iScience, Volume *24*

Supplemental Information

Structural characterization of nonstructural protein 1 from SARS-CoV-2

Cameron Semper, Nobuhiko Watanabe, and Alexei Savchenko

Supplemental Information for Semper *et al*

Transparent Methods

Cloning

The portion of Orf1a encoding full-length Nsp1 was codon-optimized for *E. coli* expression, synthesized (Codex DNA) and cloned into the pMCSG53 expression vector at the SspI site via Gibson assembly.

>**cDNA sequence for SARS-CoV-2 Nsp1**
ATGGAAAGTCTGGTACCTGGGTTCAACGAGAAAACCCATGTTCAGCTGAGTTTACCG
GTTCTGCAAGTTCGTGATGTTCTGGTTCGCGGTTTTGGTGATAGCGTTGAAGAAGTTC
TGAGTGAAGCACGTCAACACCTGAAAGATGGTACATGCGGCTTAGTGGAAGTGGAA
AAAGGTGTTTTACCTCAGCTGGAACAGCCGTACGTGTTCATTAAACGCAGCGATGCA
AGAACAGCACCTCATGGTCATGTTATGGTTGAACTGGTGGCAGAACTGGAAGGTAT
CCAGTATGGTAGATCTGGTGAAACACTGGGTGTTTTAGTTCCGCATGTGGGCGAAAT
TCCTGTGGCATACCGTAAAGTGCTGCTGCGTAAAAATGGCAATAAAGGTGCAGGTG
GTCACAGCTATGGTGCCGATCTGAAAAGCTTTGATCTGGGCGATGAATTAGGTACAG
ATCCGTATGAGGACTTCCAGGAAAACTGGAACACCAAGCATAGTAGCGGTGTTACC
CGTGAATTAATGCGCGAACTGAATGGTGGTTAA

The fragment encompassing amino acid residues 13-127 of Nsp1 was PCR amplified using Phusion polymerase (NEB) and cloned into the same vector via ligation-independent cloning.

Primers for amplification of Nsp1$_{13\text{-}127}$

FWD - TACTTCCAATCCAATGCCCATGTTCAGCTGAGTTTACCG
REV - TTATCCACTTCCAATGTTAGCCATTTTTACGCAGCAG

Expression and purification of SARS-CoV2 Nsp1

Plasmids were transformed into the *E. coli* strain BL21(DE3)-Gold for protein expression. The same procedure was used for purification of full-length Nsp1 and Nsp1$_{13\text{-}127}$. Cells were grown at 37°C and 200 rpm to an OD600 of 0.8, cooled to 20°C then induced with 1 mM IPTG and incubated for 16 hours. Cells were harvested via centrifugation at 5000 x *g*, resuspended in

binding buffer (300 mM NaCl, 50 mM HEPES pH 7.5, 5 mM imidazole, 5% glycerol) and lysed via sonication. Lysates were centrifuged at 20,000 x *g* for 45 minutes at 4°C and the supernatant was incubated with Ni-NTA resin and rotated for 1 hour at 4°C. Nsp1 and $Nsp1_{13\text{-}127}$ were eluted in elution buffer (300 mM NaCl, 50 mM HEPES pH 7.5, 250 mM imidazole, 5% glycerol) then incubated with Tobacco-etch virus (TEV) protease overnight to cleave the N-terminal polyhistidine tag while dialyzing to remove imidazole. The proteins were then passaged over a second Ni-NTA to remove impurities. $Nsp1_{13\text{-}127}$ was immediately dialyzed into precrystallization buffer (300 mM NaCl, 10 mM HEPES pH 7.5), while full-length Nsp1was further purified via gel filtration using a Superdex75 column (300 mM NaCl, 10 mM HEPES pH 7.5).

Crystallization

Crystals of $Nsp1_{13\text{-}127}$ were grown at 298 K in 0.2 M sodium formate, 20% PEG3350 via the vapour diffusion sitting-drop method. Prior to data collection the crystals were soaked in 0.2 M sodium formate, 20% PEG3350, 30% glycerol and flash frozen in liquid nitrogen.

Data collection, structure determination and refinement

X-ray diffraction data of crystals of $Nsp1_{13\text{-}127}$ were collected at Advanced Photon Source Beamline 19 ID (Argonne, Illinoi USA) under cryo-stream at 93.15 K. Diffraction data were processed with HKL3000 suit (Minor et al., 2006). Initial phase for $Nsp1_{13\text{-}127}$ was obtained by molecular replacement with Molrep using NMR structure of SARS-CoV Nsp1 (PDB: 2HSX) as a search model (Winn et al., 2011). Subsequently, the initial electron density map was improved though density modification by parrot and the model was built using buccaneer (Winn et al., 2011). Final model was produced by the cycle of manual model building and refinement using Phenix.refine and Coot (Adams et al., 2010, Emsley and Cowtan, 2004). All geometry was

verified using Phenix validation tools (Ramachandran statistics: Favoured (97.1%), additionally allowed (2.9%), disallowed (0.0%)) and the wwPDB server.

Structure analysis and homology-modeling

Electrostatic potential surfaces were calculated using Pymol using APBS(Baker et al., 2001). Structure similarity searches of the Protein Data Bank were performed using the Dali server (Holm and Laakso, 2016).Secondary structure prediction of Nsp1 homologues was done using PSIPRED (Buchan and Jones, 2019). Homology modeling of MERS-CoV was performed by I-TASSER using the crystal structure of SARS-CoV-2 $Nsp1_{13-127}$ as the threading template (Yang and Zhang, 2015, Zhang et al., 2017). The model produced by I-TASSER was subjected to further energy minimization using the YASARA energy minimization server to improve stereochemical property (Krieger et al., 2009). Stereochemical properties of the final model was validated with Ramachandran plot using PROCHECK server and 90% of residues are in the favoured/additionally allowed region (Laskowksi, 1993). The model of the full-length SARS-CoV-2 Nsp1 was generated using the Robetta structural predication server with the SARS-CoV-2 $Nsp1_{13-127}$ structure defined as the template (Song et al., 2013).

Methods References

ADAMS, P. D., AFONINE, P. V., BUNKOCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L. W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-21.

BAKER, N. A., SEPT, D., JOSEPH, S., HOLST, M. J. & MCCAMMON, J. A. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A,* 98**,** 10037-41.

BUCHAN, D. W. A. & JONES, D. T. 2019. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res,* 47**,** W402-W407.

EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

HOLM, L. & LAAKSO, L. M. 2016. Dali server update. *Nucleic Acids Res,* 44**,** W351-5.
KRIEGER, E., JOO, K., LEE, J., LEE, J., RAMAN, S., THOMPSON, J., TYKA, M., BAKER, D. & KARPLUS, K. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins,* 77 Suppl 9**,** 114-22.
LASKOWKSI, R. A., MACARTHUR, M.W., MOSS, D. S, THORNTON, J.M. 1993. PROCHECK - a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26**,** 283-291.
MINOR, W., CYMBOROWSKI, M., OTWINOWSKI, Z. & CHRUSZCZ, M. 2006. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr,* 62**,** 859-66.
SONG, Y., DIMAIO, F., WANG, R. Y., KIM, D., MILES, C., BRUNETTE, T., THOMPSON, J. & BAKER, D. 2013. High-resolution comparative modeling with RosettaCM. *Structure,* 21**,** 1735-42.
WINN, M. D., BALLARD, C. C., COWTAN, K. D., DODSON, E. J., EMSLEY, P., EVANS, P. R., KEEGAN, R. M., KRISSINEL, E. B., LESLIE, A. G., MCCOY, A., MCNICHOLAS, S. J., MURSHUDOV, G. N., PANNU, N. S., POTTERTON, E. A., POWELL, H. R., READ, R. J., VAGIN, A. & WILSON, K. S. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr,* 67**,** 235-42.
YANG, J. & ZHANG, Y. 2015. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res,* 43**,** W174-81.
ZHANG, C., FREDDOLINO, P. L. & ZHANG, Y. 2017. COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. *Nucleic Acids Res,* 45**,** W291-W299.

Table S1. Amino acid sequence similarity to SARS-CoV-2 Nsp1 of Nsp1 proteins from human pathogenic coronaviruses, related to Figure 3

Protein Name	Organism	% Identity to SARS-CoV-2 Nsp1
MERS-CoV Nsp1	Middle East respiratory syndrome-related coronavirus	19
HCoV-OC43 Nsp1	Human coronavirus OC43	19
HCoV-HKU1 Nsp1	Human coronavirus HKU1	18
HCoV-229E Nsp1	Human coronavirus 229E	9
HCoV-NL63 Nsp1	Human coronavirus NL63	6

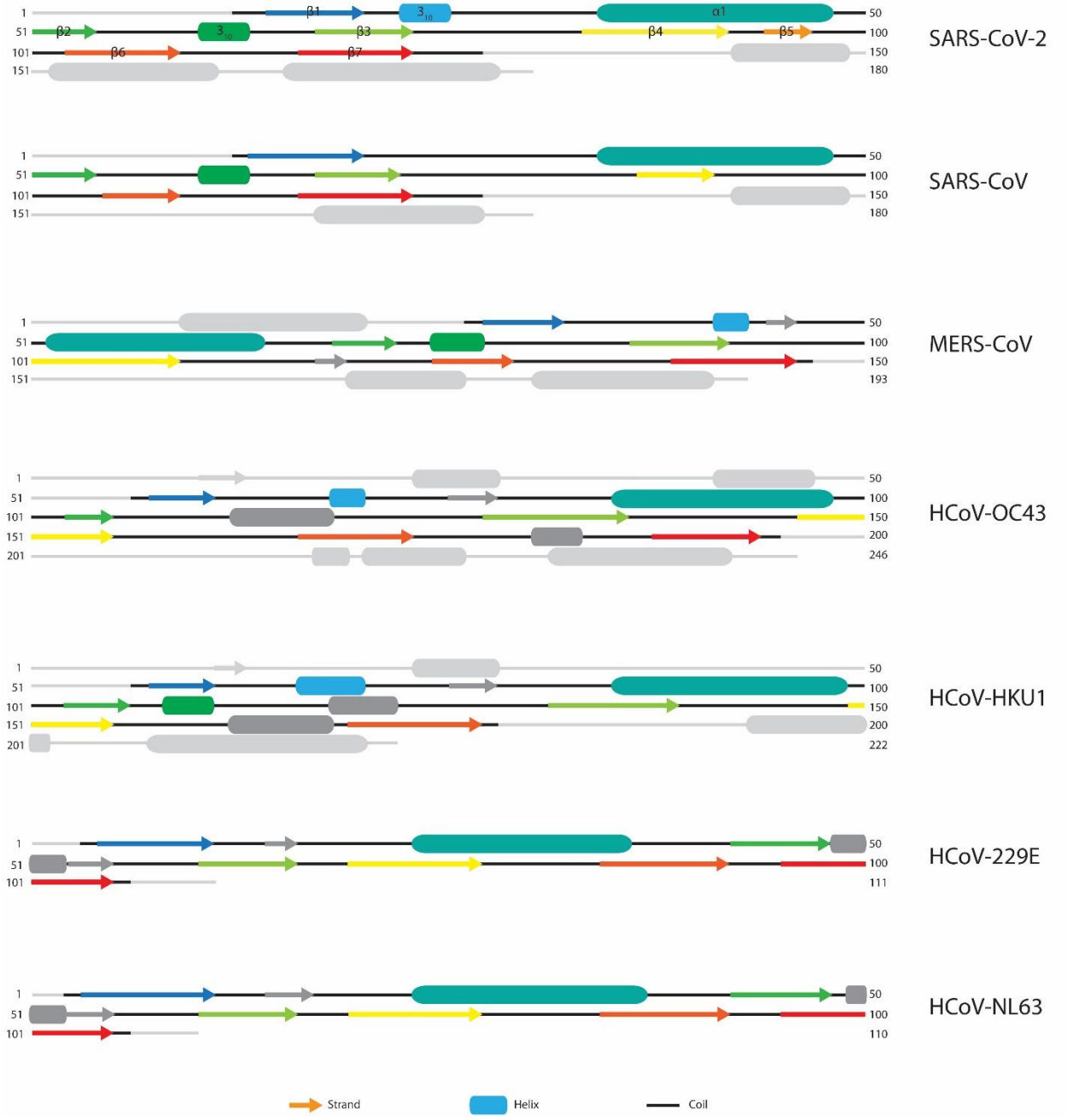

Fig S1. Secondary structure analysis of Nsp1 from human-infecting Coronaviruses, related to Figure 3. Within each sequence, the topological fingerprint containing elements that may facilitate formation of the capped β-barrel structure are coloured, while sequence outside of the fingerprint region is depicted in light grey. Key secondary structure elements observed in the SARS-CoV-2 $Nsp1_{13\text{-}127}$ crystal structure are labeled and are coloured accordingly in the ortholog sequences. Unique secondary structure features that lack an equivalent in the SARS-CoV-2 structure but that fall within the predicted β-barrel region are coloured as dark grey.