Supporting Information

Supplemental Figures and Experimental Protocols

NMR-based analysis of nanobodies to SARS-CoV-2 Nsp9 reveals a

possible antiviral strategy against COVID-19

Gennaro Esposito^{1,2}, Yamanappa Hunashal^{1‡}, Mathias Percipalle^{1,3‡}, Tomas Venit^{4‡}, Mame Massar Dieng⁴, Federico Fogolari^{5,2}, Gholamreza Hassanzadeh⁶, Fabio Piano^{4,7}, Kristin C. Gunsalus^{8,9}, Youssef Idaghdour^{4,7}, Piergiorgio Percipalle^{4,10*}

¹Chemistry Program, Science Division, New York University Abu Dhabi, Abu Dhabi 129188, United Arab Emirates; ²Istituto Nazionale Biostrutture e Biosistemi, 00136 Roma, Italy; 3 Department of Chemistry and Magnetic Resonance Center, University of Florence, 50019 Florence, Italy; ⁴Biology Program, Science Division, New York University Abu Dhabi, Abu Dhabi 129188, United Arab Emirates; ⁵Dipartimento di Scienze Matematiche, Informatiche, e Fisiche, Udine Univer-sity, 33100 Udine, Italy; ⁶VIB Nanobody Core, Vrije Universiteit Brussel, 1050 Brussels, Belgium; ⁷Public Health Research Center, New York University Abu Dhabi, Abu Dhabi 129188, United Arab Emirates; ⁸Center for Genomics and Systems Biology, New York University Abu Dhabi (NYUAD), P.O. Box 129188, Abu Dhabi, United Arab Emirates; 9 Department of Biology, Center for Genomics and Systems Biology New York University, New York, NY, 10003, USA; ¹⁰Department of Mo-lecular Bioscience, The Wenner Gren Institute, Stockholm University, Stockholm, Sweden

‡ Equal contributions

* Correspondence: Piergiorgio Percipalle, address as above

Tel.: +97126284389

Email: pp69@nyu.edu

‡ These authors equally contributed to the work

* Corresponding author

Contents

I. Experimental procedures

triSer-Nsp9 protein preparation. Recombinant wild-type (wt) Nsp9 and triSer-Nsp9 proteins carrying the three amino acid substitutions C14S, C23S, C73S were provided by ASLA Biotech (Lot No. 2007ZJ03NSP01). Unlabeled and uniformly 15N,13C-labelled SARS-CoV-2 C14S,C23S,C73S NSP9 (triSer-NSP9) was expressed with an additional methionine at the N-terminus (M0). A construct was used for expression where the histag is located after the stop codon and therefore is not expressed. The scheme for this genetic construct was: pET22b(NdeI)-NSP9(113 aa + M0)-pET22b(XhoI) where CATATG is for Ndel; CTCGAG for Xhol; TAA is a stop codon, corresponding to DNA sequence:

CATATGAACAACGAACTGAGTCCGGTGGCACTGCGTCAGATGAGTAGTGCCGCCGGCACC ACCCAGACCGCCAGTACAGATGATAATGCCCTGGCCTATTATAATACCACCAAAGGTGGTC GCTTTGTTCTGGCACTGCTGAGCGATCTGCAGGATCTGAAATGGGCACGTTTTCCGAAAAG CGATGGTACCGGCACCATCTATACCGAACTGGAACCGCCGAGTCGTTTTGTGACCGATAC CCCGAAAGGTCCGAAAGTGAAATATCTGTATTTTATTAAGGGTCTGAACAACCTGAATCGT GGTATGGTGCTGGGTAGTCTGGCAGCAACCGTTCGCCTGCAGTAACTCGAGCACCACCAC CACCACCACTGA.

After IPTG induction of the Escherichia coli BL21 (DE3) transformed strain, the cells were grown at 22°C for 18 hours in M9 medium containing 13C-labeled glucose and 15N‐ labeled ammonium chloride. The cell pellet was collected by centrifugation and resuspended in lysis buffer (1/40 volume: 1xPBS, protease inhibitor cocktail), sonicated and then centrifuged at 32000xg for 30 min. The cell pellet was dissolved in 6M GuCl,

50mM NaPi, pH 7.3 solution and incubated for 40 min. The cell lysate was again centrifuged at 32000xg. The recombinant protein was further purified by cation exchange chromatography using a HiTrap SP HP column (GE Healthcare, 20 ml). The protein rich fractions were pooled, concentrated and applied to a Superdex 200 size-exclusion column (GE Healthcare, 120 ml). The protein was refolded by drop dilution into refolding buffer (100mM Tris, 5mM EDTA, 0.75M arginine, pH 8) on ice. The protein was further dialyzed twice against PBS (1:40 volume), and then again concentrated and dialyzed by Amicon Ultra 15 (Merck Millipore) against 20 mM ammonium acetate. The final protein preparation was lyophilized.

The uniformly 15N,13C-labelled wild-type SARS-CoV-2 NSP9 was obtained according to the protocol kindly provided by the NMR COVID-19 Consortium, ^[1] leading to a final product with an additional GlyAlaMetGly tetrapeptide at the N-terminus.

For immunization, about 3 mg of lyophilized triSer-Nsp9 protein were dissolved in 20 mM ammonium acetate (pH 6.7). As the lyophilized protein was not fully dissolved in this buffer, the supernatant obtained by centrifugation, was recovered and stored for further use in immunization, panning and ELISA screening. The protein pellet (insoluble protein fraction) was dissolved in a small amount of dimethyl sulfoxide (DMSO), and the protein in DMSO was immediately diluted in PBS. Both the supernatant obtained by dissolving the protein in ammonium acetate, followed by centrifugation, and the protein dissolved in DMSO & diluted in PBS were used separately in the immunization, panning and ELISA screening experiments described below. In these experiments, we refer to the protein in

ammonium acetate as NSP, and the protein dissolved in DMSO and diluted in PBS as INS.

Immunization. A llama was subcutaneously injected on days 0, 7, 14, 21, 28 and 35, each time with about 100 µg of recombinant triSer-NSP9 dissolved in ammonium acetate (here referred to as NSP) & 100 µg of recombinant triSer-NSP9 protein dissolved in DMSO/PBS (here referred to as INS). Each injection was performed at 2 sites at the end of the neck (cranially or caudally to the scapula bone) and at 2 sites at the two back limbs. The protein dissolved in ammonium acetate (NSP) was injected on the left side of the animal and the protein dissolved in DMSO/PBS (INS) was injected on the right side of the animal. The adjuvant used was Gerbu adjuvant P. On day 40 (5 days after last immunization), about 100 ml anticoagulated blood was collected from the llama for peripheral blood lymphocytes (PBLs) preparation and library generation.

Construction of a VHH (Nanobody) library. A VHH library was constructed from PBLs to screen for the presence of antigen-specific nanobodies. To this end, total RNA was prepared from PBLs and about 50 µg of total RNA was used as template for first strand cDNA synthesis with oligodT primers. Using this cDNA, the VHH encoding sequences were amplified by PCR, digested with PstI and NotI, and cloned into the PstI and NotI sites of the phagemid vector pMECS. In pMECS vector, the nanobody sequence is followed by a linker, HA tag and His₆ tag (Nanobody-AAAYPYDVPDYGSHHHHHH). Electro-competent E. coli TG1 cells were transformed with the recombinant pMECS vector resulting in a VHH (nanobody) library of about 2 x 10⁹ independent transformants. This library is named "Core 145/146 library". About 87% of the transformants from this

library harbored the vector with the right insert size (size of VHH-encoding sequences), as demonstrated by PCR analysis of 95 randomly picked colonies.

Nanobodies isolation. The Core 145/146 library was panned separately on NSP & INS protein samples immobilized on solid-phase (100 µg/ml in 100 mM NaHCO3 pH 8.2). Three rounds of panning were performed on each protein sample. The enrichment for antigen-specific phages was assessed after each round of panning by comparing the number of phagemid particles eluted from antigen-coated wells with the number of phagemid particles eluted from negative control (uncoated blocked) wells.

For panning on protein sample dissolved in ammonium acetate (NSP), the phage population was enriched for antigen-specific phages about 20-fold, 100-fold and 700-fold after the 1st, 2nd and 3rd round, respectively. In total, 190 colonies from 2nd and 3rd rounds (95 colonies from each round) were randomly selected and analyzed by ELISA for the presence of antigen-specific Nanobodies in their periplasmic extracts (ELISA using crude periplasmic extracts including soluble Nanobodies). The ELISA tests were performed on both protein dissolved in ammonium acetate (NSP) and the protein dissolved in DMSO/PBS (INS). Uncoated blocked wells served as negative control (blank) for ELISA. Out of these 190 colonies, 165 colonies scored positive for NSP and/or INS. Based on sequence data of the 165 positive colonies, 71 different Nanobodies were identified, belonging to 18 different CDR3 groups (B-cell lineages). Out of these 71 Nanobodies, 65 Nanobodies are positive for both NSP and INS, while 6 Nanobodies bind only to NSP (See Excel file). The 71 different nanobodies specific for Nsp9 which resulted from panning on protein dissolved in ammonium acetate bear the code "NSP" in their names.

For panning on protein sample dissolved in DMSO/PBS (INS), the phage population was enriched for antigen-specific phages about 30-fold and 80-fold after the 2^{nd} and 3^{rd} round, respectively. No clear enrichment was observed after the 1st panning round. Here also, 190 colonies from $2nd$ and $3rd$ rounds (95 colonies from each round) were randomly selected and analyzed by ELISA for the presence of antigen-specific Nanobodies in their periplasmic extracts (ELISA using crude periplasmic extracts including soluble Nanobodies). The ELISA tests were performed on both the protein dissolved in ammonium acetate (NSP) and the protein dissolved in DMSO/PBS (INS). Uncoated blocked wells served as negative control (blank) for ELISA. Out of these 190 colonies, 156 colonies scored positive for INS and/or NSP. Based on sequence data of the 156 positive colonies, 65 different nanobodies were identified, belonging to 22 different CDR3 groups (B-cell lineages). Out of these 65 nanobodies, 59 nanobodies are positive for both INS and NSP, while 2 nanobodies are only positive for INS and 4 nanobodies bind only to NSP (See Excel file). The 65 different nanobodies specific for SARS-CoV-2 Nsp9 which resulted from panning on protein dissolved in DMSO/PBS bear the code "INS" in their names.

In summary, the above-described experiments resulted in 136 different nanobodies, belonging to 40 different CDR3 groups (B-cell lineages). Out of these 136 nanobodies, 124 specifically recognize both protein dissolved in ammonium acetate and protein dissolved in DMSO/PBS, while 12 nanobodies recognize either of these protein samples. The most likely explanation for differential binding of the later 12 nanobodies to 2 different protein samples is the experimental variabilities since the ELISA data here are all from single experiments. This is strongly supported by the fact that within the same CDR3

group (the same target epitope), there are both nanobodies binding to two different protein samples and also nanobodies which bind only to one protein sample. nanobodies belonging to the same CDR3 group (same B-cell lineage) are very similar and their amino acid sequences suggest that they are from clonally-related B-cells resulting from somatic hypermutation or from the same B-cell but diversified due to RT and/or PCR error during library construction. Nanobodies belonging to the same CDR3 group recognize the same epitope but their other characteristics (e.g. affinity, potency, stability, expression yield, etc.) can be different.

 \overline{D}

*Expression and purification of nanobodies***.** Each nanobody was expressed in about 2L of TB medium (6 x 330 ml). Small scale overnight cultures were started in 20 mL of LB medium supplemented with ampicillin (100 µg/ml) and glucose (1%). The overnight cultures were used to inoculate the shaker flasks, each containing 330 mL of TB medium supplemented with ampicillin (100 µg/ml) and glucose (0.1%), and grown at 220 rpm at 30°C. After the cultures reached an OD600nm of about 0.8, expression was induced by addition of IPTG to final concentration of 1 mM. The induced cultures were then incubated at 20°C at 220 rpm. The following day, the cultures were spun down. The supernatants were discarded and the cell pellets were subject to osmotic shock by resuspending each pellet from 330 ml culture in 4 mL TES buffer & incubating for 2 h at 4°C with gentle shaking, followed by addition of 8 mL of water and further 4 h incubation at 4°C. The periplasmic extracts (PEs) were collected by centrifugation and stored at 4°C. The cell pellets were used for a second osmotic shock cycle. When all extracts were collected, 500 µL of His-select matrix was added to each individual periplasmic extract (PE) and incubated for 1 h at 4°C. Each PE was then applied to an empty PD-10 column. The flow

through was applied to the column for a second time. Subsequently, the matrix was washed with a solution of 20 mM imidazole in PBS. Finally, the bound Nanobodies were eluted using a solution of 500 mM imidazole in PBS, in 5 fractions of 1 mL. These fractions were monitored by Nanodrop at 280 nm to quantify the amount of nanobody collected. The fractions containing substantial amounts of protein were pooled and loaded onto a Superdex 75 16/60 or a Superdex 75 10/300 GL Increase size exclusion chromatography (SEC) columns.

Saliva samples collection and storage. Saliva samples from individuals with known SARS-CoV-2 status (positive and negative) were used for this study. Approximately 1 ml of saliva samples were collected in sterile 5 ml falcon tubes (Fisher scientific) without transport medium under NYUAD IRB-approved protocol HRPP-2020-48 (PI Idaghdour), stored directly at 4°C until shipment to a BSL2 laboratory for processing following guidelines from the US Center for Disease Control and Prevention (CDC). The samples were stored at -80°C until processing.

Saliva RNA isolation and SARS-CoV-2 Detection by RT-qPCR. For each sample, 300µl of saliva were used to extract RNA in a BSL2 laboratory using an automated system, Chemagic 360, and Viral DNA/RNA 300 Kit H96, both from Perkin Elmer. Extracted RNA was eluted in 80 μL of elution buffer and used right away for SARS-CoV-2 detection or stored at -80°C until use as per manufacturer's instruction protocol. The full protocol for SARS-CoV-2 detection using RT-qPCR and the Fluidigm BioMark HD system is described in Xie *et al*, 2020. [2] CDC recommended assays (primers and probes) for SARS-CoV-2 detection, and human RNase P (RP) control for RNA extraction and RT-

qPCR reactions were used. The assays were synthesized by Integrated DNA Technology (IDT) and the sequences are as follow: 2019-nCoV_N2 (Forward: TTA CAA ACA TTG GCC GCA AA; Reverse: GCG CGA CAT TCC GAA GAA; Probe: ACA ATT TGC CCC CAG CGC TTC AG); human RNase P (RP) (Forward: AGA TTT GGA CCT GCG AGC G; Reverse: GAG CGG CTG TCT CCA CAA GT; Probe: TTC TGA CCT GAA GGC TCT GCG CG). For each sample, five microliters of extracted RNA were reverse transcribed (RT) to cDNA using 1.25 μL RT Master Mix following the manufacturer protocol (Fluidigm). The converted cDNA was pre-amplified using a pool of the three assays (N1, N2, and RP) diluted into low TE buffer (Thermo Scientific) to a final concentration of 100.5 and 25.5 nM for the primers and probes, respectively, and then mixed with 2.5 μL of Preamplification Master Mix, and 0.635 μL PCR grade water (Fluidigm). After the preamplification step, the PCR reactions were further diluted 1:5 in Low TE buffer (Thermo Scientific) resulting in a final volume of 62.5 μL and ready to use for qPCR. The qPCR mix was prepared using 1.8 μL of diluted cDNA and 2 μL of 2X TaqMan Fast Advanced Master Mix (Thermo Scientific) and 0.2 μL 20X GE Sample Loading Reagent (Fluidigm). Next, 3 μL qPCR mix from each sample was loaded into the sample inlet in the 192.24 integrated fluid circuits (IFC, Fluidigm). For each assay, 3 μL of primer/probe mix (13.5X) was mixed with 1 μL of 4X Assay Loading Reagent (Fluidigm), and 3 μL of each assay mixed was loaded in the assay inlet in the 192.24 or Flex Six IFC chip (Fluidigm). The chip was then placed in an integrated fluidic circuit (IFC) controller RX machine to preload the samples and the assays, and then, loaded onto the BioMark HD instrument (Fluidigm) for RT-qPCR using 35 cycles. The raw amplification data were acquired using the Fluidigm data collection software and analyzed using the Fluidigm Real-Time PCR

Analysis software 3.0.2. The expression of viral N2 protein was normalized to human RNase P expression and compared between samples. Each RT-qPCR reaction was repeated at least 4 times.

Western blot analysis. Purified wtNsp9 was diluted in 1xPBS buffer to prepare 1mg/ml stock solution with protease inhibitors. The stock wtNsp9 solution was further diluted with 1xPBS containing 1mg/ml bovine serum albumin (BSA) to the final concentration of 15 µg of total protein per loading with decreasing amount of wtNsp9 per sample. Following SDS PAGE, protein samples were visualized by PageBlue Protein Staining solution (Thermo Fisher) according to the manufacturer's instructions. For preparation of saliva samples, 20µl of saliva was mixed with 2x RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 % NP-40, 0.5 % sodium deoxycholate, 0.1 % SDS) in 1:1 ratio and total protein concentration measured by the Pierce BCA Assay Kit (Thermo Fisher Scientific). Subsequently 150 µl of the saliva sample was mixed with 50 µl of 4x Laemmli buffer with protease inhibitors and based on the protein concentration measurement by BCA assay, approximately 15µg of total protein per sample was loaded to the 15% SDS-PAGE gel. The extracts were separated under reducing conditions and transferred to polyvinylidene difluoride (PVDF) membrane and blocked with 3% BSA in 1xTBST buffer (20 mM Tris, 150 mM NaCl, 0,1% Tween 20) for 1 hour. After blocking, membranes were incubated with nanobodies 2NSP23 and 2NSP90 (dilution 1000x in 1xTBST buffer) overnight at 4°C and subsequently washed 4 times for 15 minutes in 1xTBST buffer. Immunoblots were then stained with HRP-conjugated secondary antibodies (dilution 2000x in 1xTBST buffer) recognizing either 6xHis (ab237339, Abcam) or VHH epitopes

(128-035-230, Jackson ImmunoResearch). Detection was by chemiluminescence using the ECL Western Blot Substrate (Bio-Rad) and imaged by a ChemiDoc MP Imaging system (Bio-Rad).

NMR spectroscopy. TCEP (tris(2-carboxyethyl)phosphine), sodium phosphate, hen egg white lysozyme (HEWL), D₂O and DMSO-d6 were all from Sigma Aldrich (St. Louis, MO, USA). Perdeuterated glycine was from Cambridge Isotope Laboratories (Tewksbury, MA, USA). The uniformly ¹⁵N,¹³C-labelled wild-type SARS-CoV-2 Nsp9 and triSer-Nsp9 were obtained as above described in Experimental **Protocols**. The wild-type protein samples were typically prepared in H₂O/D₂O 95/5 with phosphate buffer 10-30 mM, 60-150 mM NaCl, 0.4-1.2 mM TCEP, 0.004-0.01% NaN3, pH 7.0-7.1. For the triSer-Nsp9 samples, additional buffers and pH conditions were explored besides the phosphate (25 mM with 75 mM NaCl, pH 7.0). In particular triSer-Nsp9 was first prepared and lyophilized from 20 mM ammonium acetate solution. Upon redissolving in H₂O/D₂O, pH was 6.1 (residual salt was observed) and the quality of the spectra was poor. Addition of saline phosphate buffer resulted in pronounced precipitation. This was also the case with the protein lyophilized from pure water that exhibited better solubility when dissolved without further additions (pH 4.7), or after dialysis against water (pH 3.7), or with 5% addition (in volume) of DMSOd6, or in 100 mM perdeuterated glycine, pH 6.1. Under all these conditions, no major differences were detected with respect to phosphate, except for the solubility. In general, for the protein samples, concentrations ranged from 18 to 431 µM, as determined by UV absorption at 280 nm. No relevant difference was observed at different concentrations. For titrations with nanobodies (2NSP23 and 2NSP90) two mother solutions of 18 µM ¹⁵N,¹³C-labeled NSP9 without and with 36 µM unlabeled nanobody were progressively

 $|O|$

mixed. The buffer was 10 mM phosphate, 150 mM NaCl, 0.4 mM TCEP, 0.004% NaN3 with pH 7.2. No precipitation was observed during the whole titrations. The NMR experiments of the isolated proteins were mostly acquired at 298 K. With triSer-NSP9 some experiments were also recorded at higher temperatures (305, 310 and 315 K) in the absence and presence of 5% DMSO-d6 to check for linewidth improvements. For titration of the wild-type sequence with nanobodies, the temperature was lowered to 278 and 276 K before starting the experiments which were in fact conducted at those two temperatures with 2NSP23 and 2NSP90, respectively. The NMR data were collected at 14.0 T (¹H at 600.19 MHz, ¹⁵N at 60.82 MHz) on a Bruker Avance III NMR system equipped with triple resonance cryoprobe. Two-dimensional ¹⁵N-¹H HSQC experiments ^[3] were carried out using sensitivity-improved Echo/Antiecho-TPPI pure phase detection in F1, gradient coherence selection and flip-back pulse for solvent suppression $[4-6]$. Data were collected over 40-50 ppm and 12-16 ppm windows in F1 and F2, respectively, using 128 points in t_1 and 2,048 points in t_2 , with 32, 64 or 256 scans/ t_1 increment and 16 or 32 dummy scans to achieve steady state. Prior to 2D Fourier transform, squared sinebell apodization, shifted at least by $\pi/4$ in either dimensions, and linear prediction to double the data size in t_1 were applied, with zero filling to obtain 2Kx1K matrices of reals. ¹³C was decoupled by band-selective adiabatic π pulses applied at the center of t₁ evolution. ^[7] DOSY ^[8] spectra for determination of diffusion coefficients were acquired by 2D 1H DSTEBPP (Double STimulated Echo BiPolar Pulse) experiments with compensation for convective motions. [9] The z-axis gradient strength was varied linearly from 20 to 90% of its maximum value (~ 60 G/cm) and matrices of 2,048 by 40 points were collected by accumulating 32-64 scans per gradient increment. Water suppression

was achieved by appending to the DSTEBPP sequence a pair of WATERGATE [10] elements in the excitation-sculpting mode [11], as for 1D experiments. Following Fourier transform of the 1D traces, DOSY data fitting to extract the diffusion coefficients was carried out using Dynamics Center (Bruker). The related hydrodynamics calculations were done according to Hansen [12].

MD simulations. The dimer of SARS-COV-2 Nsp9 RNA-Replicase was taken from the structure deposited in the Protein Data Bank (pdb id: 6w4b). The mutation of all cysteines into serines was performed using the program DeepView4.10. [13] The structures were soaked in a box of TIP3P water [14] and 0.150 M NaCl up to at least 14 Å from any solute atom using the program VMD. [15] All molecular simulations were performed using the program NAMD2. [16] First each system was energy minimized by 5,000 steepest descent steps. The dynamics was started at 0 K and the temperature was increased to the target value in 10 ps and was equilibrated for 1 ns rescaling every 0.1 ps. During this phase and in all simulations, pressure was kept constant at 1.0 atmosphere by the Langevin piston method with period 200 ps and decay time 100 ps, at the target temperature. [17,18] In all simulations, except for the heating phase described above, the temperature was controlled through Langevin dynamics with damping constant 2.5 ps[−]1. Interactions were gradually switched off at 12 Å starting at 10 Å. The time step was 1 fs for bonded interactions and 2 fs for nonbonded interactions. Hydrogen bond lengths were restrained by the algorithm Settle. [19] Exactly the same protocol was applied to the wildtype and mutant protein. Simulations were performed for 200 ns.

II. References

[1] E. F. Dudás, R. Puglisi, S.M. Korn, C. Alfano, M.L. Bellone, F. Dal Piaz, G. Kelly, E. Monaca, A. Schlundt, H. Schwalbe, A. Pastore, *Biomol. NMR Assign.* **2021**, doi:10.1007/s12104-021-10011-0.

- [2] X. Xie, T. Gjorgjieva, Z. Attieh, M. M. Dieng, M. Arnoux, M. Khair, Y. Moussa, F. Al Jallaf, N. Rahiman, C. A. Jackson, L. El Messery, K. Pamplona, Z. Victoria, M. Zafar, R. Ali, F. Piano, K. C. Gunsalus, Y. Idaghdour, *Processes* **2020**, *8*, 1425.
- [3] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* **1980**, *69*, 185.
- [4] A. G. Palmer, J. Cavanagh, P. E. Wright, M. Rance, *J. Magn. Reson.* **1991**, *93*, 151.
- [5] J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser,
- O. W. Sørensen, C. Griesinger, *J. Biomol. NMR* **1994**, *4*, 301.
- [6] S. Grzesiek, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 12593.
- [7] D. Rosenfeld, S. H. Panfil, Y. Zur, *J. Magn. Reson.* **1997**, *126*, 221.
- [8] K. F. Morris, C. S. Johnson, *J. Am. Chem. Soc.* **1992**, *114*, 3139.
- [9] A. Jerschow, N. Müller, *J. Magn. Reson.* **1998**, *132*, 13.
- [10] M. Piotto, V. Saudek, V. Sklenár, *J. Biomol. NMR* **1992**, 2, 661.
- [11] T. L. Hwang, A. J. Shaka, *J. Magn. Reson. A* **1995,** 112, 275.
- [12] S. Hansen, *J. Chem. Phys.* **2004**, *121*, 9111.
- [13] T. Schwede, J. Kopp, N. Guex, M. C. Peitsch, *Nucl. Ac. Res.* **2003**, *31*, 3381.

[14] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.

[15] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.

- [16] L. Kale, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, K. Schulten, *J. Comp. Phys.* **1999**, *151*, 283.
- [17] G. Martyna, D. Tobias, M. Klein, *J. Chem. Phys.* **1994**, *101*, 4177.
- [18] S. Feller, Y. Zhang, R. Pastor, B. Brooks, *J. Chem. Phys.* **1995**, *103*, 4613.
- [19] S. Miyamoto, P. A. Kollman, *J. Comp. Chem.* **1992**, *13*, 952.
- [20] C. Zhang, Y. Chen, L. Li, Y. Yang, J. He, C. Chen, D. Su, *Mol. Biomed.* **2020**, *1*, 5, doi:10.1186/s43556-020-00005-0.

III. Supplementary figures and tables

Figure S1. (A) Model of triSer-Nsp9 with red highlights at the positions of the three serines replacing the cysteines in the wild-type sequence. (B) Average pairwise RMSD values upon monomer-monomer superposition along the sequences of the homodimeric subunits of WT SARS-CoV-2 Nsp9 (red) and the triSer-Nsp9 mutant (black), calculated from snapshots selected every 1 ns from a 200 ns molecular dynamics (MD) trajectory. The two subunits of each homodimer are indicated by separate traces. Deviations were observed for one or both subunits of the wild-type homodimer at approximately residues 55-65, as well as at the N- and C- termini, corresponding to a slightly larger conformational spread of wild-type with respect to the corresponding dispersion of the mutant. Overall, however, the observation of a similar MD pattern confirmed that no major difference should occur between the parent and the mutant subunits of the homodimers. Therefore, the triSer-Nsp9 was deemed suitable for raising the HCAbs response.

Figure S2. Representative size exclusion chromatogram profiles of selected nanobodies. The highlighted peaks represent the elution of each nanobody and the respective fractions collected.

Green - glycine, thiol, hydroxyl, amine; Red - small, hydrophobic; Blue - acidic; Magenta - Basic * - conserved position : - strong positive identity . - weak positive identity

B

Figure S3. (A) Annotated amino acid sequences of eight selected nanobodies based on their DNA sequences. Alignment and annotations were done using the ProtParam tool [\(https://web.expasy.org/protparam/\)](https://web.expasy.org/protparam/). (B) SDS-PAGE (top panels) and Western blot (bottom panels) analysis of the 8 purified nanobodies. Western blotting was performed with anti-His tag antibodies. For quality control, 5µg of each purified nanobody were loaded onto two 12.5% SDS-PAGE gels for Coomassie Blue staining or for Western blotting with mouse anti-His tag antibodies detected with a goat anti-mouse-HRP antibody. All purified Nanobodies revealed a single band profile on both gel and Western blot in both reducing and non-reducing conditions.

Figure S4. (A) Aliphatic region overlay from the DOSY contour maps of SARS-CoV-2 triSer-Nsp9 (black), wild-type Nsp9 (red) and hen egg white lysozyme (HEWL, cyan). The *D* values are 1.07±0.01×10−10 m2/s and 1.13±0.02×10⁻¹⁰ m²/s for the mutant and wild-type Nsp9, respectively. For HEWL, considered as a marker, a value of 1.24±0.02×10⁻¹⁰ m²/s is obtained. Protein concentrations in H₂O/D₂O 95/5 were ~ 130 µM for Nsp9 variants and 120 µM for HEWL. The overlay shows that the translational diffusion is slower for triSer-Nsp9 (black trace) compared to wild-type Nsp9 (red trace). Considering the crystallographic Nsp9 dimensions [20], the prolate ellipsoids of revolution representing monomers and dimers (with axial ratios 1.304 and 1.805, respectively) correspond to spheres with radius of 1.71 nm and 2.28 nm [12] with an expected D value of 1.06×10⁻¹⁰ m²/s and 1.41×10⁻¹⁰ m²/s, respectively. Therefore, the experimentally determined D values and the reported NMR evidence consistently indicate that the extent of oligomerization is more pronounced for triSer-Nsp9, with an essentially complete dimerization for the visible signals. Instead, the observable NMR signals of wild-type Nsp9 also bear some contribution from monomers, which indicates a difference in the extent of dimerization between triSer-Nsp9 and the natural sequence with three cysteines. (B) Average pairwise RMSD values upon dimer-dimer superposition along the sequences of the homodimer subunits of SARS-CoV-2 Nsp9 (black) and triSer-Nsp9 mutant (red), calculated from the snapshots selected every 1 ns from a 200 ns MD trajectory. The two subunits of each homodimer are indicated by separate traces. The green traces represent the average pairwise RMSD calculated upon dimer-dimer superposition of all mutant snapshots with all wild-type snapshots. The conclusion inferred from DOSY is consistent with the MD simulation results. The average pairwise RMSD values, upon mutant versus wild-type dimer-dimer superposition, are larger than the single species counterparts at both Nterminal and C-terminal tracts, implying that wild type and mutated Nsp9 fluctuate around different conformers. As those tracts are involved at the inter-monomer interface, [20] the pairwise RMSD profile is consistent with difference in the dimerization extent between the species.

Figure S5. Overlay of the 15N-1H HSQC maps of SARS-CoV-2 Nsp9 at 298K (black contours) and 278 K (olive contours).

Figure S6. Fitting of the chemical shift changes observed at the peripheral residues of SARS-CoV-2 Nsp9 upon titration with 2NSP90. As reported in main text, besides the cross-peak loss observed with titrant increase, we could also detect progressive chemical shift changes. Those titration chemical shifts involve mostly the N-terminal and C-terminal residue signals, namely A8, L9, R111 and Q113 and a couple of other locations (C73, V76). This pattern is compatible with the intermediate exchange regime observed for all the other residues of Nsp9 and may arise for intrinsically mobile molecular locations where the chemical shift is effectively averaged by the local dynamics, leading to a very small difference between the limiting chemical shift values and matching therefore local fast exchange regime. The experimental data were fitted using a Hill-Langmuir model with equation:

$$
\Delta \delta = \frac{\Delta \delta_{max} \times c^n}{K_A^n + c^n}
$$

where ∆δ and ∆δmax are the observed and maximum chemical shift change, *c* is the nanobody concentration, *KA* is half occupation constant (related to the apparent dissociation constant) and *n* is the Hill coefficient. For A8, the fitting gave *n* = 2.9±0.6 (*p* value 1.5×10[−]3) and KA = 9.7 µM (*p* value 3.2×10[−]6), whereas for Q113, *n* = 4.0±0.5 (*p* value 3.7×10[−]5) and KA = 11 µM (*p* value 4.1×10[−]9). Albeit statistically significant, the fitting parameters should be regarded with some caution. This is expected due to difficulty of appreciating the limiting ∆δ values and the experimental titration errors with concentrations as small as 18 µM (constant for Nsp9) and 36 µM (maximum for 2NSP90).

Table S1. The ELISA tests were performed on both Nsp9 dissolved in ammonium acetate (NSP) and dissolved in DMSO/PBS (INS). Uncoated blocked wells served as negative control (blank) for ELISA. Ratios of ELISA readouts between wells coated with Nsp9 and uncoated wells were taken as estimates of each nanobody's binding strength towards the antigen. Nanobodies exhibiting Strong INS/control or Nsp9/control ratios above 10 are considered strong binders.

Tables S2. Amino acid sequence analysis of the Nanobodies using the ProtParam tool allows the prediction of several theorical proprieties. MW, Molecular Weight; pI, isoelectric point; ε, extinction coefficient. Production yields of the Nsp9-specific nanobodies in the expression vector pHEN6 are also included.

Table S3. Peak loss and attenuation in ¹⁵N-¹H HSQC spectra of NSP9 titrations with nanobodies

- (a) Although quite similar in the relative interactions, the two nanobodies exhibit slight differences for the involved epitopes, with 2NSP23 addressing first the surface encompassing the fragments 50- 53, 86-89, and 2NSP90 selecting only fragment 50-53.
- (b) The peaks with the steepest decrease in intensity were identified when the slope of their relative intensity attenuations was larger than the average attenuation slope increased by one standard deviation.