Supplementary Material

**Supplementary Tables**

**Supplementary Table 1. Protein sequences**

|  |  |  |
| --- | --- | --- |
| **Protein** | **GenBank ID** | **Sequence** |
| SARS-CoV-2 N | YP\_009724397 | MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALLLLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDDFSKQLQQSMSSADSTQA |
| sdAb-B6 | N/A | MAEVQLQASGGGLVQAGDSLRLSCVAVSGRTISTFAMGWFRQAPGKEREFVATINWSGSSARYADPVEGRFTISRDDAKNTVYLEMSSLKPGDSAVYYCASGRYLGGITSYSQGDFAPWGQGTQVTVSSAAALEHHHHHH |
| sdAb-C2 | N/A | MAEVQLQASGGGLVRPGGSLRLSCAASGFTFSSYAMMWVRQAPGKGLEWVSAINGGGGSTSYADSVKGRFTISRDNAKNTLYLQMNSLKPEDTAVYYCAKYQAAVHQEKEDYWGQGTQVTVSSAAALEHHHHHH |
| sdAb-E2 | N/A | MAEVQLQASGGGLVQAGGSLRLSCAASGRTDSTQHMAWFRQAPGKEREFVTAIQWRGGGTSYTDSVKGRFTISRDNAKNTVYLEMNSLKPEDTAVYYCATNTRWTYFSPTVPDRYDYWGQGTQVTVSSAAALEHHHHHH |
| sdAb-N3 | N/A | MAKVKLQQSGGGLVQTGGSLRLSCAPSGRTFSVNAMGWYRQAPGKQREPVALISTGGSTNYADSVKGRFTISRDNVKNTVYLQMNTLKPEDTAVYYCYTERWKPRGIERDWGQGTQVTVSSAAALEHHHHHH |

1Underlined text shows the N-terminal MA sequence remaining after signal peptide cleavage, and the C-terminal His6-tag used for purification of sdAbs.

**Supplementary Table 2. Data collection and refinement statistics**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data collection** | **sdAb-B6:N49-174** | **sdAb-C2:N49-174** | **sdAb-E2:N269-364** |
| Beamline | APS 24ID-C | APS 24ID-E | ALS 5.0.2 |
| Data collection date | March 24, 2021 | February 11, 2021 | May 14, 2021 |
| Resolution (Å) | 112 – 2.51 | 77 – 1.42 | 50 – 1.80 |
| Wavelength (Å) | 0.97918 | 0.97918 | 1.00004 |
| Space Group | C2221 | P21 | P212121 |
| Unit Cell Dimensions (a, b, c) Å | 99.61, 175.06, 111.96 | 46.67, 77.12, 71.93 | 75.85, 131.556, 140.05 |
| Unit cell Angles (α,β,γ) ° | 90, 90, 90 | 90, 95.14, 90 | 90, 90, 90 |
| *I*/σ (last shell) | 9.7 (0.7) | 25.7 (4.5) | 8.1 (0.4) |
| 1*R*sym (last shell) | 0.214 (4.143) | 0.037 (0.281) | 0.089 (5.417) |
| 2*R*meas (last shell) | 0.223 (4.305) | 0.041 (0.307) | 0.096 (5.884) |
| 3*R*pim (last shell) | 0.060 (1.161) | 0.016 (0.121) | 0.038 (2.282) |
| 4CC1/2 (last shell) | 0.993 (0.419) | 0.999 (0.976) | 0.999 (0.344) |
| Completeness (last shell) % | 100.0 (100.0) | 99.1 (94.4) | 99.7 (99.5) |
| Number of reflections | 463981 | 632696 | 850156 |
| *unique* | 33832 | 95122 | 129860 |
| Multiplicity (last shell) | 13.7 (13.6) | 6.7 (6.2) | 6.5 (6.5) |
| Wilson B-factor | 68.88 | 23.97 | 32.70 |
| **Refinement** |  |  |  |
| Resolution (Å) | 87.5 – 2.51 | 52.5 – 1.42 | 50 – 2.20 |
| No. of reflections | 33544 | 95054 | 64008 |
| *working* | 31948 | 90426 | 60713 |
| *free* | 1596 | 4628 | 3295 |
| 5*R*work (last shell) (%) | 22.64 (38.56) | 14.05 (16.89) | 23.94 (34.11) |
| 5*R*free (last shell) (%) | 27.32 (39.25) | 16.62 (19.81) | 27.14 (39.04) |
| **Structure/Stereochemistry** |  |  |  |
| No. of atoms | 10939 | 8295 | 11100 |
| *hydrogen* | 5340 | 3738 | 0 |
| *solvent* | 0 | 672 | 862 |
| *ligand* | 0 | 15 | 5 |
| r.m.s.d. bond lengths (Å) | 0.011 | 0.008 | 0.003 |
| r.m.s.d. bond angles (°) | 1.259 | 0.964 | 0.683 |
| Average B-factor | 82.89 | 22.58 | 51.52 |
| Ramachandran favored/allowed (%) | 97.37/100 | 98.37/100 | 99.61/100 |
| MolProbity Clashscore/Overall score | 4.94/1.68 | 2.49/1.03 | 4.30/1.40 |
| 6Protein Data Bank ID | 7R98 | 7N0R | 7N0I |
| 7SBGrid Data Bank ID | 837 | 836 | 835 |

1*R*sym = ∑∑j|*I*j – 〈*I*〉|/∑*I*j, where *I*j is the intensity measurement for reflection j and 〈*I*〉 is the mean intensity for multiply recorded reflections.

2*R*meas = ∑h [ √(*n*/(*n*-1)) ∑j [*I*hj - 〈*I*h〉] / ∑hj 〈*I*h〉 where *I*hj is a single intensity measurement for reflection h, 〈*I*h〉 is the average intensity measurement for multiply recorded reflections, and *n* is the number of observations of reflection h.

3*R*pim = ∑h [1/(/*n* - 1)]1/2 ∑i| <*I*h> - *I*hi|/∑h ∑i *I*hi where *I*hi is a single intensity measurement for reflection h, 〈*I*h〉 is the average intensity measurement for multiply recorded reflections, and *n* is the number of observations of reflection h.

4CC1/2 is the Pearson correlation coefficient between the average measured intensities of two randomly-assigned half-sets of the measurements of each unique reflection (Karplus and Diederichs, 2012). CC1/2 is considered significant above a value of ~0.15.

5*R*work, free = ∑||Fobs| – |Fcalc||/|Fobs|, where the working and free *R*-factors are calculated using the working and free reflection sets, respectively.

6Coordinates and structure factors have been deposited with the Protein Data Bank (http://www.pdb.org) with the noted accession codes.

7Diffraction data have been deposited with the SBGrid Data Bank (http://data.sbgrid.org) with the noted accession codes.

**Supplementary Table 3. N protein mutations in 489,605 SARS-CoV-2 genomic samples**

*See attached Excel spreadsheet*

## Supplementary Figures

A computer screen capture

Description automatically generated with medium confidence

**Supplementary Figure 1. sdAbs target individual domains of SARS-CoV-2 N.**

(a) Ni2+ pulldown with C-terminally His6-tagged sdAb-N3 and seven truncations of SARS-CoV-2. Positive interactions are shown with blue arrowheads. Schematic (bottom) indicates that sdAb-N3 binds the C-terminal IDR of N (residues 365-419).

(b) Ni2+ pulldown with C-terminally His6-tagged sdAb-B6 and seven truncations of SARS-CoV-2. Positive interactions are shown with pink arrowheads. Schematic (bottom) indicates that sdAb-B6 binds the RNA-binding domain of N (residues 49-174).

(c) Ni2+ pulldown with C-terminally His6-tagged sdAb-C2 or sdAb-E2 and seven truncations of SARS-CoV-2 (control pulldowns without sdAbs shown at left). Positive interactions are shown with pink (C2) or blue (E2) arrowheads. Schematic (bottom) indicates that sdAb-C2 binds the RNA-binding domain of N (residues 49-174), and sdAb-E2 binds the C-terminal dimerization domain of N (residues 247-364).

A screenshot of a computer

Description automatically generated with medium confidence

**Supplementary Figure 2. N protein variation and recognition by sdAbs.**

(a) Histogram showing mutation frequency of N protein in a dataset of 489,605 SARS-CoV-2 genomes (see **Materials and Methods**). N protein domains are labeled, with NRBD highlighted in pink and NCTD highlighted in blue. Red arrowheads indicate positions mutated in one or more variants of concern: B.1.1.7 (D3L, R203K, G204R, S235F), B.1.351 (T205I), B.1.427 (T205I), B.1.429 (T205I), B.1.526 (P199L, M234I), B.1.526.1 (T205I, M234I), B.1.526.2 (P13L, S202R), B.1.617 (del204, del215), B.1.617.1 (R203M, D377Y), B.1.617.2 (D63G, R203M, D377Y), B.1.617.3 (P67S, R203M, D377Y), P.1 (P80R, R203K, G204R), and P.2 (A119S, R203K, G204R, M234I).

(b) N protein variation (black bars, left y axis) versus residue-by-residue buried surface area for the NRBD:sdAb-C2 (orange bars, right y axis) and NRBD:sdAb-B6 (blue bars, right y axis) complexes.

(c) N protein variation (black bars, left y axis) versus residue-by-residue buried surface area for the NCTD:sdAb-E2 complex (blue bars, right y axis). NCTD chain A is shown in dark blue, and chain B in light blue.

A picture containing text, indoor, monitor, screenshot

Description automatically generated

**Supplementary Figure 3. Single-domain antibodies affect RNA binding by SARS-CoV-2 N**

(a) Electrophoretic mobility shift assays (EMSA) showing binding of 7mer single-stranded RNA by full-length N protein alone (black boxes) and mixed 1:1 with sdAb-B6 (blue) or sdAb-C2 (orange). Protein concentrations (left to right): 0, 0.15625, 0.3125, 0.625, 1.25, 2.5, and 5 μM.

(b) Binding curves for N protein binding 7mer single-stranded RNA. All measurements were performed in triplicate, and binding curves were fit using a cooperative binding model with a Hill coefficient of 5 (see **Materials and Methods**). Data points at 5 μM protein concentration (outline symbols) were not used for curve fitting because of high error.

(c) EMSA showing binding of 7mer double-stranded RNA by full-length N protein alone (black boxes) and mixed 1:1 with sdAb-B6 (blue) or sdAb-C2 (orange).

(d) Binding curves for N protein binding 7mer double-stranded RNA.