Structure, Volume *31*

Supplemental Information

Structural insights into the activity regulation of full-length non-structural protein 1 from SARS-CoV-2

Ying Wang, John Kirkpatrick, Susanne zur Lage, and Teresa Carlomagno

Supplemental Figures

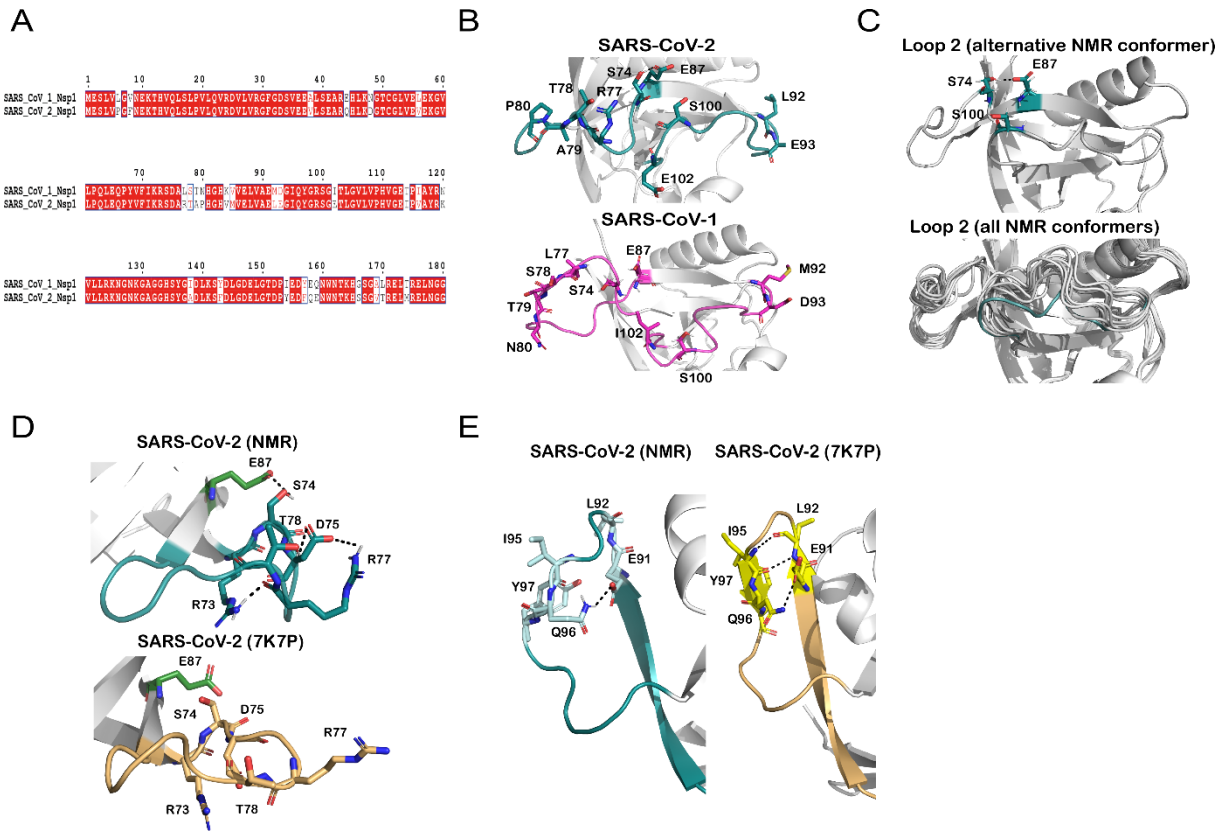

Figure S1. Comparison of the solution structure of the NTD of SARS-CoV-2 FL-Nsp1 with the solution structure of SARS-CoV-1 Nsp1-NTD and the crystal structure of SARS-CoV-2 Nsp1-NTD, related to Figure 1. (A) Sequence alignment of Nsp1 from SARS-CoV-1 and SARS-CoV-2. The alignment was done with Clustal Omega, and the colored figures were generated by ESPript 3.0 [S1, S2]. Residues boxed in red are identical. **(B)** Conformational differences in loop1 and loop2 between Nsp1 from SARS-CoV-2 (turquoise, upper panel) and SARS-CoV-1 (magenta, lower panel). Residues that are different between the two species are labeled and shown as sticks. (C) Variability of loop 2 conformations in the NMR structure of SARS-CoV-2 FL-Nsp1. The upper panel shows the conformation of loop 2 of one member of the NMR ensemble, where it deviates substantially from that of representative conformer shown in (B). The lower panel shows an overlay of the conformations of loop 2 for all members of the NMR ensemble. The structure of the upper panel is in turquoise. (D) Conformational differences of loop 1 between the solution NMR structure of SARS-CoV-2 FL-Nsp1 (turquoise, top) and the crystal structure of PDB 7K7P (light-orange, bottom). PDB 7K7P is shown as representative of all available crystal structures of SARS-CoV-2 Nsp1-NTD (PDB codes: 7K7P, 7K3N and 7EQ4). Residue E87 is colored in green in both structures and residues that are

involved in the formation of hydrogen-bonds are shown as sticks. Hydrogen-bonds are shown as black dashed lines. (E) The β4 strand and loop 2 in the NMR structure of SARS-CoV-2 FL-Nsp1 (turquoise, left) and in the crystal structure of PDB code 7K7P (light-orange, right). Residues that are implicated in the formation of a small β-sheet in the crystal structures are shown as sticks in both structures (light-blue and yellow, respectively). Polar contacts and hydrogen bonds are represented as black dashed lines.

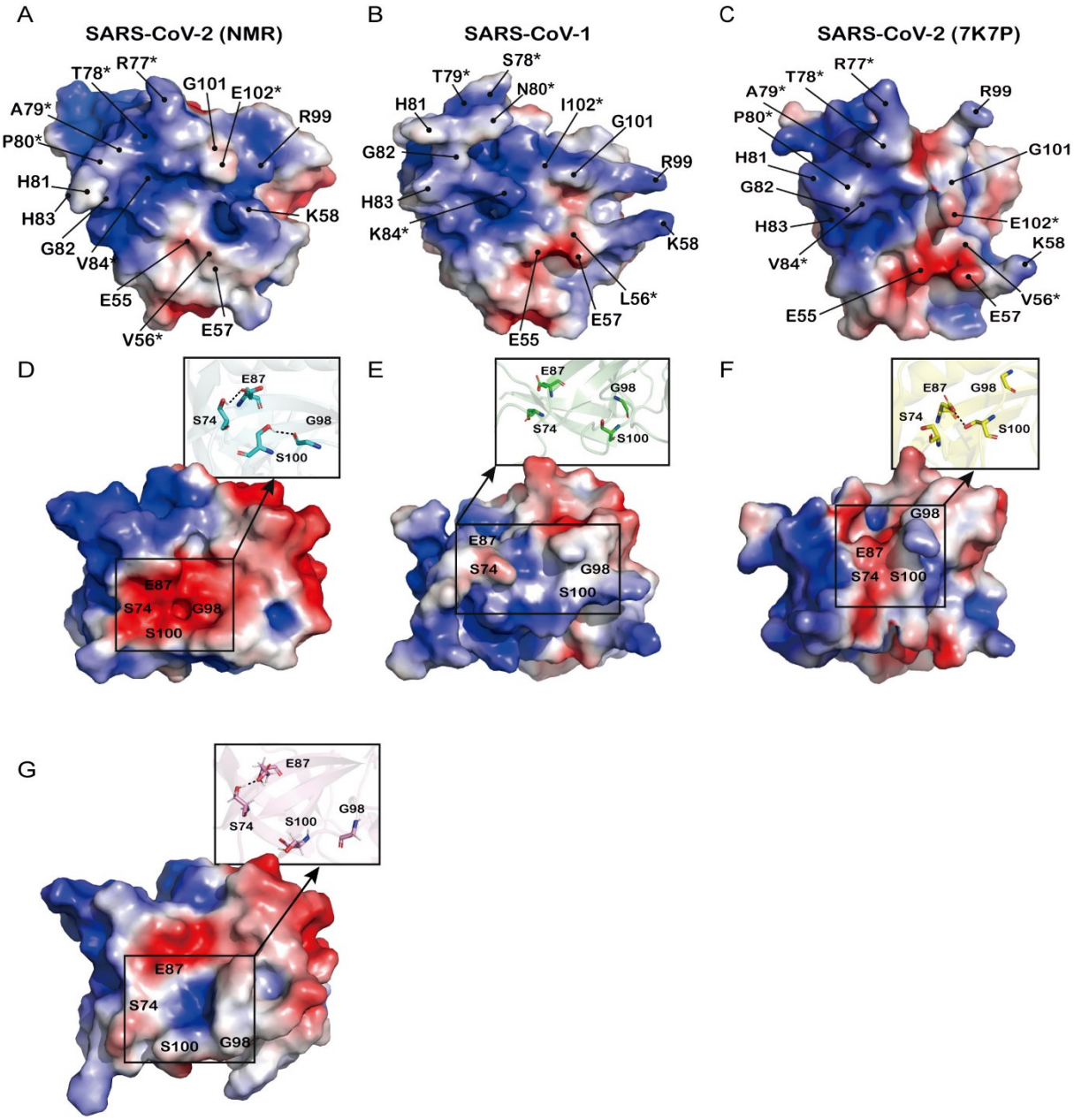

Figure S2. Comparison of electrostatic surfaces of the solution structures of SARS-CoV-2 and SARS-CoV-1 Nsp1-NTD and the crystal structure of SARS-CoV-2 Nsp1-NTD, related to Figure 3. The tails of the solution structure of SARS-CoV-2 Nsp1-NTD were removed *in silico* in all figures. **(A–C)** The electrostatic surface is broadly similar for all three structures, despite differences in the amino-acid sequence between the SARS-CoV-1 and SARS-CoV-2 proteins (labeled with asterisks). PDB 7K7P is shown as representative of all available crystal structures of SARS-CoV-2 Nsp1-NTD. Due to structural differences in residues 55–57, the electrostatic surfaces differ at this site. **(D–F)** Emergence of a negatively charged pocket in the solution structure of SARS-CoV-2 Nsp1-NTD. The residues forming the pocket (S74, E87E, G98 and S100) are colored in cyan and green for the solution structures

of SARS-CoV-2 and SARS-CoV-1 Nsp1-NTD, respectively, and yellow for the crystal structure of SARS-CoV-2 Nsp1-NTD. Hydrogen bonds are represented as black dashed lines. **(G)** The negative pocket is absent from the NMR conformers of SARS-CoV-2 Nsp1-NTD where the conformation of loop 2 differs from that of the representative structure of the ensemble. The conformer shown here is the same as in Figure S1C. The sequence boundaries of the NTD domains as shown are as follows: SARS-CoV-2 Nsp1-NTD (NMR), residues 10–128; SARS-CoV-1 Nsp1-NTD, residues 13–127; SARS-CoV-2 Nsp1-NTD (7K7P), residues 10–126.

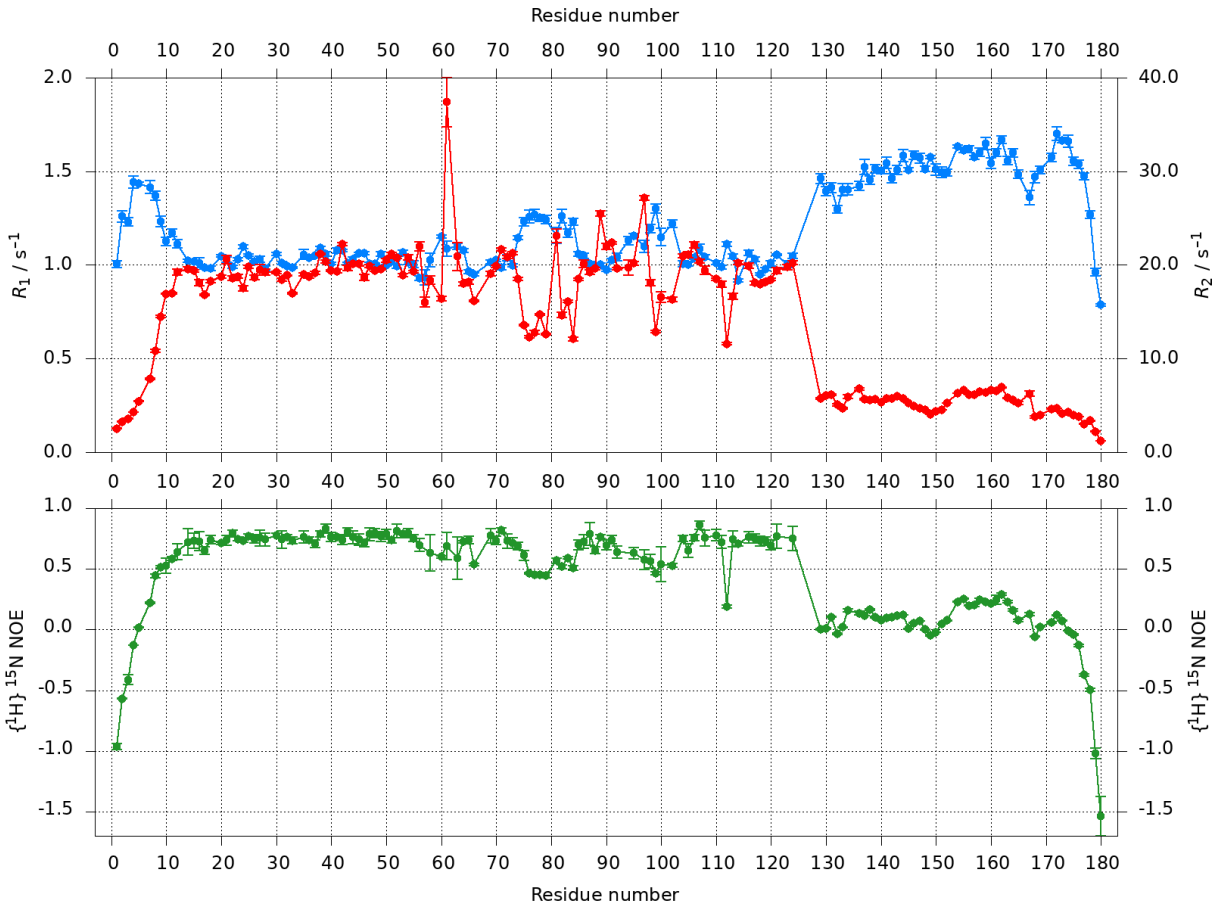

Figure S3. Backbone ^{15}N relaxation parameters of SARS-CoV-2 FL-Nsp1 in solution, related to Figure 2. Upper plot: ^{15}N R_1 (blue; left-axis) and R_2 (red; right-axis) rate-constants; lower plot: {^{1}H}^{15}N heteronuclear NOEs (green). R_2 rate-constants were back-calculated from the measured R_1 and $R_{1\rho}$ rate-constants as described in the Methods section. Parameter values and associated uncertainties (shown as error-bars in the plots above) are listed in Table S1. For the R_1 and R_2 rate-constants, the associated uncertainties are derived from the least-squares fitting of the experimental intensity-decay curves to the respective exponential-decay functions as implemented by FuDA (D. Flemming Hansen; https://www.ucl.ac.uk/hansen-lab/fuda/). For the {^{1}H}^{15}N heteronuclear NOEs, the associated uncertainties are calculated from the uncertainties of the peak-intensities in the saturated and reference sub-spectra (according to standard error-propagation formulae); the uncertainties in the peak-intensities themselves are derived from the least-square fitting of the spectral peak-shapes, as implemented by FuDA.

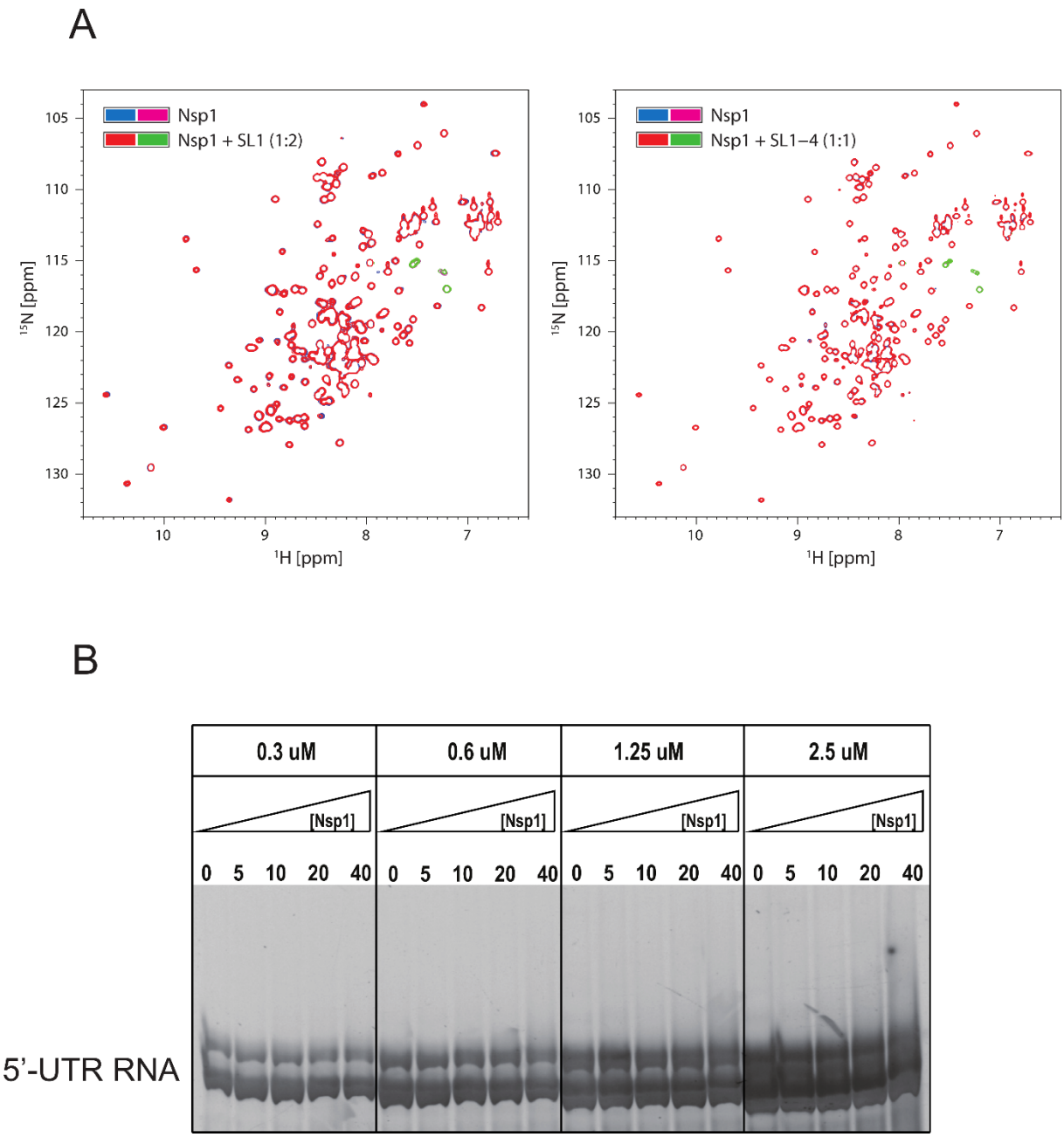

Figure S4. SARS-CoV-2 FL-Nsp1 does not show a binary interaction with the 5'-UTR, related to Figure 3 and 4. (A) Left: Overlay of ^{15}N-HSQC spectra at 600 MHz of ^{15}N-labelled FL-Nsp1 in isolation ([Nsp1] = 400 µM; positive/negative contours in blue/pink) and in combination with a two-fold stoichiometric excess of SL1 RNA ([Nsp1] = 270 µM; positive/negative contours in red/green). Right: Overlay of ^{15}N-HSQC spectra at 850 MHz of ^{15}N-labelled FL-Nsp1 in isolation ([Nsp1] = 200 µM; positive/negative contours in blue/pink) and in combination with equimolar SL1–4 RNA ([Nsp1] = 170 µM; positive/negative contours in red/green). The sample buffer was as for other NMR experiments, except that the concentration of NaCl was lowered from 200 mM to 50 mM. **(B)** EMSAs of the SARS-CoV-2

5’-UTR RNA in the presence of increasing stoichiometric ratios of FL-Nsp1. Four different RNA concentrations were tested (0.3, 0.6. 1.2 and 2.4 μM) in the presence of Nsp1 at 0-, 5-, 10-, 20- and 40-fold excess. The very bottom of the gel, which showed bands corresponding to shorter RNAs arising from degradation of the full-length 5’-UTR, has been excised from this image.

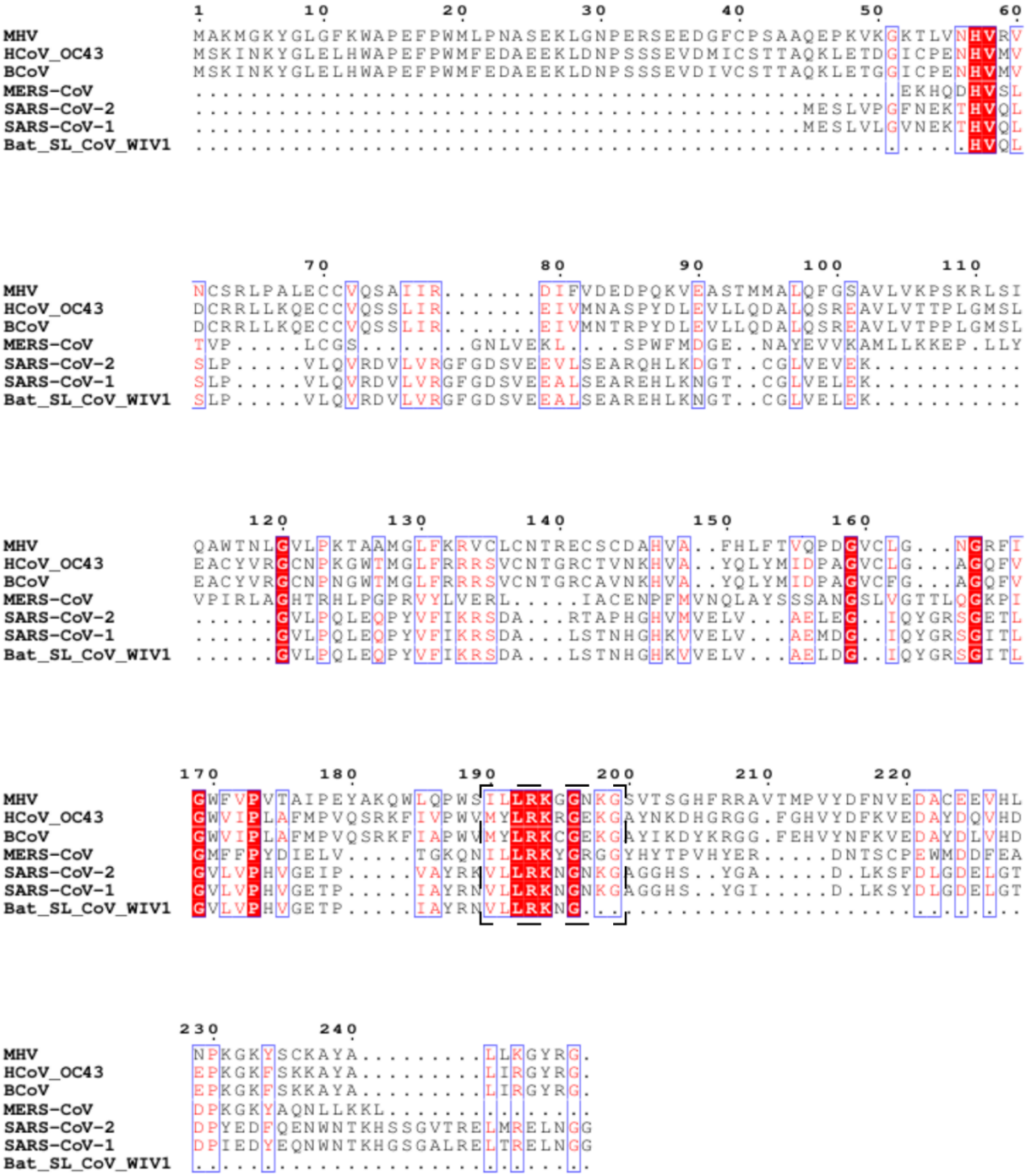

Figure S5. Sequence alignment of SARS-CoV-2 Nsp1 with other Nsp1 proteins derived from beta-coronaviruses, related to Figure 3. Alignments were done with Clustal Omega, and the colored figures were generated using ESPript 3.0 [S1, S2]. Residues boxed in red are identical. The conserved LRKxGxKG motif is indicated with a black dashed box.

Supplemental References:

[S1] X. Robert, and P. Gouet. (2014). “Deciphering key features in protein structures with the new ENDscript server,” Nucleic Acids Res *42*, 320-324. https://doi.org/10.1093/nar/gku316.

[S2] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, and J. Söding. (2011). “Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega,” Mol. Syst. Biol *7*, 539. https://doi.org/10.1038/msb.2011.75.