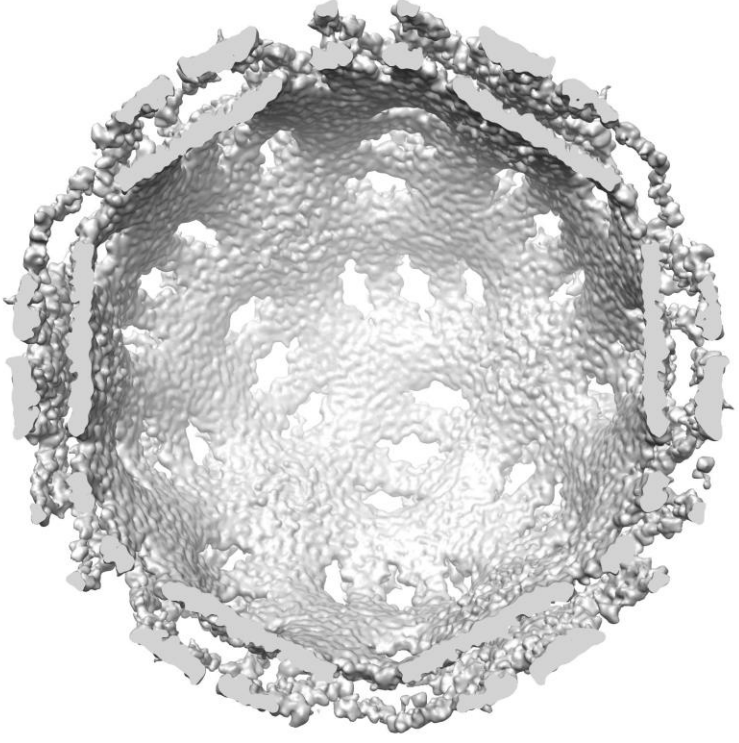
Supplementary Figure 5 Atomic models of E:M:M:E heterotetramers. (**a**) Superposition of the three quasi-equivalent copies of E and of M. Extensive overlap of the three copies of the E:M heterodimer in each asymmetric unit (large triangles in Fig. 1b,c) demonstrates quasi-equivalence. The E:M:M:E tetramer (i.e., dimer of the E:M dimer) is shown using the color scheme in the inset encoding the Cα RMSD in Å of each residue among the three copies. (**b**, **c**) Naming scheme for secondary structure elements. (**b**) The E:M:M:E heterotetramer is duplicated from Figure 2b, but secondary structure elements are labeled with the scheme defined by earlier works (*e.g.*, ref. 10). (**c**) The E:M:M:E tetramer is shown as in Figure 4a, with secondary structure elements in one E and one M labeled. (**d** and **e**) The region in the purple box in (**a**) is here magnified and rotated to highlight the loop that links domains I and III.


a
M–H
E–H2
E–H1
E–H3
E–T1
E–T2
M–T1
M–T2
Membrane bilayer
Phospholipid
headgroups
b
c
*
*
5Å

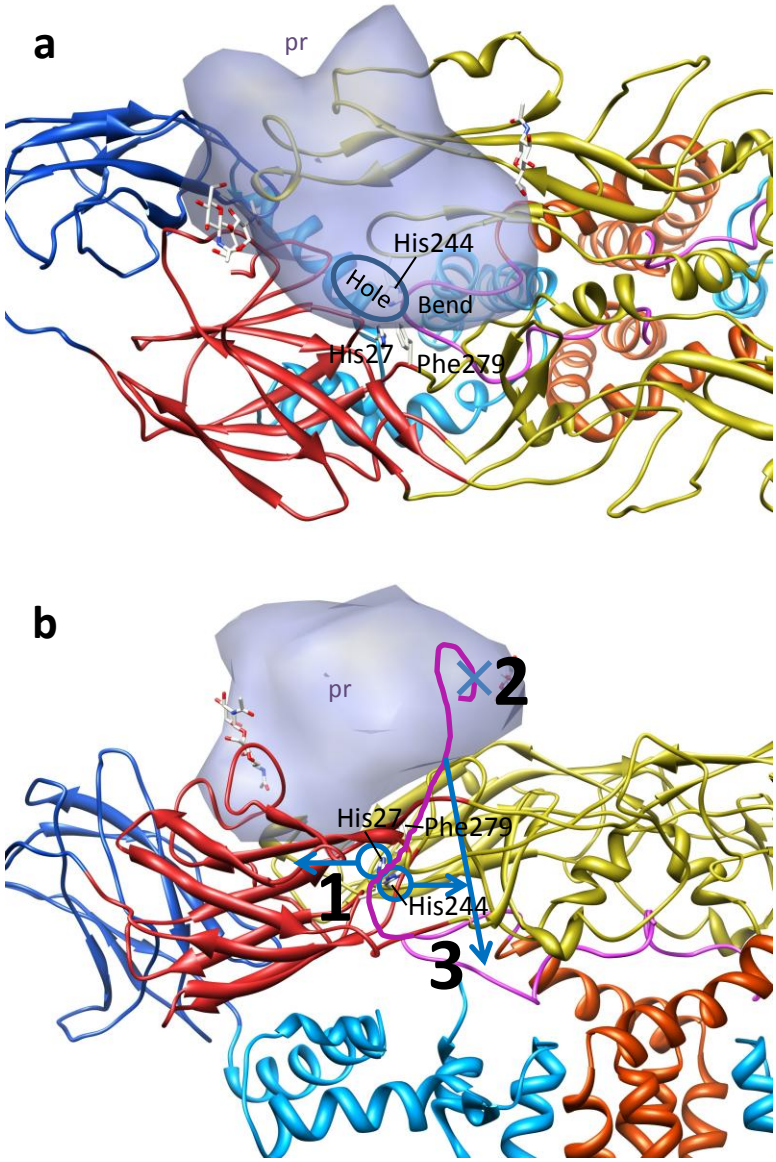
d



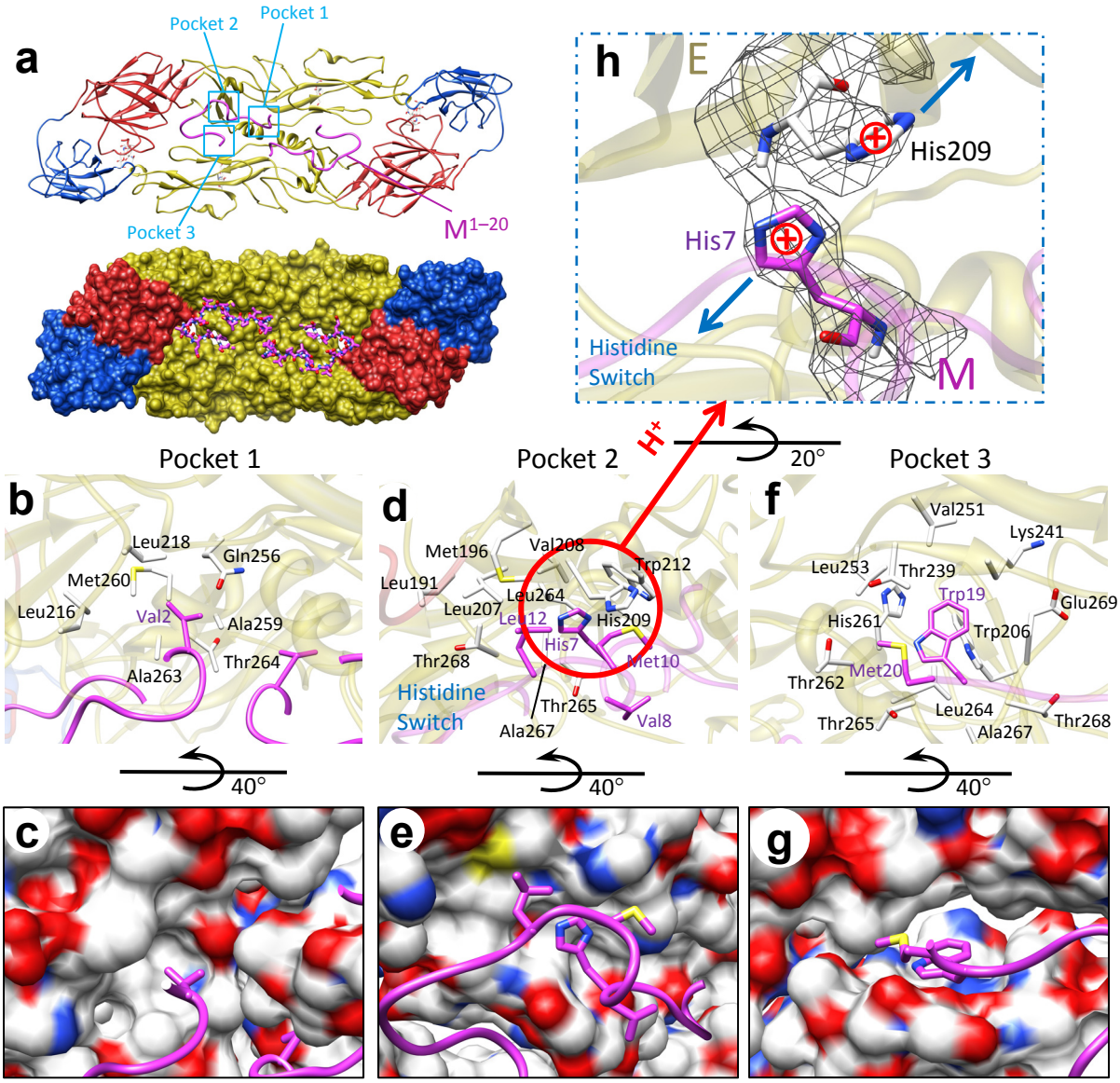
e



Supplementary Figure 6 Orderliness among the phospholipids of the membrane. (**a**) Overview of the E:M:M:E heterotetramer with associated membrane. Transmembrane and perimembrane helices are labeled. The density map prior to Fourier amplitude rescaling and B-factor correction is displayed in gray by volume rendering, superimposed on the atomic model (ribbons). For clarity, the membrane density is segmented to include only the region within 15Å of the E:M:M:E heterotetramer, and the densities for the proteins are removed by zeroing all densities within 2Å of the atoms of the modeled proteins E and M. (**b**) Membrane and E and M proteins viewed from inside the virion. The orderliness of the membrane remains noticeable after the three sets of E:M:M:E tetramers in the rhombic region (Fig. 1c) are averaged. Here, the density map of the averaged tetramer has been low-pass filtered to ~6 Å. Densities that belong to E and M are colored as in Figures 2a–c. Densities that belong to the membrane are shown in grey. (**c**) Same as (**b**) but restricting the depth of view for clarity. The densities attributable to transmembrane helices are shown in semi-transparent contour superimposed with their atomic models (ribbons). The asterisks mark two gaps in the membrane, each ~100 $Å^2$ in area. These gaps correspond to the absence of phospholipid between the four loops (colored) that interconnect the eight transmembrane helices, two per monomer, in an E:M:M:E tetramer. No density is found inside these gaps. (**d**) A cut-away view of the virion showing the arrangement of the holes in (**b**) and (**c**) in the inner leaflet of the viral envelope. The density map of the virion is low-pass filtered to about 8 Å. The colors of density regions follow the scheme in main text figures. (**e**) Another view of (**d**) with all protein densities removed to show the edged membrane structure.



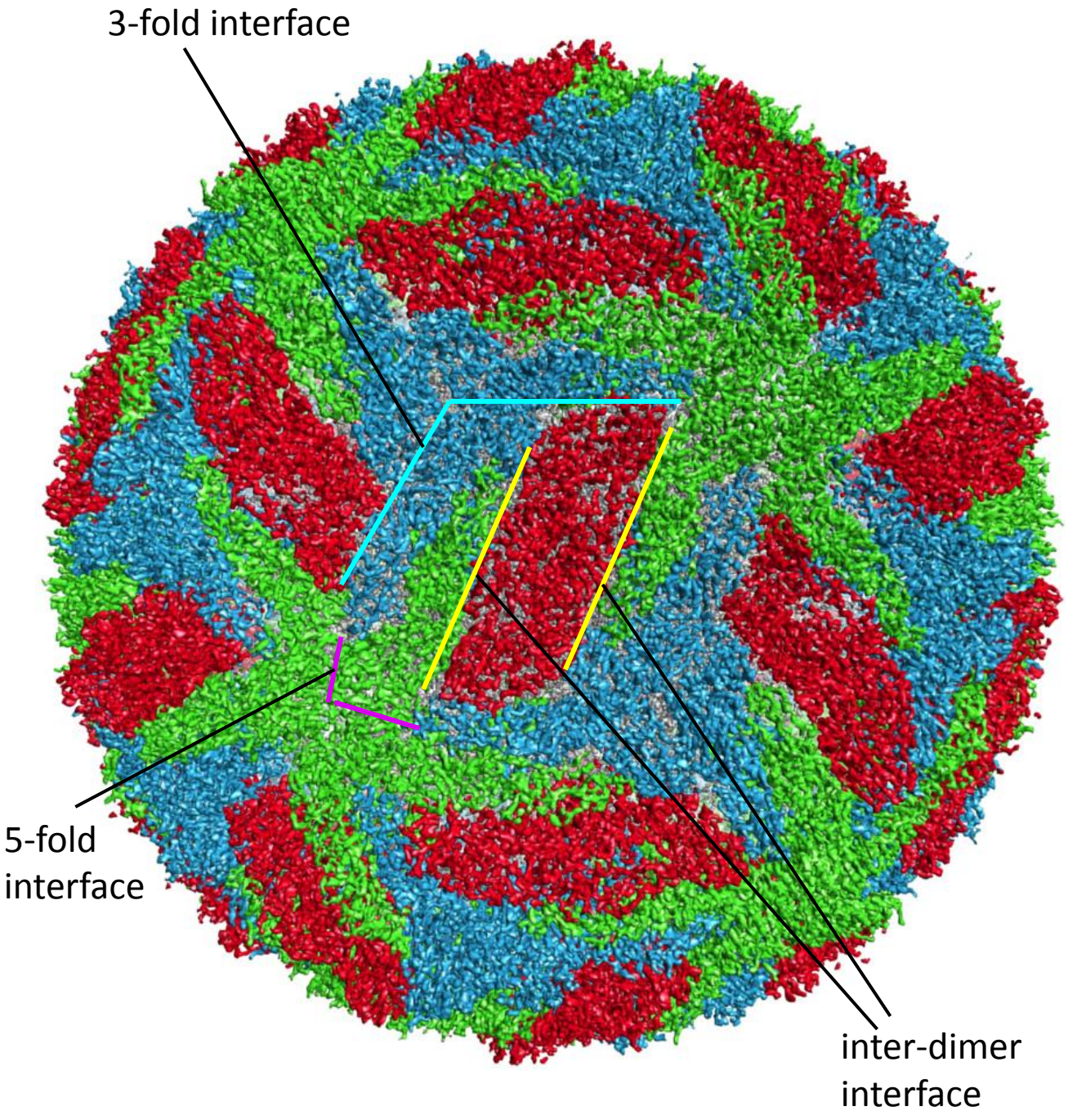
Supplementary Figure 7 The 'hole' that allows $M^{1–20}$ to internalize. (**a**) Binding site of pr (5-Å solvent accessible surface in semi-transparent grey) on E (blue, red, yellow and cyan ribbons). The position of pr is deduced by matching E in the pseudo-atomic model of the (low-pH) smooth immature virion (PDB: 3C6R [3]) onto our atomic model of E in the mature virus. Coinciding with the 'bend' at Thr16 in $M^{1–20}$ (Fig. 2e and Supplementary Fig. 3), there is a 'hole' (navy blue oval) in the E (ribbon) structure. Two histidine residues of E, His27 and His244, oppose each other across the hole to form a double door blocking the hole in the mature virus at neutral pH. (**b**) Illustration of a possible mechanism of the first step in the maturation from the smooth immature particle (Stage 2 in Fig. 5b): (1) At the low pH environment of the TGN, the histidine double door opens due to protonation and consequent charge repulsion between His27 and His244; (2) furin cleaves prM to yield M and pr; (3) this cleavage allows the N-terminal loop of M, namely $M^{1–20}$, to pass through the histidine double door to the same side of the door as its membrane anchor.



Supplementary Figure 8 Key interactions between E and M. (**a**) E:M:M:E heterotetramer from inside the virus. The ribbon model shows three pockets (cyan boxes) on E where M binds. (Transmembrane domains omitted.) The space-filling model of E shows the groove where M (stick model: atoms C-magenta, N-blue, O-red, S-yellow, H-white) binds. See Supplementary Movie 2. (**b–g**) Enlargement of pockets 1–3 viewed along directions that best depict interactions. (**b**, **c**) Val2 and the first few residues of M sit in a big cavity in the inner surface of E. (**d**, **e**) His7, Met10 and Leu12 of M form a hydrophobic core with neighboring residues in E. The two opposing histidine residues (H7 of M and H209 of E), when protonated at low pH, repel each other [pull-out panel, where the density (mesh) of these two residues are superimposed on their models]. (**f**, **g**) A conserved Trp19 (W19) from M inserts into a deep recess along the E:E dimer interface that includes the partially conserved H261 of E (see Supplementary Fig. 9). Panels **c**, **e**, and **g** are derived from panels **b**, **d** and **f** but with 40° rotation and with E shown as a molecular surface colored by atom type (C–white, N–blue, O–red, S–yellow, H–white) and M shown as a purple wire backbone with sticks for key residues. In stick models, atom types are colored as in **a**. See also Supplementary Movie 2.


SeqNo.
Sequence
Furin Signal
M
UniprotID
Name
Val2
His7
Leu12
M
E
Trp19
E
Special Asp for yellow fever (Asp209)
Trp206
Trp212
Leu218
His209
Leu216
E
Met260, His261

Supplementary Figure 9 Identification of conserved residues involved in E:M interactions by multiple sequence alignment of the polyproteins of flaviviruses. Fourteen polyproteins of flaviviruses were aligned by the Uniprot (www.uniprot.org) server. Selected parts with conserved residues are shown here. The letters and arrows above the aligned sequences indicate some boundaries of the viral proteins. The penultimate column shows the Uniprot ID of the protein sequence. The last column lists the names of the polyproteins with their viral source indicated after 'POLG_' (WNV for West Nile, KUNJ for Kunjin, MVEV for Murray Valley encephalitis, JAEV for Japanese encephalitis, STEV for St. Louis encephalitis, DEN1–4 for dengue virus subtypes 1–4, YEFV for yellow fever, TBEVW, TBEVS and POWV for the European, Far Eastern and Powassan tick-borne encephalitis and LANV for Langat virus). The conserved residues involved in E:M interaction are marked by red boxes. The dengue virus amino acid at each of these conserved positions is indicated underneath each box.



Supplementary Figure 10 Definitions of interfaces mentioned in Supplementary Table 2. The isosurface view of the virion is annotated with colored lines denoting different interfaces mentioned in Supplementary Table 2. The interface around the 5-fold axis is designated 5-fold interface (purple lines); the interface around the 3-fold axis is designated 3-fold interface (cyan lines); the interface between two dimers in a single rhombus is designated inter-dimer interface (yellow lines).

Supplementary Table 1. Interactions between E of one E:M:M:E heterotetramer and E of a neighboring E:M:M:E heterotetramer at interfaces shown in Supplementary Fig. 8.

Interface (Sup. Fig. 7)	Residue 1	Res. 1 subunit	Residue 2	Res. 2 subunit	Type of interaction
five-fold	Ser298	5f	Lys307	5f to the left	hydrogen-bond (main 1 side 2)
five-fold	Glu338	5f	Lys388	5f to the left	salt-bridge
five-fold	His346	5f	Lys344	5f to the left	hydrogen-bond (side chains)
inter-dimer	Glu184	5f	Lys388	the other 2f in rhombus	salt-bridge
inter-dimer	Ser229	5f	Gln86	2f in the same copy in rhombus	hydrogen-bond (main 1 side 2)
inter-dimer	Gln86	5f	Arg89	2f in the same copy in rhombus	hydrogen-bond (side 1 main 2)
inter-dimer	Lys394	3f	Glu195	2f in the same copy in rhombus	salt-bridge
three-fold	Asn83	3f	Met301	2f from a rhombus to the right	hydrogen-bond (side 1 main 2)
three-fold	Gln86	3f	Phe337	2f from a rhombus to the right	hydrogen-bond (side 1 main 2)
three-fold	Gln131	3f	Glu343	3f to the right	hydrogen-bond (side chains)
three-fold	Glu172	3f	Met297	3f to the right	hydrogen-bond (main chains)

Supplementary Table 2. R–factors of individual resolution bins

bins	Resolution range (Å)			R factor (%)
1	7.00	-	100	16.5
2	5.56	-	7.00	24.2
3	4.85	-	5.56	26.1
4	4.41	-	4.85	28.5
5	4.09	-	4.41	35.2
6	3.85	-	4.09	43.4
7	3.66	-	3.85	45.9
8	3.50	-	3.66	48.7

Legends for Supplementary Movies

Supplementary Movie 1. 3D visualization of various structures described in the figures. The animation begins with a surface rendering of the cryoEM density map, rotating around a 2-fold axis. Structural units containing membrane proteins E and M shown in the same color are equivalent by icosahedral symmetry. The differently colored structural units are quasi-equivalent. Specifically, the green units fall on the icosahedral 5-fold axes, the blue on the 3-fold and the red on the 2-fold. This scene is followed by a close up view of a rhombus-shaped group of six E:M dimers, fitted with the ribbon representations of its atomic model, rotating around the horizontal axis. Next, the three quasi-equivalent E:M:M:E heterotetramers are averaged. Rotated around the horizontal axis, half of this averaged tetramer is rendered as a shaded surface representation and half is shown in semi-transparent grey, superimposed with the ribbon representations of an E monomer and an M monomer, with the same color scheme as in Figure 2.

Supplementary Movie 2. Interactions between E (molecular surface) and M (sticks). First, an animated view of Figure 4b. Second, an animated view of the molecular surface of E and the ribbon and stick model of M that comprise pocket 1, as shown in Figure 4c and Supplementary Fig. 8c. A second pocket 1 in a symmetry related position is also visible in the same view. Thirs, an animated view of the molecular surface of E and the ribbon and stick model of M that comprise pocket 2, as shown Figure 4d and Supplementary Fig. 8e. The histidine in the center of this movie is His7 of M which is involved in the pH sensitive latching of E by M. The two contiguous nitrogen atoms in a nearby bulge above and to the right of this His7 belong to His209 of E. Finally, an animated view of the molecular surface of E and the ribbon and stick model of M that comprise pocket 3 as shown Figure 4e and Supplementary Fig. 8g. The surrounding of the central Trp19 of M is highly hydrophobic as indicated by the atom types on the molecular surface of E.