Supplementary Information for

Novel chimeric proteins mimicking SARS-CoV-2 spike epitopes with broad inhibitory activity.

Mario Cano-Muñoz[1], Daniel Polo-Megías[1], Ana Cámara-Artigas[2], José A. Gavira[3], María J. López-Rodríguez[1], Géraldine Laumond[4], Sylvie Schmidt[4], Julien Demiselle[5], Seiamak Bahram[4], Christiane Moog[4,6,*] and Francisco Conejero-Lara[1,*]

[1]Departamento de Química Física, Instituto de Biotecnología y Unidad de Excelencia de Química Aplicada a Biomedicina y Medioambiente (UEQ), Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain.

[2]Department of Chemistry and Physics, Agrifood Campus of International Excellence (ceiA3) and CIAMBITAL, University of Almeria, Carretera de Sacramento s/n, 04120 Almeria, Spain.

[3]Laboratory of Crystallographic Studies, IACT, (CSIC-UGR), Avenida de las Palmeras 4, 18100 Armilla, Granada, Spain.

[4]Laboratoire d'ImmunoRhumatologie Moléculaire, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR_S 1109, Institut thématique interdisciplinaire (ITI) de Médecine de Précision de Strasbourg, Transplantex NG, Faculté de Médecine, Fédération Hospitalo-Universitaire OMICARE, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg, Strasbourg, France.

[5]UMR_S1260, NanoRegMed, University of Strasbourg, Strasbourg, France.

[6]Vaccine Research Institute (VRI), Créteil, France.

***Corresponding authors:**

F. Conejero-Lara; conejero@ugr.es

C. Moog; c.moog@unistra.fr

This PDF file includes:

Tables S1 and S2

Figures S1 to S7

Table S1. Amino acid sequences of the reference HR1, HR2 of the SARS-CoV-2 spike (Uniprot P0DTC2) and the CoVS-HR1 variants.

	Amino acid sequences [b]
HR1	NVLYENQKLI ANQFNSAIGK IQDSLSSTAS ALGKLQDVVN QNAQALNTLV KQLSSNFGAI SSVLNDILSR LDKVE
HR2 (V39E) [a]	VDLGDISGIN ASVVNIQKEI DRLNEVAKNL NESLIDLQE
L3A [c]	**D**VLYENQKLI AN**E**FNSAIGK IQDSLSSTAS ALGKLQD**K**VN QNAQ**K**LNTLV KQLSSNFGAI SS**K**LNDILSR LDKGEPAKDL RSLIDNL**K**S**E** IAGFNSSLQK VLTNL**R**Q**E**NQ NV**E**D**E**LKGLA SATSSL**E**DQI KG**E**ASNFQN**R** ILKQNEYLVN SGSGNVLYEN QKLIANQFNS A**E**GKIQDSLS ST**K**SALGKLQ DVVNQN**K**QAL NTL**E**KQLSSN F**R**AISSVLND I**E**SRL**Q**KV**Q**
L3B	**D**VLYENQKLI AN**K**FNSAIGK IQDSLSSTAS **E**LGKLQD**E**VN QNAQ**D**LNTLV KQLSSNFG**R**I SS**E**LNDILSR LDKGEPAKDL RS**D**IDNL**E**S**K** IAGFNSSLQK VLTNLAQ**K**NQ NV**E**D**K**LKGL**E** S**R**TSSL**EK**QI KGIASNFQN**E** ILKQ**R**EYLVN KGSGNVLYEN QKLI**E**NQFNS AIGKIQDSLS ST**K**SALGKL**K** DVVNQN**K**QAL NTLVKQLSSN FGAISSVLND I**K**SRLDKVE
L3C	**D**VLYENQKLI AN**K**FNSAI**K**K IQDSLSSTAS **E**L**K**KLQD**E**VN QNAQ**D**LNTLV KQLSSNF**KR**I SS**E**LNDILSR LDKGEPAKDL RS**D**IDNL**E**S**K** IA**K**FNSSLQK VLTNLAQ**K**NQ NV**E**D**K**LK**T**L**E** S**R**TSSL**EK**QI K**K**IASNFQN**E** ILKQ**R**EYLVN KGSGNVLYEN QKLI**E**NQFNS AI**K**KIQDSLS ST**K**SAL**K**KL**K** DVVNQN**K**QAL NTLVKQLSSN F**S**AISSVLND I**K**SRLDKVE

[a] The V39E peptide contained a C-terminal SGGY tag and was N-acetylated and C-amydated.

[b] The substituted amino acids from the reference sequence are highlighted in bold.

[c] The three protein variants contained a N-terminal Methionine and a C-terminal Histidine-tag of sequence GGGGSHHHHHH.

Table S2. Data collection and refinement statistics

	PDB entry 7ZR2
Data collection	
Space group	C 1 2 1
Cell dimensions	
a, b, c (Å)	166.78 37.37 43.91
a, b, g (°)	90 95.67 90
Resolution (Å)	18.92 - 1.45 (1.48 - 1.45)
Rmerge	0.059 (0.720)
I / sI	9.4 (1.2)
Completeness (%)	96.3 (75.4)
Redundancy	3.0 (2.4)
Refinement	
Resolution (Å)	18.92- 1.45
No. reflections	46198 (3824)
Rwork / Rfree	0.178 (0.281)/ 0.213 (0.315)
No. atoms	
Protein	2023
Ligand/ion	3
Water	304
B-factors	
Protein	22.45
Ligand/ion	26.48
Water	31.40
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.70

***Values in parentheses are for highest-resolution shell.**

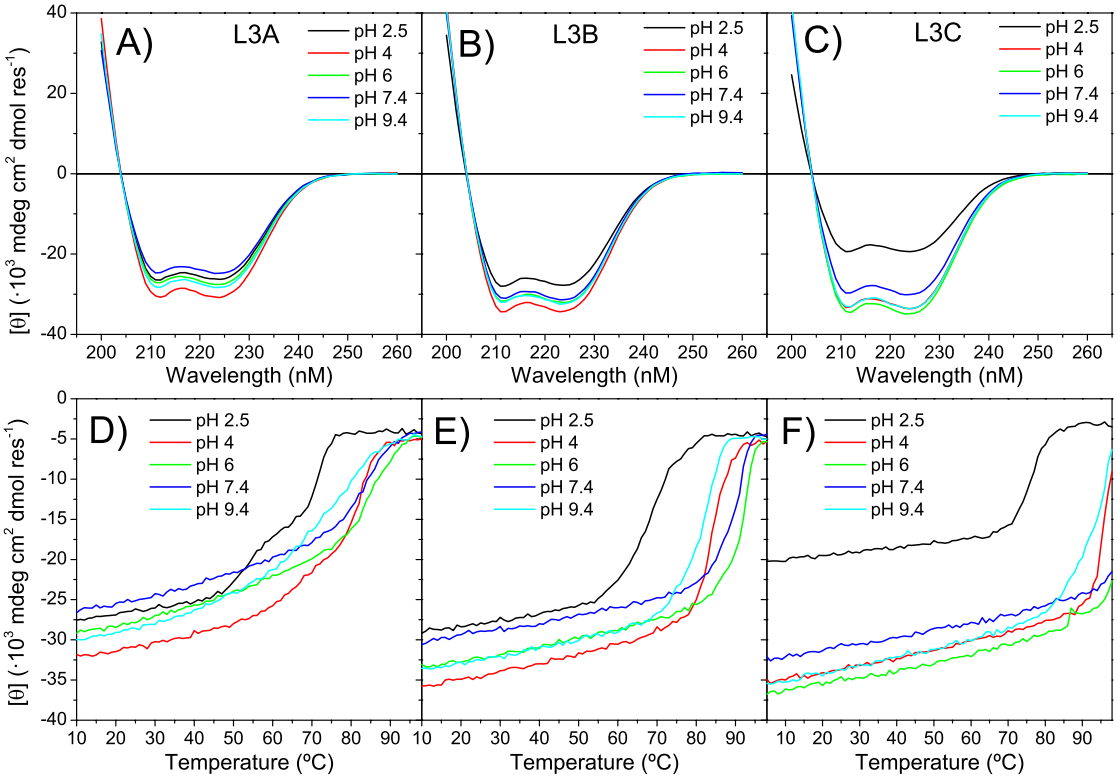

Fig S1: Secondary structure characterization of the CoVS-HR1 proteins. A), B) and C) CD spectra measured with L3A (A), L3B (B) and L3C (C) at different pH and 25°C. Experiments were carried out in 50 mM glycine/HCl pH 2.5, 50 mM sodium acetate pH 4, 50 mM sodium cacodilate pH 6, 50 mM sodium phosphate pH 7.4 and 50 mM sodium carbonate pH 9.4. The percentage of α-helix ranged between 69% at pH 7.4 and 80% at pH 4 for L3A, between 75% at pH 2.5 and 92% at pH 4 for L3B, and between 53% at pH 2.5 and 94% at pH 6 for L3C. **D), E) and F)** Thermal scans monitored by CD at 222 nm showing the thermal unfolding of the proteins L3A (D), L3B (E) and L3C (F) at different pH. CD data have been normalized per mole of amino acid residue. Protein concentration was 15 μM in all experiments.

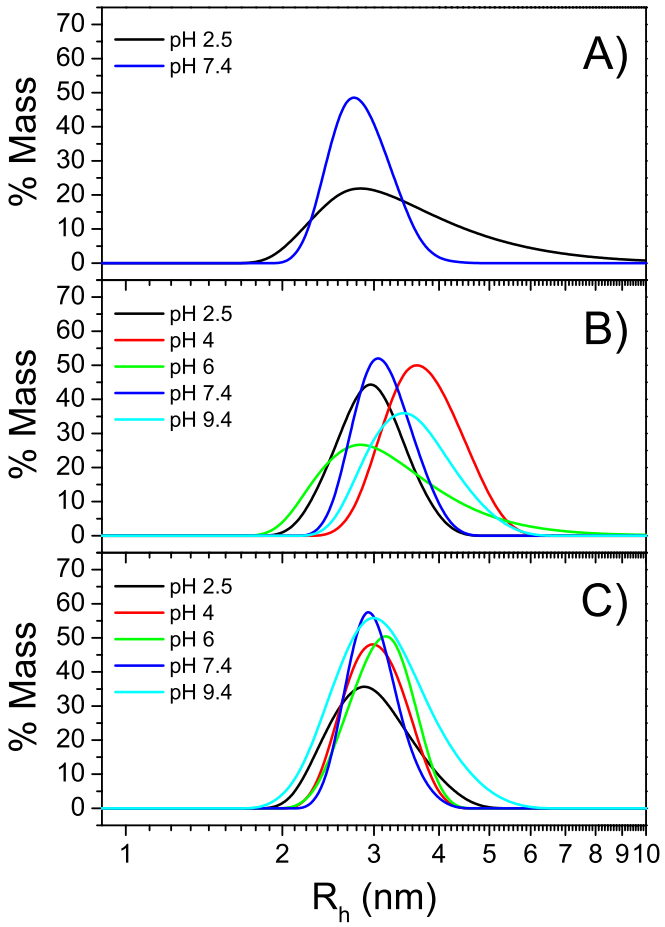

Fig S2: Molecular size characterization of the CoVS-HR1 proteins by dynamic light scattering. Hydrodynamic radii distributions were measured at 25°C with L3A **(A)**, L3B **(B)** and L3C **(C)** at 15 μM protein concentration and different pH using the same buffers as in the CD experiments. The expected R_h for the monomeric proteins is about 3.3 nm.

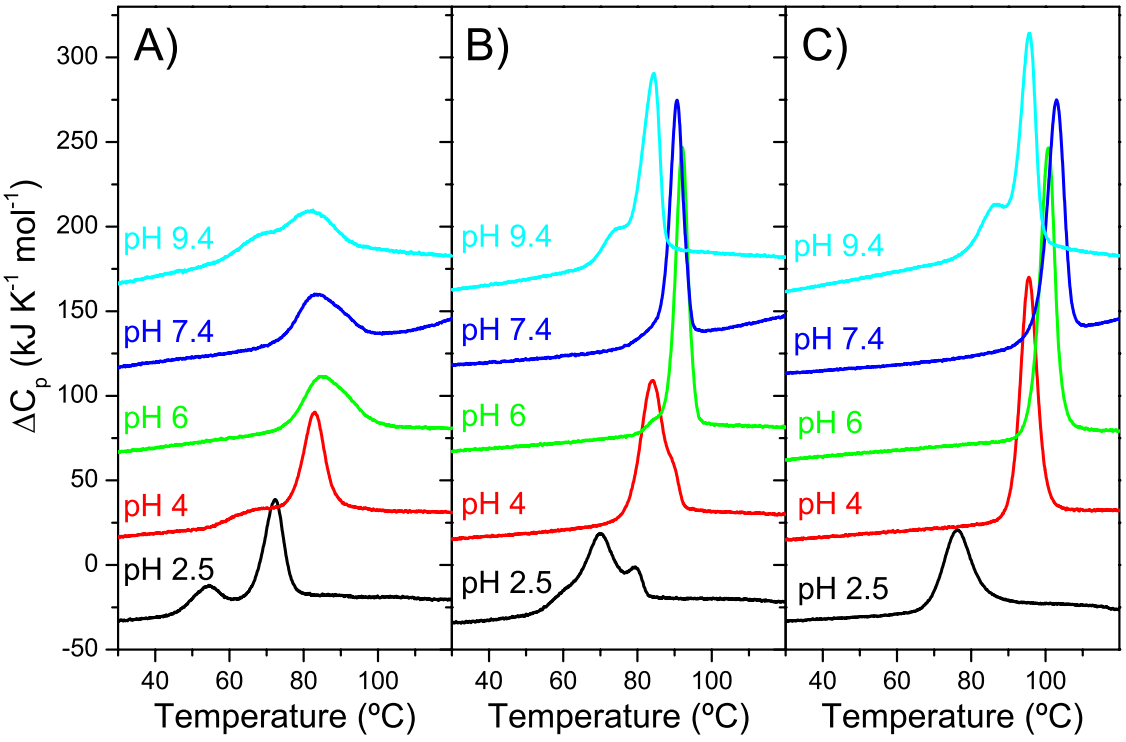

Fig S3: Thermal stability of the CoVS-HR1 proteins at different pH analyzed by DSC. Proteins L3A (A), L3B (B) and L3C (C) were analyzed at 30 μM protein concentration and different pH values using the same buffers as in experiments shown in Fig S1. Scan rate was 90 °C h^{-1} in all experiments. The curves have been displaced vertically in 50 kJ K^{-1} mol^{-1} intervals for the sake of clarity.

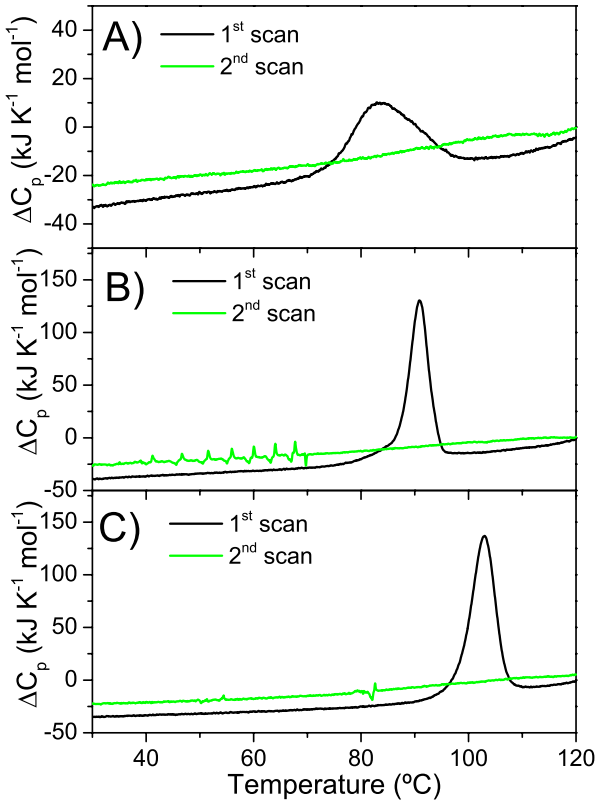

Fig S4: Reversibility test of the thermal unfolding of the CoVS-HR1 proteins. Proteins L3A **(A)**, L3B **(B)** and L3C **(C)** were analyzed by DSC at pH 7.4 in 50 mM sodium phosphate buffer at a scan rate of 90 °C h^{-1}. Protein concentration was 30 μM. Second consecutive scans (green) were carried out with the same sample after the first scans (black).

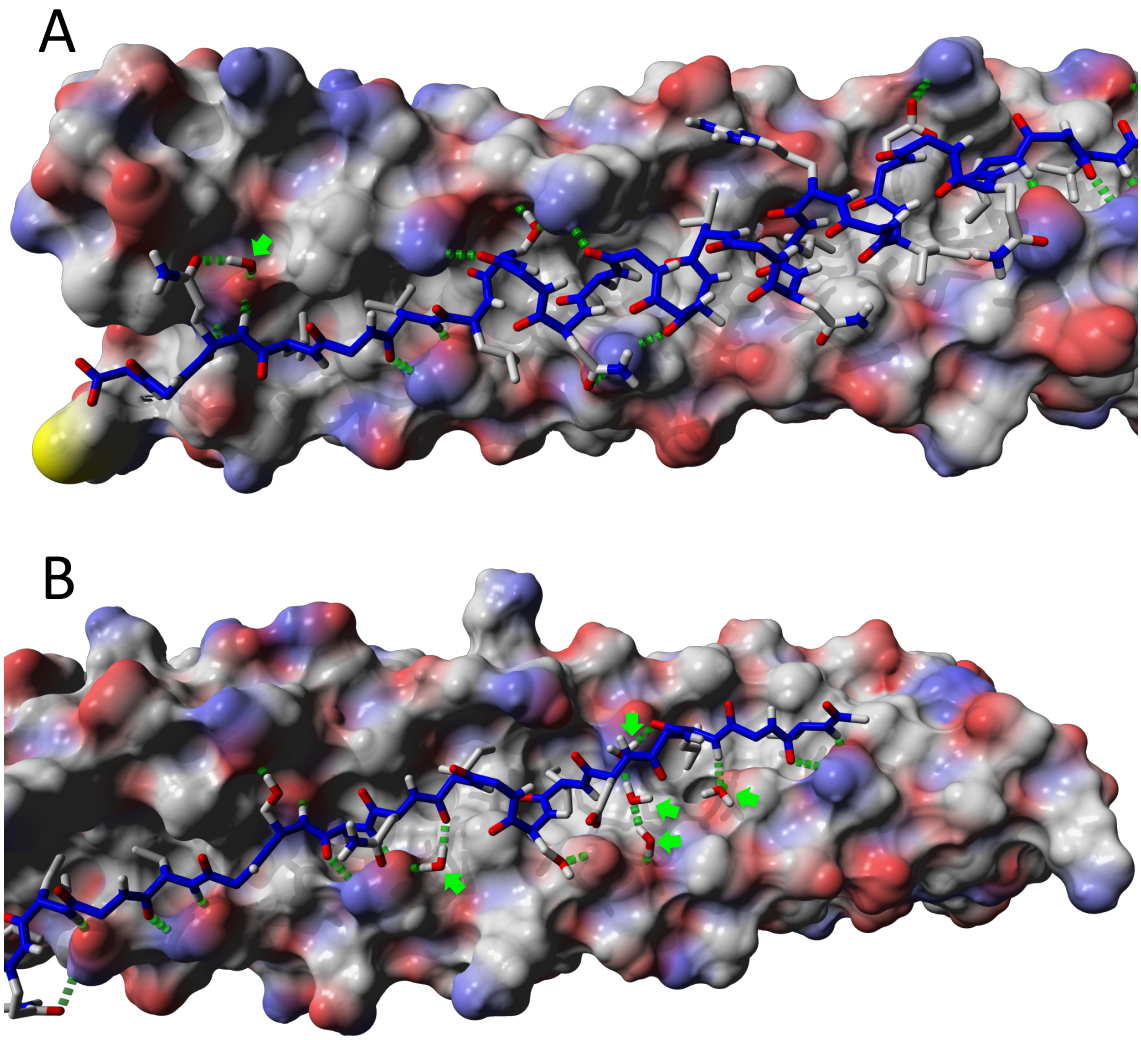

Fig S5: Interfacial interactions between the CoVS-HR1 L3B protein and the V39E peptide in their crystallographic complex structure (PDB id. 7ZR2). Panel **A)** shows the N-terminal half of the HR1 binding crevice. Panel **B)** shows the C-terminal HR1 half. The L3B protein is represented as molecular surface and colored according to element type (CPK scheme). The peptide is represented using sticks with backbone atoms colored in blue and side chains involved in interactions with the protein are colored in CPK scheme. Several buried water molecules at the protein peptide interface are highlighted with green arrows. Interfacial hydrogen bonds are represented in green dashed lines.

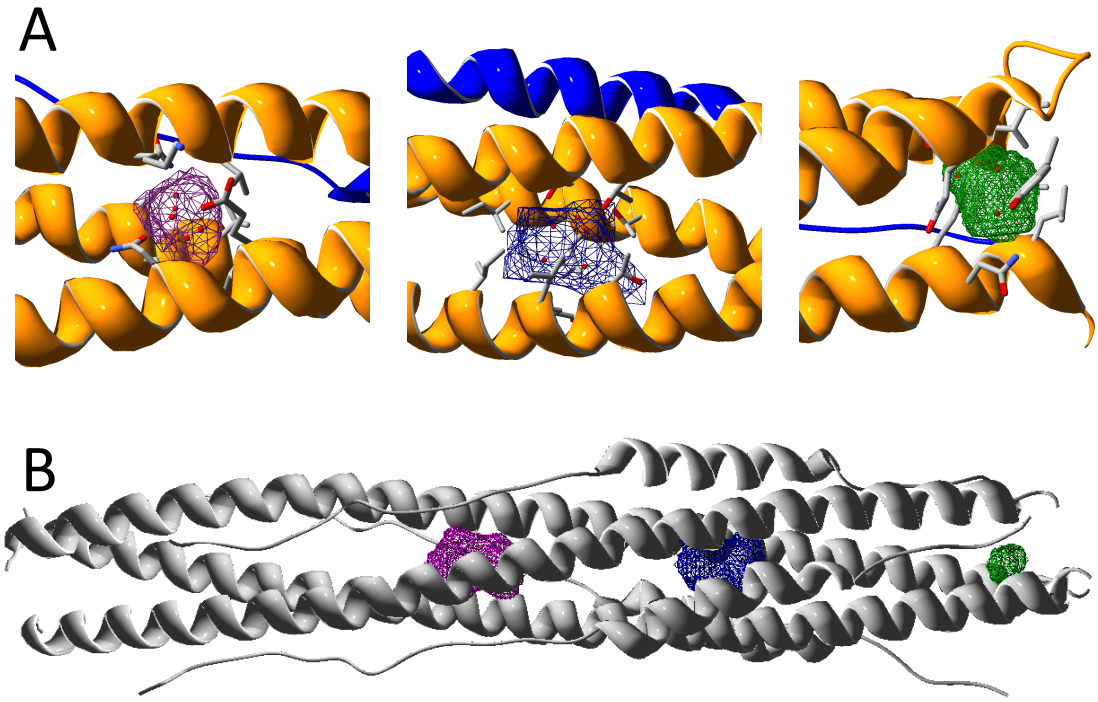

Fig S6: Internal water-filled cavities inside the HR1 coiled-coil. A) Three sections of the L3B-V39E complex structure showing internal cavities in the HR1 trimeric coiled-coil filled with several buried water molecules well resolved in the crystallographic structure. The cavities are depicted with grids and the water molecules are represented as red spheres. Side chains of HR1 residues lining the cavities are highlighted with sticks. **B)** Crystal structure of the 6-helix bundle post-fusion state of HR1 and HR2 in S2 (PDB id. 6LXT) showing internal cavities similar to those observed in the L3B-V39E complex structure (A). No water molecules were resolved in this structure because of its lower resolution (2.9Å).

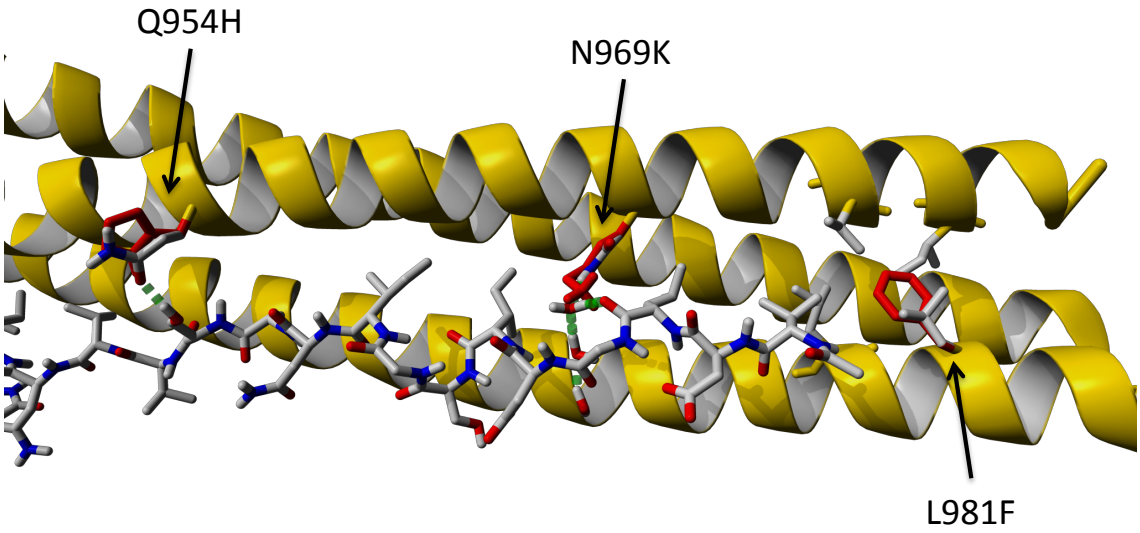

Fig S7: HR1 mutations in Omicron BA.1 variant. The side chains of amino acids in the L3B protein that are structurally equivalent to those mutated in Omicron HR1 are highlighted with sticks and the Omicron mutations are indicated with arrows. Original side chains are colored in CPK scheme and mutant side chains are colored in red. Similar side chain rotamers have been chosen for the mutant amino acids compared to the original ones. Hydrogen bonds involving the side chains are represented with green dashed lines. Due to its larger size, the Lys 969 side chain in Omicron may exclude one of the interfacial water molecules bridging the protein-peptide interaction.