Structure and receptor recognition by the Lassa virus spike complex

Extended Data for:

Structure and Receptor Recognition by the Lassa Virus Spike Complex

Michael Katz[1], Jonathan Weinstein[2], Maayan Eilon Ashkenazy[1], Katrin Gehring[1], Hadas Cohen-Dvashi[1], Nadav Elad[3], Sarel J Fleishman[2] & Ron Diskin[1*]

1 Department of Chemical and Structural Biology, Weizmann Institute of Science, Rehovot 7610001, Israel

2 Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel

3 Department of Chemical Research Support, Weizmann Institute of Science, 76100 Rehovot, Israel

* Corresponding author. Email: ron.diskin@weizmann.ac.il

This file contains:

Extended results and discussion.
Extended tables 1 & 2.
Extended figures 1-11.
Extended methods.

Extended results and discussion:

Solving the structure of the spike. The gold-standard FSC for the C3-symmetric density map indicated a resolution of 2.46 Å (Extended data Fig. 1). Careful inspection of the density, however, suggested that the actual resolution is somewhat lower. Based on the shape of the density for side chains, we initially estimated the resolution as ~3.0 Å. Indeed, fitting a model to the density map yielded a map-model FSC of 3.3 Å (Extended data figure 2). The discrepancy in the resolution estimates suggested that the map was potentially over-fitted during 3D reconstruction, perhaps due to oversampling of top views (Extended data Fig. 1). To avoid any high-frequency noise that may have been introduced to the map, we applied to it a global low-pass filter at 3.3 Å and used the filtered map to re-refine the structure (Extended data figure 2). We used the same procedure to refine the structure in the C1-symmetric focused map of the ectodomain. Namely, we fitted the model to the density map that had a GSFSC=2.62 Å (Extended data Fig. 5), achieved a map-model FSC of 3.7 Å, and re-refined the structure using a low-pass filtered map at 3.7 Å (Extended data Fig. 7).

The topology of the SSP in the membrane. Over the years, different models for the topology of SSP were proposed[1-4]. A more recent and cited study by Agnihothram S.S. *et al.* assumed a model with a bitopic configuration of SSP, in which its two hydrophobic regions cross the membrane[1]. This model implies that the membrane-spanning domain of the spike should consist of nine TM-helices. Our structural data, however, indicate the presence of only six TM-helices (Fig. 1c, Extended Data Fig. 2), an observation that is not consistent with a bitopic model. Also, our structure of the spike shows that the N-terminus of SSP is located at the outer side of the membrane (Fig. 1e, Extended Data Fig. 3a), whereas Agnihothram's model[1] suggests that it should be located at the inner side of the membrane. While the observed configuration of the SSP is not supporting the previously proposed models, it is consistent with some published biochemical data. Specifically, Cys57 was shown to participate in zinc-coordination together with the cytoplasmic tail of GP2[5], an observation that is consistent with our structure that places the C-terminus of the SSP inside the cell. Also, it was demonstrated that the first hydrophobic region of the SSP interacts with GP2 inside the membrane[6], which is also consistent with our structure. Interestingly, Eichler R. et al. already came to the conclusion that only the first hydrophobic region of SSP is inserted into the membrane[2]. This

insight is consistent with our structure, albeit Eichler R. proposed an inverted topology (i.e., N-terminus inside). Lastly, our structure indicates that the positively charged residues of SSP are indeed at the cytoplasmic side of the membrane (Fig. 1g), an organization that is favored (i.e., the "positive inside rule")[7]. Overall, the fact that the SSP undergoes a topology switch that separates its C-terminus from the N-terminus of GP1 across the membrane may have been a factor that complicated previous experimental attempts to determine its actual topology.

The structure of the LASV spike complex is consistent with previously published mutagenesis data. There is an extensive body of experimental work that explores the functional effects of mutating residues of the arenaviral spike complexes. Our structure can rationalize the effects of some of these mutations. Specifically, our structure indicates that the conserved Pro12 in SSP (Fig. 1g) serves as a helix breaker at the end of h1 of SSP (Fig. 1f). This observation implies that other residues at this position might destabilize the interaction of SSP with the spike by disfavoring the bound conformation of SSP, which requires the breakage in helicity between h1 and h2. Indeed, mutating Pro12 of LCMV to either alanine or glycine does not affect the surface expression of the spike but completely abrogate infectivity of pseudotyped viruses[8], suggesting a premature dissociation of the SSP from the spike. Our structure further indicates that Glu16 and Glu17 of SSP point toward a positively-charged patch on GP2 (Fig. 1f). Mutating the equivalent negatively-charged residues of LCMV (Asp16 and Glu17) to alanine does not affect surface expression but, like in the case of Pro12, abrogates the infectivity of pseudotyped viruses[8]. Reverting the charge of these two residues by mutating them to lysine (i.e., D16K/E17K) results in a more deleterious effect that reduces surface expression and maturation of the spike on top of abrogating infectivity of pseudotyped viruses[8]. This indicates that SSP has reduced association with the spike even inside the ER-Golgi network since this interaction is needed for the proper maturation of the spike[9]. Not only it explains deleterious effects, our structure also explains the lack of effect for mutating some residues. Specifically, mutating either Gln3 or Thr6 of LCMV (corresponding to Gln3 and Val6 in LASV) to alanine does not have any functional consequence on the spike[8]. Both of these SSP residues are predicted by our structure to be solvent exposed (Fig. 1f), rationalizing this observation.

Besides the SSP, our structure is consistent with mutagenesis data that is relevant for matriglycan binding. Among other residues, His141 and Phe147 were found to be important for binding matriglycan as mutating these residues to alanine abrogated the interaction of the LASV spike with α-DG[10]. Our structure reveals that these two residues are directly located underneath the "RRLL" motif (Extended Data Fig. 9), effectively helping to stabilize the matriglycan binding site. Mutating these residues to alanine likely changes the conformation of the "RRLL" motif, making it ill-suited for matriglycan binding. These observations add to the important role in binding matriglycan that was previously assigned to Tyr150 in LCMV[11] and in LASV by mutating it to alanine[10]. Our structure reveals the important structural role of Tyr150 in forming the binding site for matriglycan (Fig 3b). Altogether, our structure helps to rationalize the functional consequences of many previously reported mutations, which in turn validates our structure.

Degenerate and alternative states for matriglycan binding. In addition to the multiple polar interactions and the extended buried surface area that drive the binding of matriglycan by the LASV spike complex, other factors may influence the binding strength. The interaction of the LASV spike complex, which is a 3-fold symmetric complex with a linear form of matriglycan that does not have any cyclic symmetry, gives rise to multiple degenerate states for binding (Extended Data Fig. 10a). Although these states are fully equivalent and indistinguishable, their existence is important as it effectively reduces the entropic cost for binding. While we have modeled only a single state in which one elongated matriglycan chain is bound to the spikes (Fig. 3c), additional states could easily be accommodated using the very same geometry. These include two or even three separate matriglycan chains that will bind to a single trimer at the same time (Extended Data Fig. 10b & 10c). Each such state will also have multiple degenerate forms that may lower the associated entropic cost for binding. If these matriglycan chains are not held together on a single surface, a single chain that will span all three binding sites within a single trimer will have the highest affinity due to avidity / extended interaction surface. Nevertheless, simultaneous binding of more than one chain may be important for establishing the initial interaction of the virus with the host cell and should also benefit from avidity if the matriglycan chains are anchored together to the same surface. A compelling idea is that the spike complex could also bind in the middle of a long matriglycan chain. A proposed geometry for such binding is shown in Extended Data Fig. 10d.

This proposed geometry is very similar to the observed geometry of the bound matriglycan (Fig. 3c), with the exception of a flip in direction at the exit point of the chain (Extended Data Fig. 10d). If such interaction mode can indeed form, it may allow the virions to slide along the long matriglycan chains to get close to the surface of the target cell while maintaining interaction with matriglycan. Such sliding may involve hopping from one site to another. Mechanisms of sliding along linear polymers were extensively studied in the context of other biological polymers[12].

Structural rule of the SKI-I cleavage site. A hallmark of class-I viral fusogens is the proteolytic cleavage between their N-terminal receptor binding module and the C-terminal transmembrane module. Many different viruses utilize furin that recognizes an Arg-X-Lys/Arg-Arg motif for this proteolytic event, like HIV-1, for example[13]. Besides the obvious benefit of using a highly abundant protease for achieving efficient maturation, acquiring a furin site enables viral spike proteins to be recognized by neuropilin-1 and 2 (NRP-1, NRP-2) according to the C-end rule[14]. Hence, following cleavage by furin, NRPs bind to the exposed terminal basic residues of the viral spikes and effectively serve as cellular attachment factors that increase the cell-entry efficiency of these viruses. This mechanism was recently demonstrated in the case of SARS-CoV-2[15]. To be effective though, the cleavage site needs to stay exposed in the mature spike. While this is true for SARS-CoV-2, it is not the case for LASV (Extended Data Fig. 11). The critical role of the SKI-I recognition site in mediating binding of matriglycan explains why a furin cleavage site cannot be accommodated by α-DG-tropic arenaviruses. For other arenaviruses like the TfR1-tropic viruses, the SKI-I recognition motif is no longer needed for binding matriglycan, but it is likely to have a structural role in stabilizing the spike as in the case of LASV (Fig. 2a).

Binding free matriglycan by the LASV spike complex. To load newly generated spike complexes with matriglycan there are two important prerequisites: a) The spike complexes need to be able to bind matriglycan, and b) matriglycan should be available in the vicinity of the spike complexes. From our structure, it is evident that cleavage by SKI-I is a critical precondition for forming the matriglycan binding site (Fig. 3b). SKI-I is localized to the Golgi[16] and the processing of the LASV GPC by SKI-I occurs at the ER/*cis*-Golgi[17]. The enzyme that produces matriglycan is LARGE1[18,19] and is also located at the Golgi[20]. However, LARGE1 is known to synthesize matriglycan on α-DG and not as a free polymer in solution. Interestingly, incubating the enzyme β1,4 glucuronyltransferase (B4GAT1)[21] with Xyl and UDP-GlcA *in-vitro*

produces a GlcAβ1-4Xyl that is recognized and elongated by LARGE1[22]. Hence, B4GAT1 that is found in the Golgi[21] could potentially produce free substrates that LARGE1 can subsequently elongate to form free polymers of matriglycan. We therefore postulate that following SKI-I-depended maturation of the LASV spike complex at the ER/*cis*-Golgi, it reaches a Golgi compartment where free matriglycan polymers are available, and that this is where loading occurs, prior to the shuttling of the newly-formed spike complexes to the cell surface.

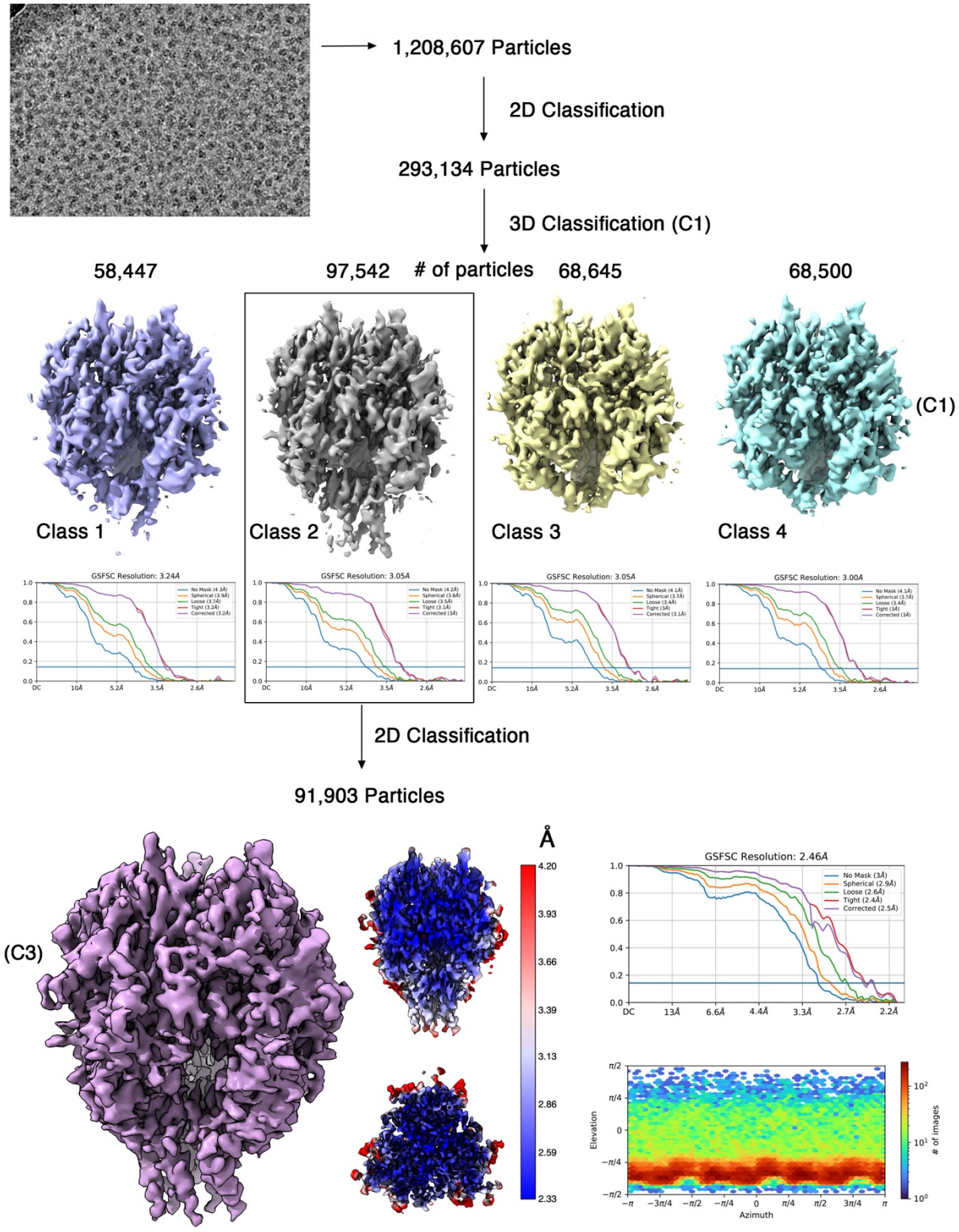
Extended Table 1

Model	LASV GPC in C3-symmetric map	
Composition (#)		
Chains	12	
Atoms	10395 (Hydrogens: 0)	
Residues	Protein: 1224 Nucleotide: 0	
Water	0	
Ligands	XYS: 6 NAG: 33 BDP: 3	
Bonds (RMSD)		
Length (Å) (# > 4σ)	0.002 (0)	
Angles (°) (# > 4σ)	0.615 (0)	
MolProbity score	2.04	
Clash score	13.87	
Ramachandran plot (%)		
Outliers	0.00	
Allowed	3.02	
Favored	96.98	
Rama-Z (Ramachandran plot Z-score, RMSD)		
whole (N = 1194)	-0.75 (0.22)	
helix (N = 489)	0.26 (0.23)	
sheet (N = 141)	-0.31 (0.43)	
loop (N = 564)	-1.18 (0.23)	
Rotamer outliers (%)	1.90	
Cβ outliers (%)	0.00	
Peptide plane (%)		
Cis proline/general	0.0/0.0	
Twisted proline/general	0.0/0.0	
CaBLAM outliers (%)	2.32	
ADP (B-factors)		
Iso/Aniso (#)	10395/0	
min/max/mean		
Protein	23.90/166.12/48.45	
Nucleotide	---	
Ligand	36.27/60.09/47.46	
Water	---	
Occupancy		
Mean	1.00	
occ = 1 (%)	100.00	
0 < occ < 1 (%)	0.00	
occ > 1 (%)	0.00	
Data		
Box		
Lengths (Å)	86.15, 92.38, 123.52	
Angles (°)	90.00, 90.00, 90.00	
Supplied Resolution (Å)	3.3	
Resolution Estimates (Å)	Masked	Unmasked
d FSC (half maps; 0.143)	---	---
d 99 (full/half1/half2)	3.4/---/---	3.4/---/---
d model	3.3	3.3
d FSC model (0/0.143/0.5)	3.2/3.2/3.4	3.2/3.2/3.4
Map min/max/mean	-0.35/0.78/0.02	
Model vs. Data		
CC (mask)	0.68	
CC (box)	0.70	
CC (peaks)	0.60	
CC (volume)	0.68	
Mean CC for ligands	0.63	

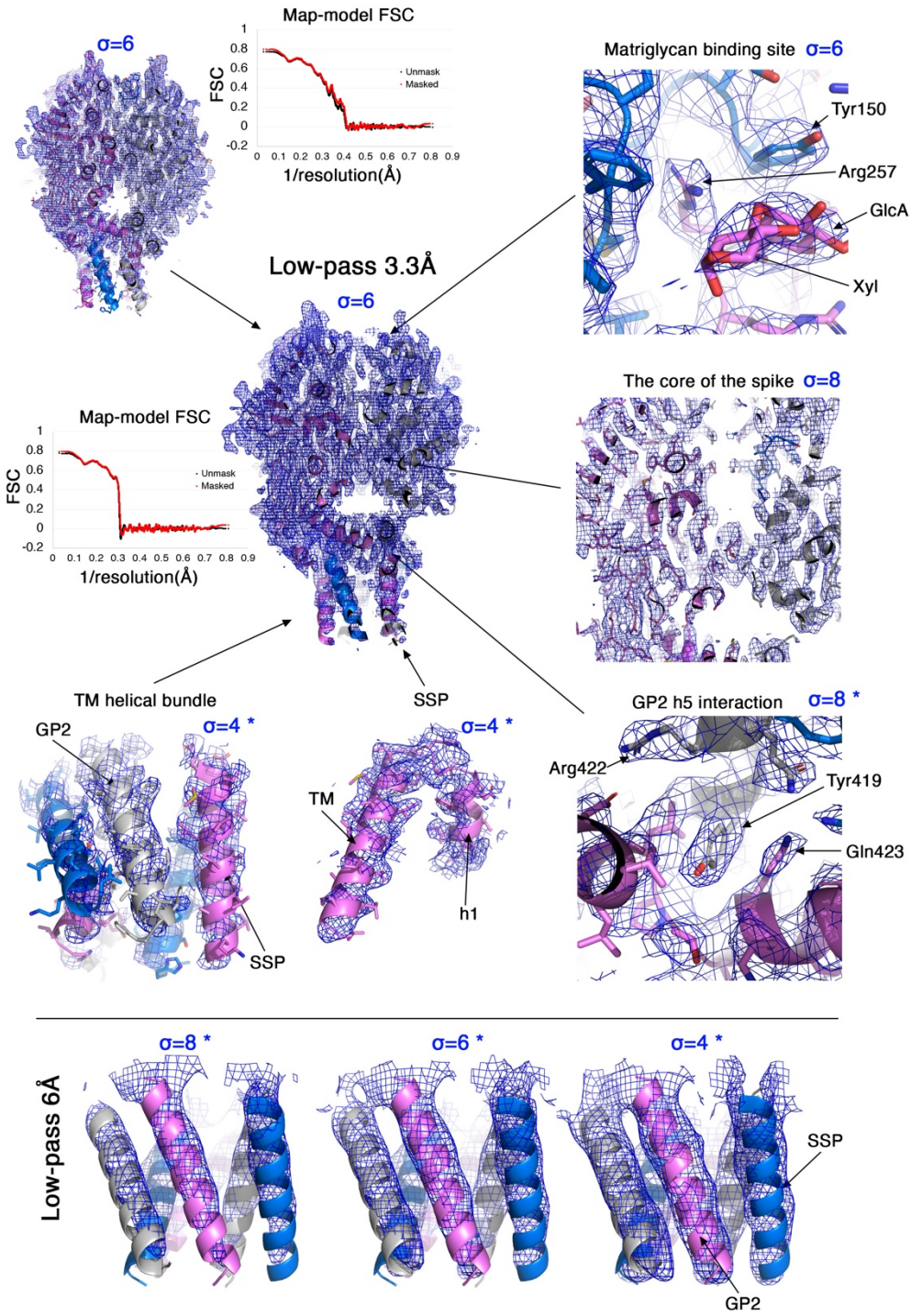
Extended Table 2

Model	LASV GPC in focused C1-symmetric map	
Composition (#)		
Chains	10	
Atoms	8902 (Hydrogens: 0)	
Residues	Protein: 1035 Nucleotide: 0	
Water	0	
Ligands	XYS: 7	
	NAG: 30	
	BDP: 6	
Bonds (RMSD)		
Length (Å) (# > 4σ)	0.004 (0)	
Angles (°) (# > 4σ)	0.668 (0)	
MolProbity score	1.89	
Clash score	13.11	
Ramachandran plot (%)		
Outliers	0.00	
Allowed	3.86	
Favored	96.14	
Rama-Z (Ramachandran plot Z-score, RMSD)		
whole (N = 1011)	-1.16 (0.24)	
helix (N = 351)	-0.36 (0.27)	
sheet (N = 150)	-0.45 (0.42)	
loop (N = 510)	-1.09 (0.24)	
Rotamer outliers (%)	0.32	
Cβ outliers (%)	0.00	
Peptide plane (%)		
Cis proline/general	0.0/0.0	
Twisted proline/general	0.0/0.0	
CaBLAM outliers (%)	2.43	
ADP (B-factors)		
Iso/Aniso (#)	8902/0	
min/max/mean		
Protein	25.02/94.34/48.00	
Nucleotide	---	
Ligand	34.66/72.16/54.77	
Water	---	
Occupancy		
Mean	1.00	
occ = 1 (%)	100.00	
0 < occ < 1 (%)	0.00	
occ > 1 (%)	0.00	
Data		
Box		
Lengths (Å)	86.15, 92.38, 96.53	
Angles (°)	90.00, 90.00, 90.00	
Supplied Resolution (Å)	3.7	
Resolution Estimates (Å)	Masked	Unmasked
d FSC (half maps; 0.143)	---	---
d 99 (full/half1/half2)	3.8/---/---	3.8/---/---
d model	3.7	3.7
d FSC model (0/0.143/0.5)	3.5/3.6/3.7	3.6/3.6/3.8
Map min/max/mean	-0.38/0.74/0.02	
Model vs. Data		
CC (mask)	0.67	
CC (box)	0.69	
CC (peaks)	0.58	
CC (volume)	0.65	
Mean CC for ligands	0.63	

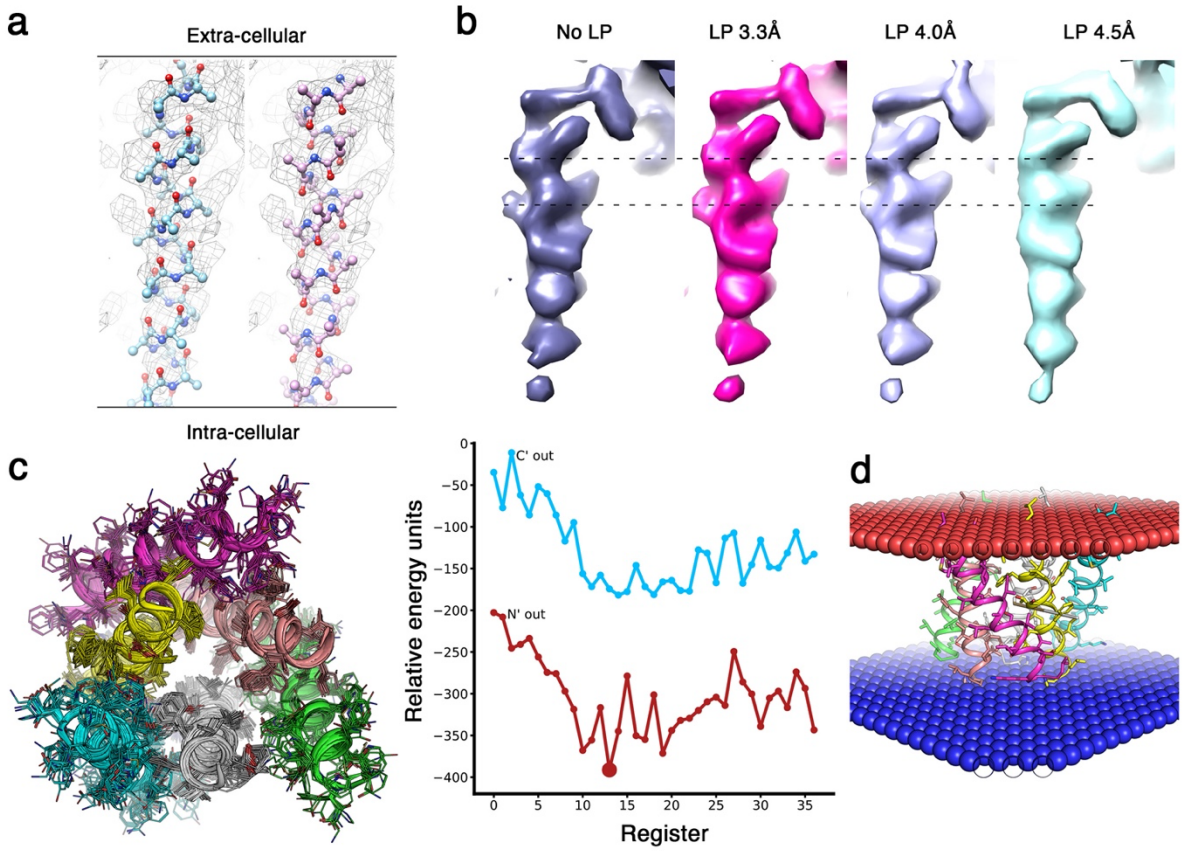
Extended Figures:



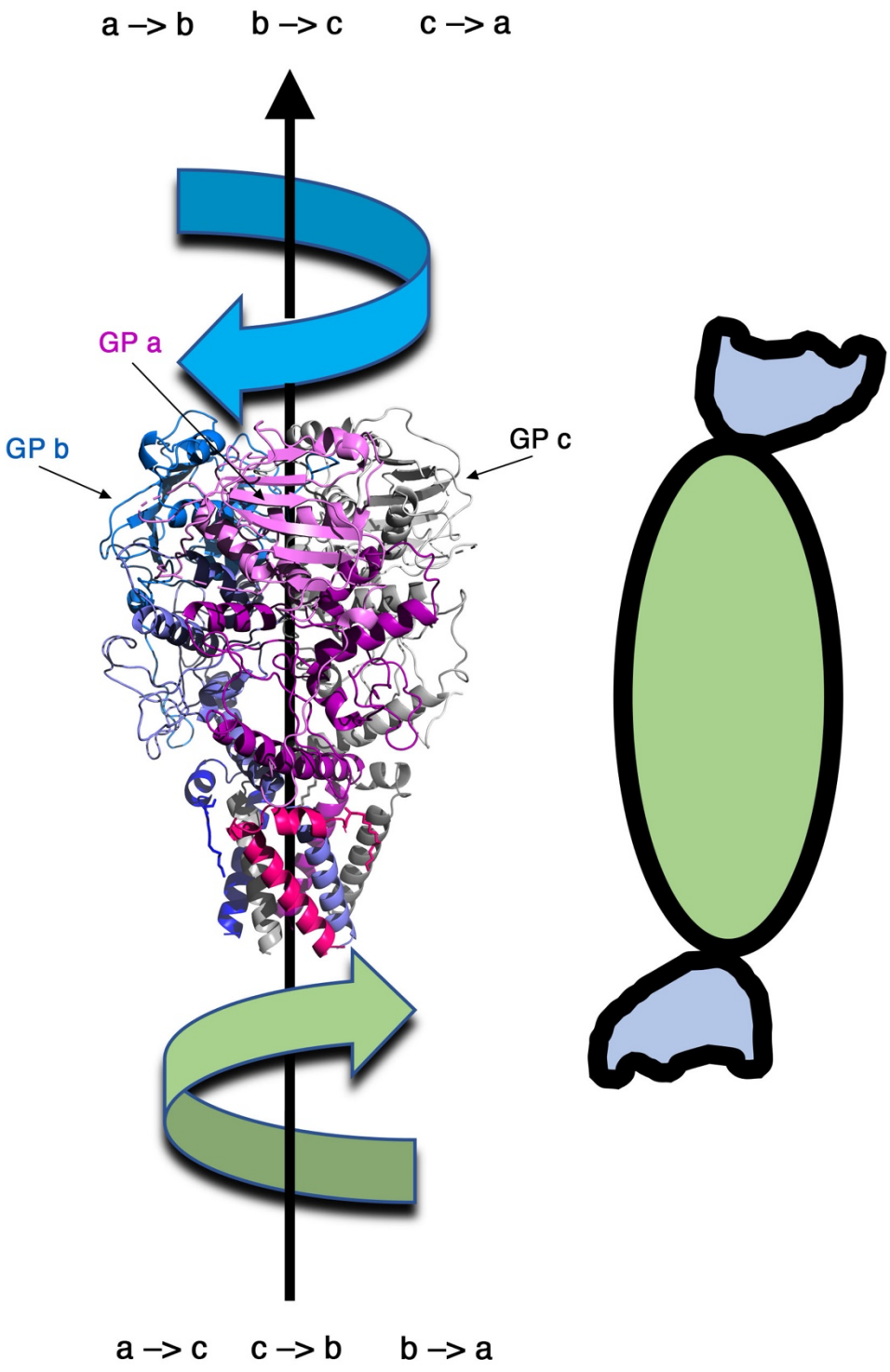
Extended data Fig. 1 | Reconstruction of density map. The course of data processing and reconstruction of a C3-symmetric density map is visually summarized. Local resolution estimates, gold-standard FSC curve and orientational distribution are shown for the final map.



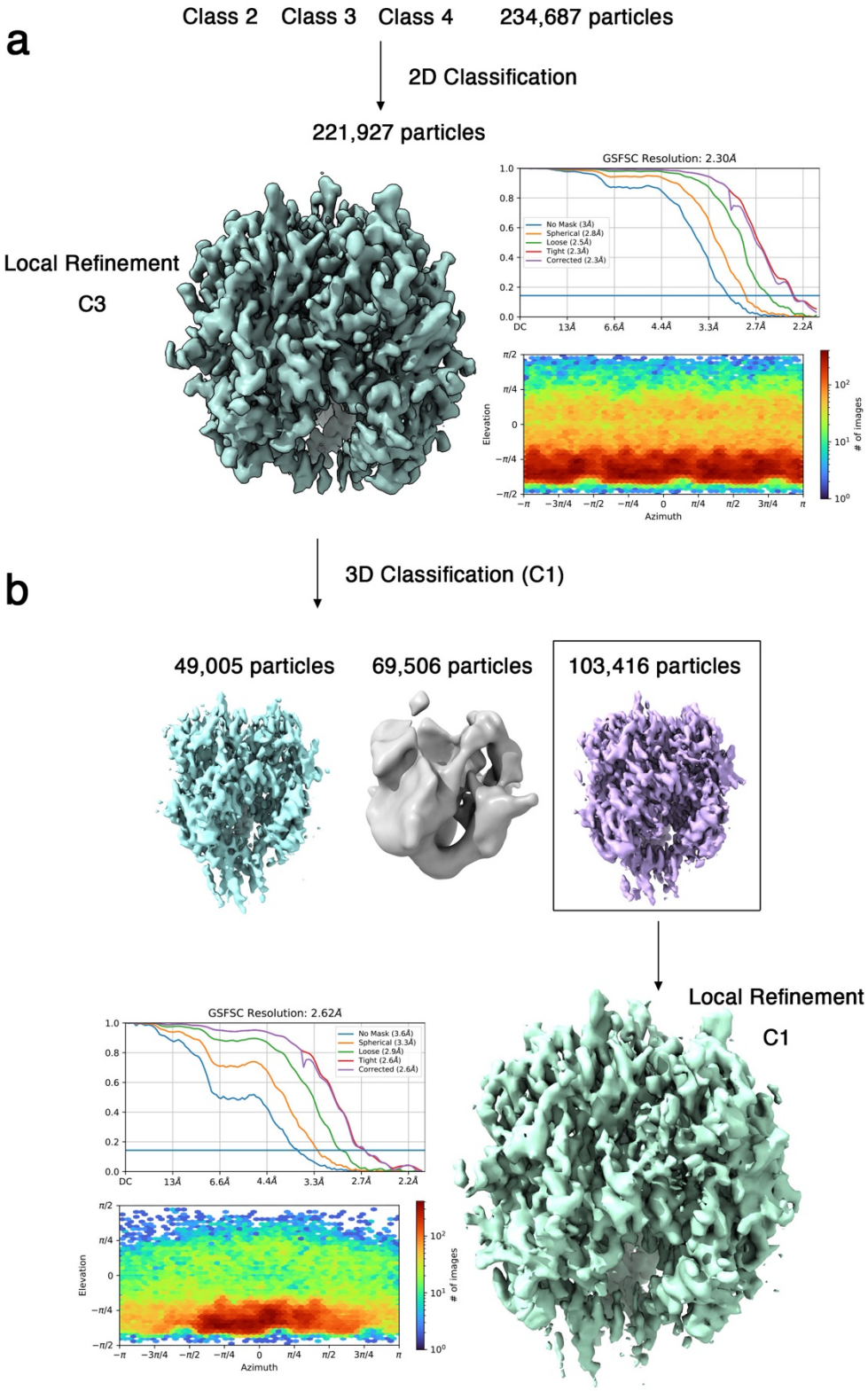
Extended data Fig. 2 | Overall fit of the model to the density map. Model is shown as a ribbon with or without stick-represented side chains. Density map is shown using a blue mesh at the indicated sigma levels. Asterisks indicate that the density map was carved around the model. Upper left side, the original model was fitted into a working map at GSFSC=2.5 Å and achieved map-model FSC of 3.3 Å. Center, final refined model at a 3.3 Å low-passed filtered map with a map-model FSC of 3.3 Å. Around the central model, close-up views of regions of interest as indicated at the 3.3 Å low-passed filtered map. Key residues or structural elements are labeled. Bottom part, the six-helical transmembrane bundle is shown in a 6 Å low-pass filtered map at the indicated three different sigma levels.



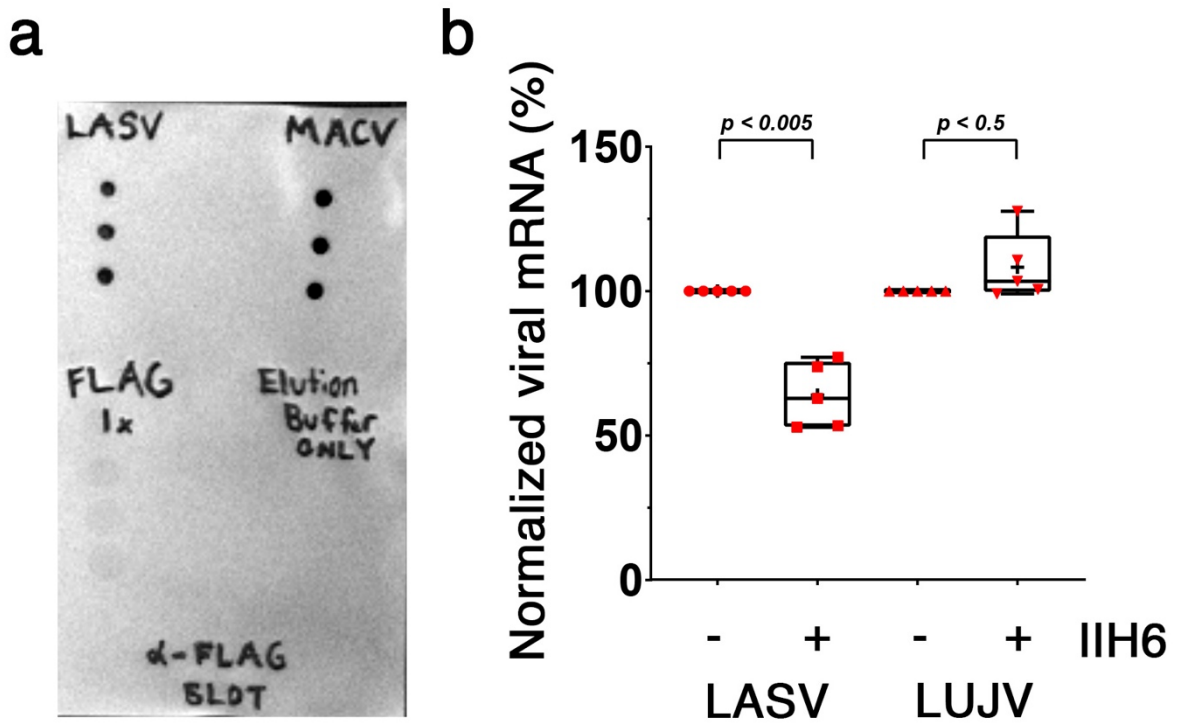
Extended data Fig. 3 | Assigning the overall directionality and registry of the SSP. a. The amino-terminus of the SSP's trans-membrane α-helix is facing extracellularly. The density map of the trans-membrane α-helix is shown as a mesh, and two α-helices are rendered as poly-ala; one has a carboxy-terminus facing up (left) and the second has an amino-terminus facing up (right). Densities for the side-chains are pointing toward the extracellular direction (up), indicating that the correct directionality of this α-helix in the membrane is with its amino-terminus at the extracellular region (right). **b.** Side-chain densities are visible in low-pass filtered maps. Densities for the α-helix of SSP are shown with maps that were low-pass (LP) filtered to the indicated resolution values. Dashed horizontal lines help to visualize the direction to which the side-chains are pointing. The directionality of the side-chains is apparent even at 4.5 Å LP-filtered map. **c.** Determining the registry of the SSP. On the left, superimposition of the 37 best-scoring computer-generated models for each threading (N-terminus out, only), shown in a 'top' view. The inner three α-helices (yellow, brown, and gray) belong to the GP2 subunits and the outer three α-helices (magenta, cyan, and green) belong to the SSP subunits. On the right, energy profiles (relative Rosetta energy units) of the best-scoring computed model for each threading option as a function of the register, starting with the indicated residue at the amino-terminus of the trans-membrane α-helix (first methionine is considered as register "0"). The light-blue curve shows threading with the C-terminus out and the red curve shows threading with the N-terminus out. Large red dot indicates the lowest energy solution. **d.** The top-ranked model for the SSP registry.



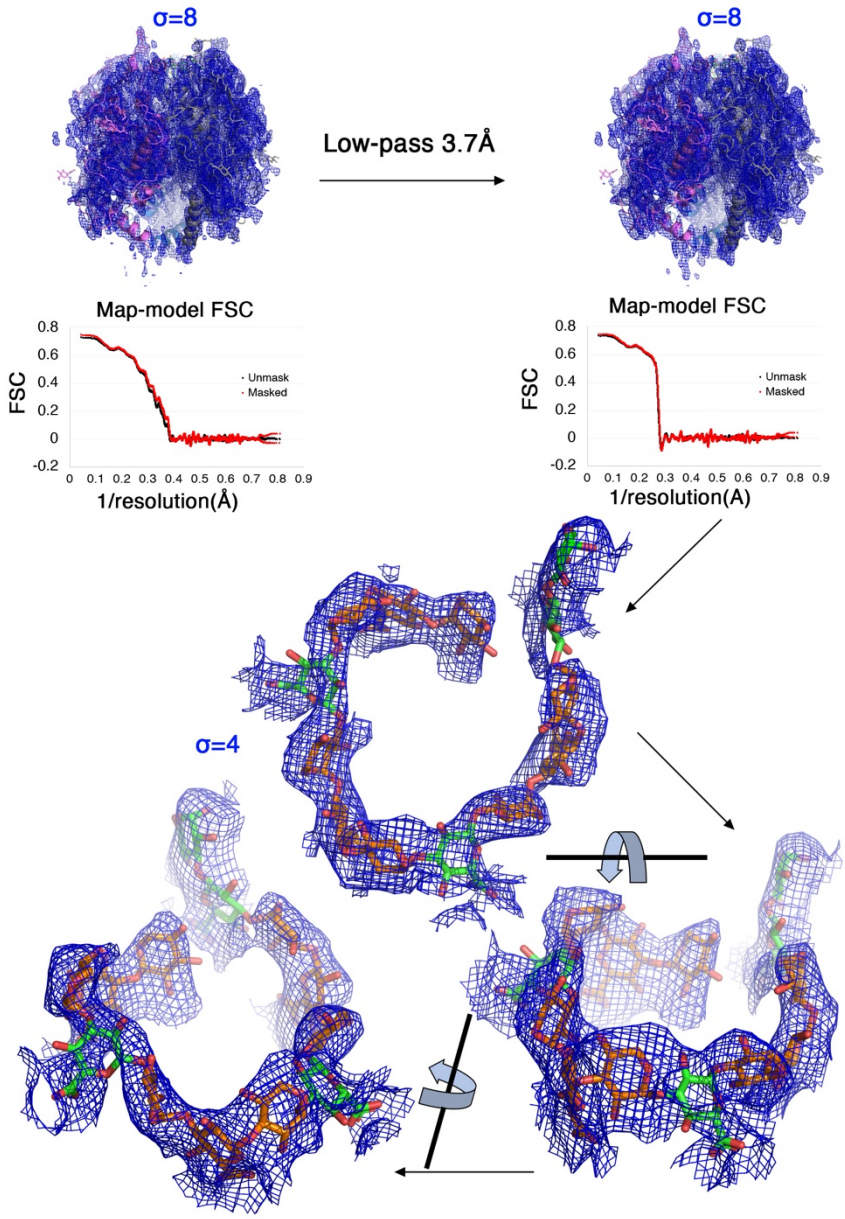
Extended data Fig. 4 | Domain swapping stabilization of the spike in a double-sided candy twist-wrap fashion. The directionality of the domain swapping in the trimer is illustrated. The domain swapping at the apex of the trimer (top) and at the membrane proximal region (bottom) have opposing directionalities, which is reminiscence to a candy's double-sided twist wrap (schematically shown on the right). Donation of a structural element from one subunit to another is marked by '–>'.



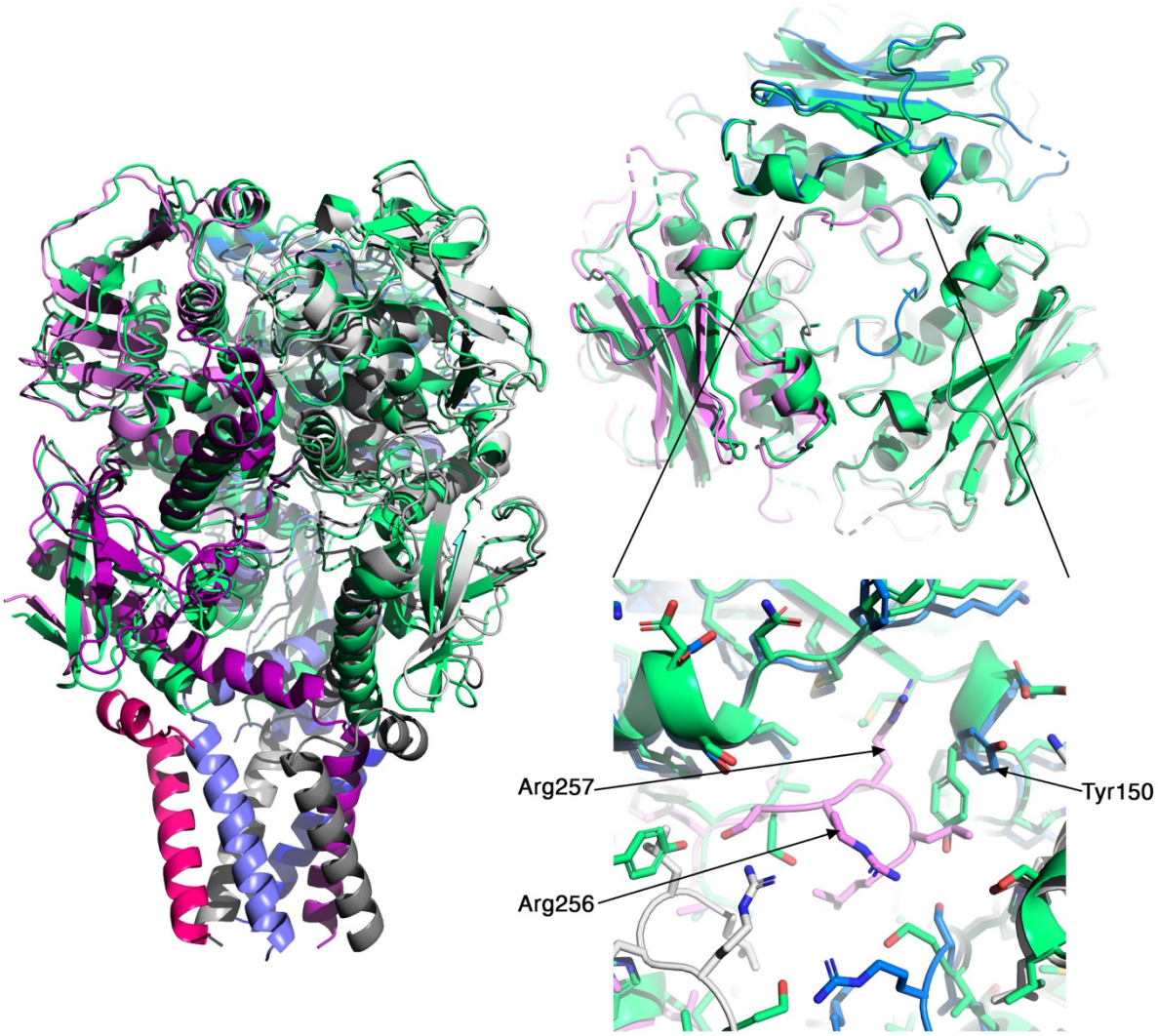
Extended data Fig. 5 | Reconstruction of focused density maps. The process of reconstructing focused density maps is visually illustrated.



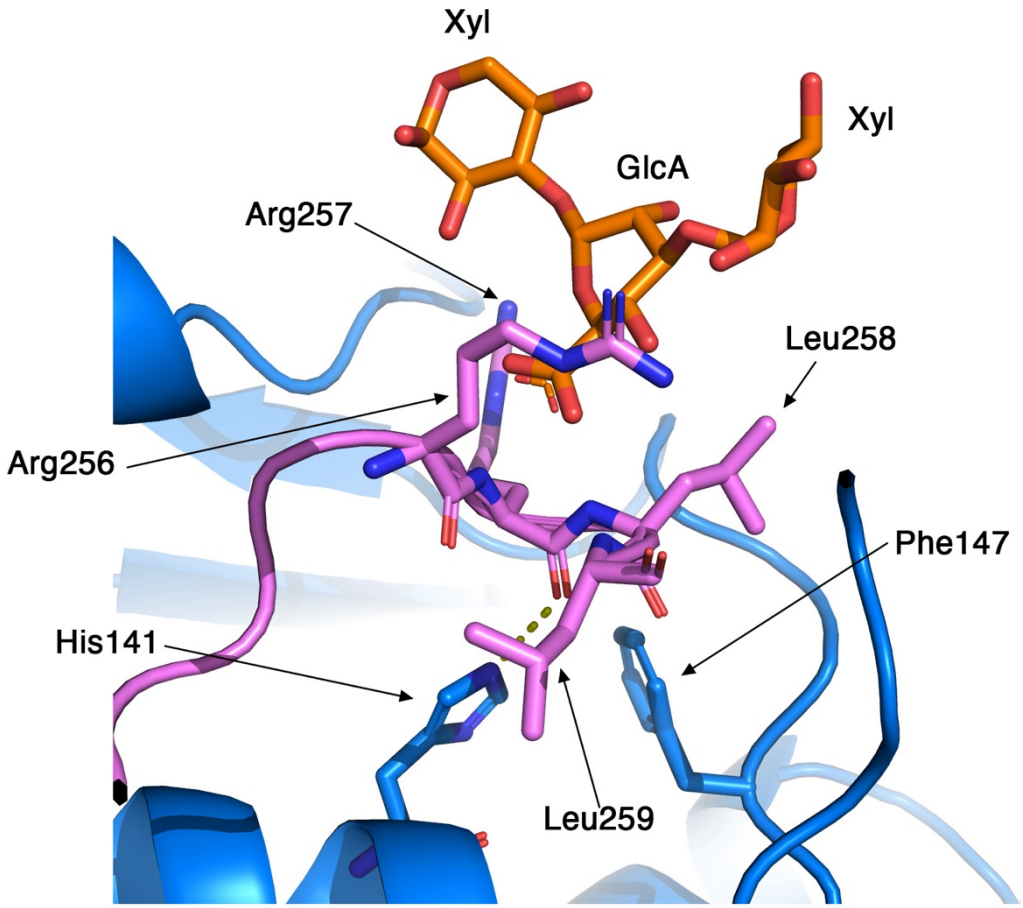
Extended data Fig. 6 | LASV spikes bind matriglycan also on pseudotyped viruses. a. Free Flag peptide is not recognized by anti-Flag antibody in a dot-blot settings. To test if the anti-Flag antibody could recognize the Flag peptide that was used to elute the spike complexes, we spotted on a nitrocellulose membrane the LASV and MACV spike complexes as well as the elution buffer alone, or the elution buffer with the Flag peptide (FLAG 1x). The elution buffer with the free Flag peptide produces only slight background, indicating that the Flag peptide is not adhering to the membrane, and hence do not contribute to the signal of the anti-Flag antibody. **b**. LASV pseudotyped viruses are depleted from solution by IIH6 conjugated to beads. To test if LASV spike complex are loaded with matriglycan in the context of pseudotyped viruses, we incubated MLV-based pseudotyped viruses bearing the spikes of LASV or of the neuropilin-2-tropic Lujo virus (LUJV) with protein-L beads alone, or after coating the beads with the IIH6 antibody. The amount of RNA encoding a luciferase reporter gene was then quantified in the supernatants of each sample using RT-qPCR. Each dot represents an average normalized value of several technical replicates in a single independent experiment (five experiments in total). The measurements in each experiment were normalized to the respective uncoated-beads control. Pseudotyped viruses with the LASV spike are significantly (two tailed Student's t-test) depleted from solution by IIH6 on beads. Whiskers indicate the min and max values, central line indicates the mean value, and the box indicates the interquartile range. The slight increase in the amount of LUJV results from minor unspecific absorption of viral particles to uncoated protein-L beads, which lowers the amount of particle in the reference control.



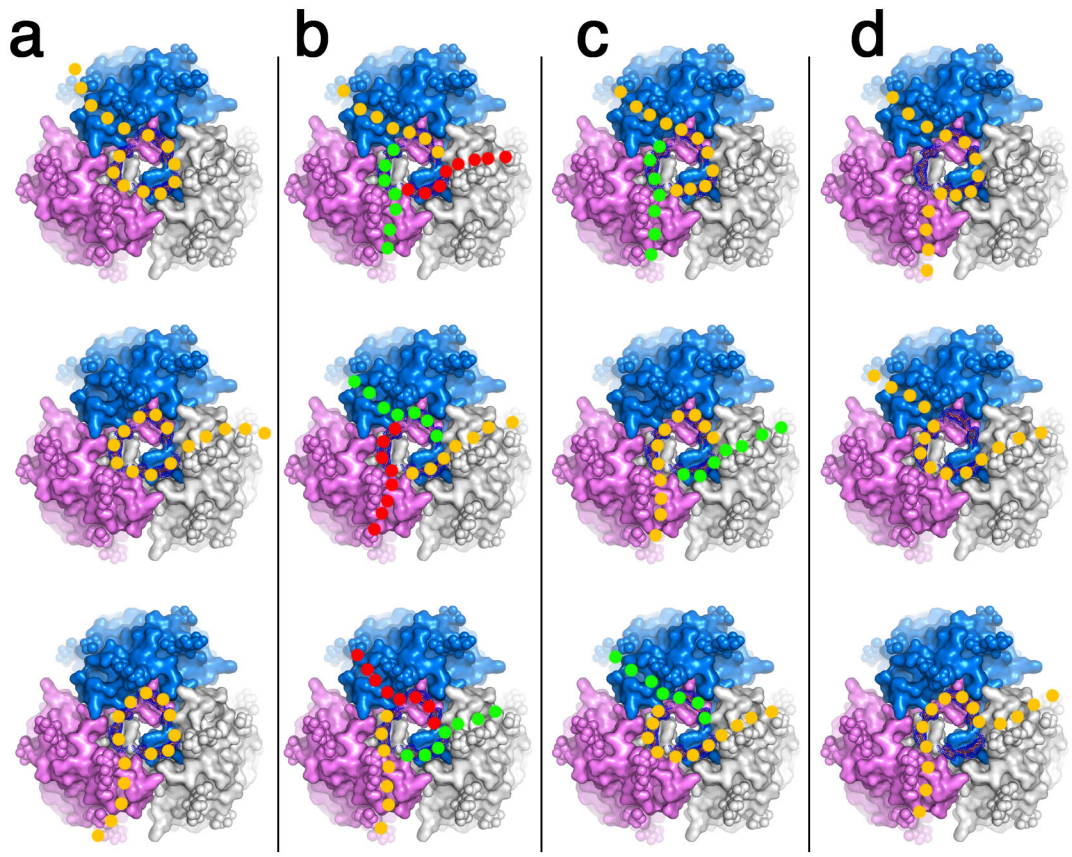
Extended data Fig. 7 | A 13-mer matriglycan in density map. The ectodomain portion of the spike was fitted into the C1-focused map and yielded a map-model FSC of 3.7 Å. The map was low-passed filtered to 3.7 Å and the model was refined again using this map. The linear matriglycan chain is shown in a top and side views. The density map is shown as a blue mesh at σ=4, carved at 2.5 Å around the matriglycan. The Xyl-GlcA-Xyl moieties in the symmetric sites are in orange. The GlcA moieties in the asymmetric sites as well as the termini Xyl-GlcA are shown in green.



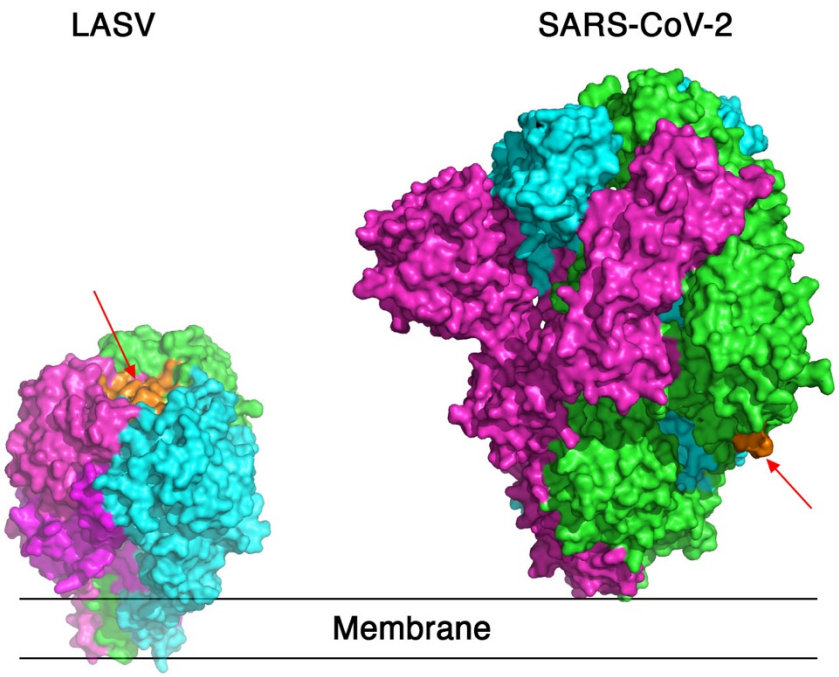
Extended data Fig. 8 | Comparison of the cryo-EM structure of LASV spike with the crystal structure of its ectodomain. The cryo-EM model from this study (similar color scheme as in Fig. 1) is show with the crystal structure of the spike's ectodomain (PDB: 5VK2, green). The two structures are shown as ribbons from a side view (left) and from a top view (top right). Overall, the structures are similar (RMSD=2.04 Å, for 1009 shared CA atoms). Key differences are near the membrane spanning domain that is missing from the crystal structure and at the receptor binding site that is not formed (enlarged view, bottom right), following the replacement of the "RRLL"-SKI-I recognition site with a furin recognition sequence. Tyr150, Arg256, and Arg257 that participate in binding of matriglycan are labeled.



Extended data Fig. 9 | Residues that affect matriglycan binding. His141 and Phe147 that were found to be important for matriglycan binding via mutagenesis studies are located directly below the "RRLL" motif that binds matriglycan. The conformation of the "RRLL" motif is partially shaped by the interactions it makes with His141 and Phe147.



Extended data Fig. 10 | Degenerate and alternative states for matriglycan binding. A top view of the LASV trimeric spike showing a surface representation of the three GP1 subunits in pink, blue, and gray. Colored dots represent the potential traces of bound matriglycan chains. Different colors represent different individual chains of matriglycan.



Extended data Fig. 11 | cleavage site between the receptor binding module and the transmembrane module in different viruses. The trimeric spike complexes of LASV (left) and of SARS-CoV-2 (right, PDB: 6VXX), are shown in relative scale in surface representations and in three different colors to their three protomers. The protease cleavage sites (i.e., SKI-I site for LASV and furin site for SARS-CoV-2) or their immediate vicinities are colored in orange and are pointed with red arrows.

Extended methods:

Command line for Rosetta modeling

rosetta_scripts_executable -database path_to_rosetta_database -parser:protocol thread_and_score.xml -out:suffix _window_sequence -s template_pdb -parser:script_vars cst_full_path=path_to_Ca_coordint_constraints start_position_A=3A start_position_B=3B start_position_C=3C seq_to_thread=windown_sequence span_starts=3A,3B,3C,428a,428b,428c span_ends=24A,24B,24C,448a,448b, 448c span_oris=in2out,in2out,in2out,out2in,out2in,out2in -nstruct 10

Where thread_and_score.xml is the RosettaScripts XML at the bottom; window sequence is the sequence to model; 3A,3B, 3C and 24A, 24B and 24C are the start and end positions of the TM to model in the PDB, respectively; 428a,428b and 428c are the start positions for the GP TM. and the span_oris parameter are the orientations of all six spans with respect to the membrane.

```
<ROSETTASCRIPTS>
 <SCOREFXNS>
  <ScoreFunction name="full" weights="ref2015_memb">
   <Reweight scoretype="coordinate_constraint" weight="0.64"/>
     <Reweight scoretype="res_type_constraint" weight="0.5"/>
  </ScoreFunction>
  <ScoreFunction name="soft" weights="ref2015_soft">
   <Reweight scoretype="mp_res_lipo" weight="1"/>
     <Reweight scoretype="coordinate_constraint" weight="0.64"/>
   <Reweight scoretype="res_type_constraint" weight="0.5"/>
  </ScoreFunction>
 </SCOREFXNS>
 <TASKOPERATIONS>
  <InitializeFromCommandline name="init"/>
  <RestrictToRepacking name="rtr"/>
 </TASKOPERATIONS>
 <MOVERS>
  <SimpleThreadingMover name="thread_A" start_position="%%start_position_A%%"
thread_sequence="%%seq_to_thread%%" scorefxn="full"/>
  <SimpleThreadingMover name="thread_B" start_position="%%start_position_B%%"
thread_sequence="%%seq_to_thread%%" scorefxn="full"/>
  <SimpleThreadingMover name="thread_C" start_position="%%start_position_C%%"
thread_sequence="%%seq_to_thread%%" scorefxn="full"/>
  <AddMembraneMover name="add_memb" membrane_core="10" steepness="4"
span_starts_num="%%span_starts%%" span_ends_num="%%span_ends%%"
span_orientations="%%span_oris%%"/>
  <PackRotamersMover name="soft_repack" scorefxn="soft" task_operations="init,rtr"/>
  <PackRotamersMover name="hard_repack" scorefxn="full" task_operations="init,rtr"/>
  <RotamerTrialsMinMover name="RTmin" scorefxn="full" task_operations="init,rtr"/>
  <MinMover name="soft_min" scorefxn="soft" chi="1" bb="1" jump="0"/>
  <MinMover name="hard_min" scorefxn="full" chi="1" bb="1" jump="0"/>
```

```
  <ConstraintSetMover name="add_CA_cst" cst_file="%%cst_full_path%%"/>
  <ParsedProtocol name="refinement_block"> #10 movers
    <Add mover_name="soft_repack"/>
    <Add mover_name="soft_min"/>
    <Add mover_name="soft_repack"/>
    <Add mover_name="hard_min"/>
    <Add mover_name="hard_repack"/>
    <Add mover_name="hard_min"/>
    <Add mover_name="hard_repack"/>
    <Add mover_name="RTmin"/>
    <Add mover_name="RTmin"/>
    <Add mover_name="hard_min"/>
  </ParsedProtocol>
  <LoopOver name="iter4" mover_name="refinement_block" iterations="4"/>
 </MOVERS>
 <FILTERS>
  <Time name="timer"/>
 </FILTERS>
 <PROTOCOLS>
  <Add filter="timer"/>
  <Add mover="add_memb"/>
  <Add mover="add_CA_cst"/>
  <Add mover="thread_A"/>
  <Add mover="thread_B"/>
  <Add mover="thread_C"/>
  <Add mover="iter4"/>
  <Add filter="timer"/>
 </PROTOCOLS>
 <OUTPUT scorefxn="full"/>
</ROSETTASCRIPTS>
```

References

1 Agnihothram, S. S., York, J., Trahey, M. & Nunberg, J. H. Bitopic membrane topology of the stable signal peptide in the tripartite Junin virus GP-C envelope glycoprotein complex. *J Virol* **81**, 4331-4337, (2007).

2 Eichler, R. *et al.* Lassa virus glycoprotein signal peptide displays a novel topology with an extended endoplasmic reticulum luminal region. *The Journal of biological chemistry* **279**, 12293-12299, (2004).

3 Froeschke, M., Basler, M., Groettrup, M. & Dobberstein, B. Long-lived signal peptide of lymphocytic choriomeningitis virus glycoprotein pGP-C. *The Journal of biological chemistry* **278**, 41914-41920, (2003).

4 Schrempf, S., Froeschke, M., Giroglou, T., von Laer, D. & Dobberstein, B. Signal peptide requirements for lymphocytic choriomeningitis virus glycoprotein C maturation and virus infectivity. *J Virol* **81**, 12515-12524, (2007).

5 Briknarova, K., Thomas, C. J., York, J. & Nunberg, J. H. Structure of a zinc-binding domain in the Junin virus envelope glycoprotein. *The Journal of biological chemistry* **286**, 1528-1536, (2011).

6 Messina, E. L., York, J. & Nunberg, J. H. Dissection of the role of the stable signal peptide of the arenavirus envelope glycoprotein in membrane fusion. *J Virol* **86**, 6138-6145, (2012).

7 Heijne, G. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J* **5**, 3021-3027, (1986).

8 Saunders, A. A. *et al.* Mapping the landscape of the lymphocytic choriomeningitis virus stable signal peptide reveals novel functional domains. *J Virol* **81**, 5649-5657, (2007).

9 Eichler, R. *et al.* Identification of Lassa virus glycoprotein signal peptide as a trans-acting maturation factor. *EMBO reports* **4**, 1084-1088, (2003).

10 Acciani, M. *et al.* Mutational Analysis of Lassa Virus Glycoprotein Highlights Regions Required for Alpha-Dystroglycan Utilization. *J Virol* **91**, (2017).

11 Hastie, K. M. *et al.* Crystal structure of the prefusion surface glycoprotein of the prototypic arenavirus LCMV. *Nat Struct Mol Biol* **23**, 513-521, (2016).

12 Bigman, L. S., Greenblatt, H. M. & Levy, Y. What Are the Molecular Requirements for Protein Sliding along DNA? *J Phys Chem B* **125**, 3119-3131, (2021).

13 Hallenberger, S. *et al.* Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature* **360**, 358-361, (1992).

14 Teesalu, T., Sugahara, K. N., Kotamraju, V. R. & Ruoslahti, E. C-end rule peptides mediate neuropilin-1-dependent cell, vascular, and tissue penetration. *Proc Natl Acad Sci U S A* **106**, 16157-16162, (2009).

15 Cantuti-Castelvetri, L. *et al.* Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **370**, 856-860, (2020).

16 Pullikotil, P., Benjannet, S., Mayne, J. & Seidah, N. G. The proprotein convertase SKI-1/S1P: alternate translation and subcellular localization. *The Journal of biological chemistry* **282**, 27402-27413, (2007).

17 Burri, D. J. *et al.* Molecular characterization of the processing of arenavirus envelope glycoprotein precursors by subtilisin kexin isozyme-1/site-1 protease. *J Virol* **86**, 4935-4946, (2012).
18 Inamori, K. *et al.* Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. *Science* **335**, 93-96, (2012).
19 Yoshida-Moriguchi, T. & Campbell, K. P. Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* **25**, 702-713, (2015).
20 Grewal, P. K., McLaughlan, J. M., Moore, C. J., Browning, C. A. & Hewitt, J. E. Characterization of the LARGE family of putative glycosyltransferases associated with dystroglycanopathies. *Glycobiology* **15**, 912-923, (2005).
21 Willer, T. *et al.* The glucuronyltransferase B4GAT1 is required for initiation of LARGE-mediated alpha-dystroglycan functional glycosylation. *Elife* **3**, (2014).
22 Osman Sheikh, M. *et al.* Cell Surface Glycan Engineering Reveals that Matriglycan Alone can Recapitulate Dystroglycan Binding and Function. *bioRxiv*, 2021.2005.2010.443358, (2021).