1

2

3

Supplementary Information for

5

Structural insights into ribonucleoprotein dissociation by nucleocapsid

protein interacting with non-structural protein 3 in SARS-CoV-2

8

Xincheng Ni, Yinze Han, Renjie Zhou, Yanmei Zhou, Jian Lei*
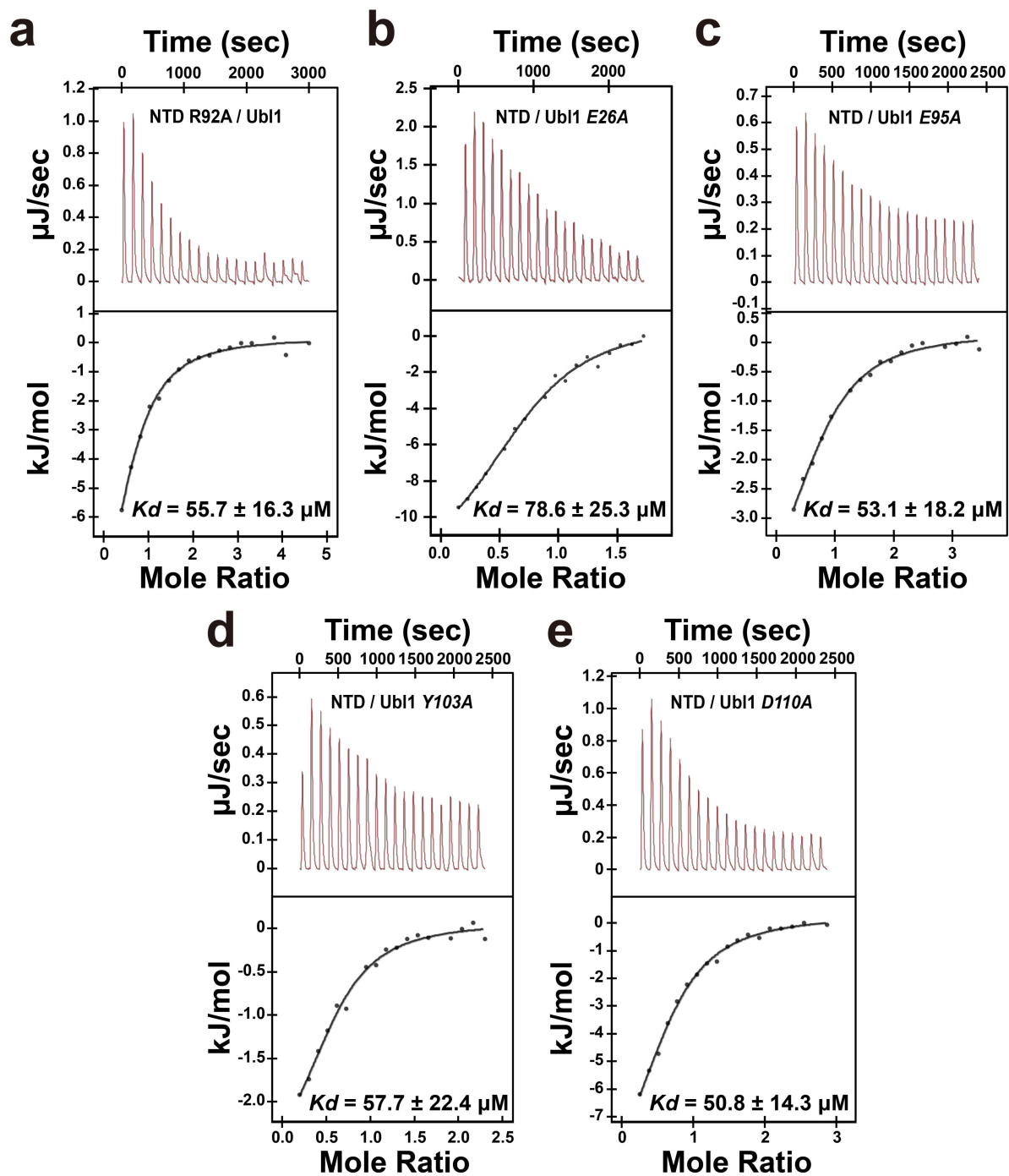
*Corresponding author. Email: leijian@scu.edu.cn

11

This file includes:

Supplementary Figures 1–6

Supplementary Tables 1–3

References 1–3

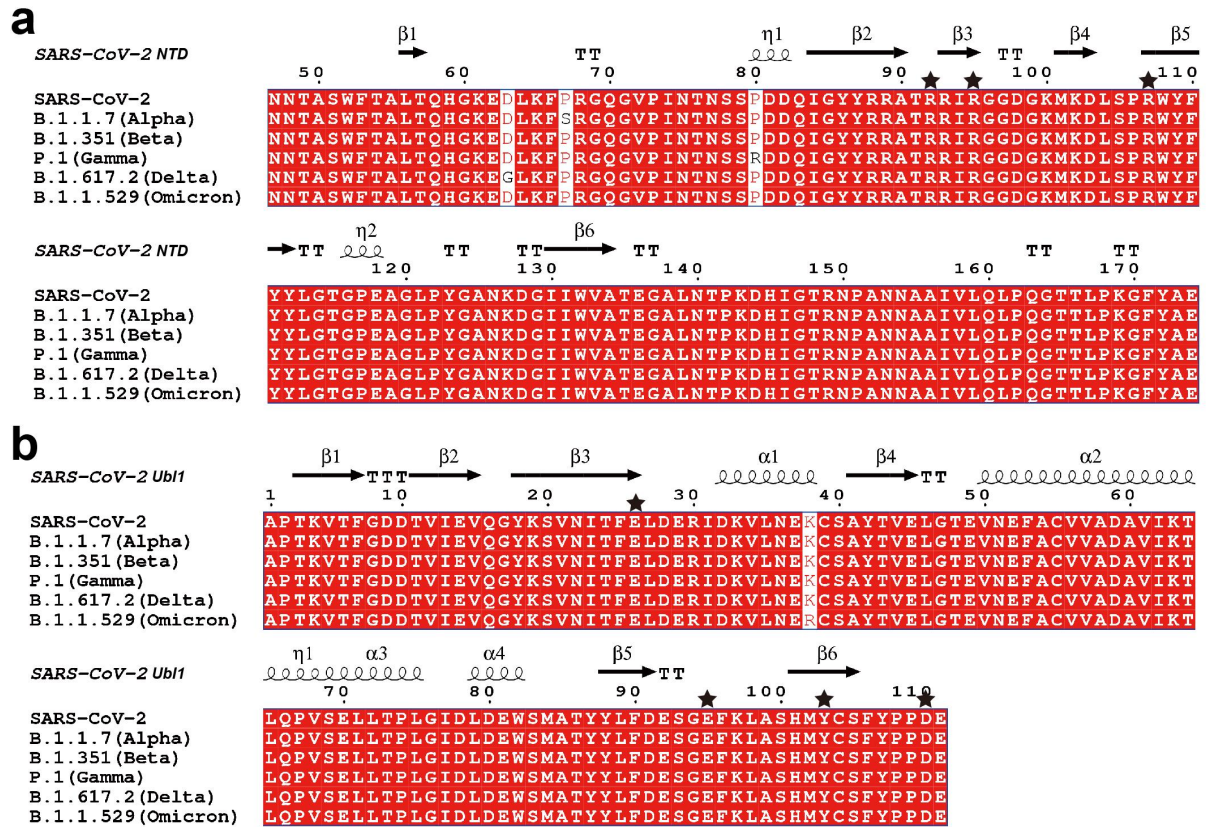16

Supplementary Fig. 1. Affinity measurements of SARS-CoV-2 NTD and Ubl1 mutations.

a Affinity between NTD R92A and Ubl1. **b-e** Affinity between NTD and Ubl1 *E26A* (**b**), *E95A* (**c**), *Y103A* (**d**), and *D110A* (**e**). The raw calorimetric curve is shown in the top panel, while the fitted binding isotherm curve is displayed in the bottom panel. *Kd*: dissociation constant.

a

SARS-CoV-2 NTD — secondary structure: β1, TT, η1, β2, β3, TT, β4, β5 (positions 50–110)

SARS-CoV-2	NNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYF
B.1.1.7(Alpha)	NNTASWFTALTQHGKEDLKFSRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYF
B.1.351(Beta)	NNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYF
P.1(Gamma)	NNTASWFTALTQHGKEDLKFPRGQGVPINTNSSRDDQIGYYRRATRRIRGGDGKMKDLSPRWYF
B.1.617.2(Delta)	NNTASWFTALTQHGKEGLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYF
B.1.1.529(Omicron)	NNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYF

SARS-CoV-2 NTD — secondary structure: TT, η2, TT, TT, β6, TT (positions 120–170)

SARS-CoV-2	YYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAE
B.1.1.7(Alpha)	YYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAE
B.1.351(Beta)	YYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAE
P.1(Gamma)	YYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAE
B.1.617.2(Delta)	YYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAE
B.1.1.529(Omicron)	YYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAE

b

SARS-CoV-2 Ubl1 — secondary structure: β1, TTT, β2, β3, α1, β4, TT, α2 (positions 1–60)

SARS-CoV-2	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKT
B.1.1.7(Alpha)	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKT
B.1.351(Beta)	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKT
P.1(Gamma)	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKT
B.1.617.2(Delta)	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKT
B.1.1.529(Omicron)	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNERCSAYTVELGTEVNEFACVVADAVIKT

SARS-CoV-2 Ubl1 — secondary structure: η1, α3, α4, β5, TT, β6 (positions 70–110)

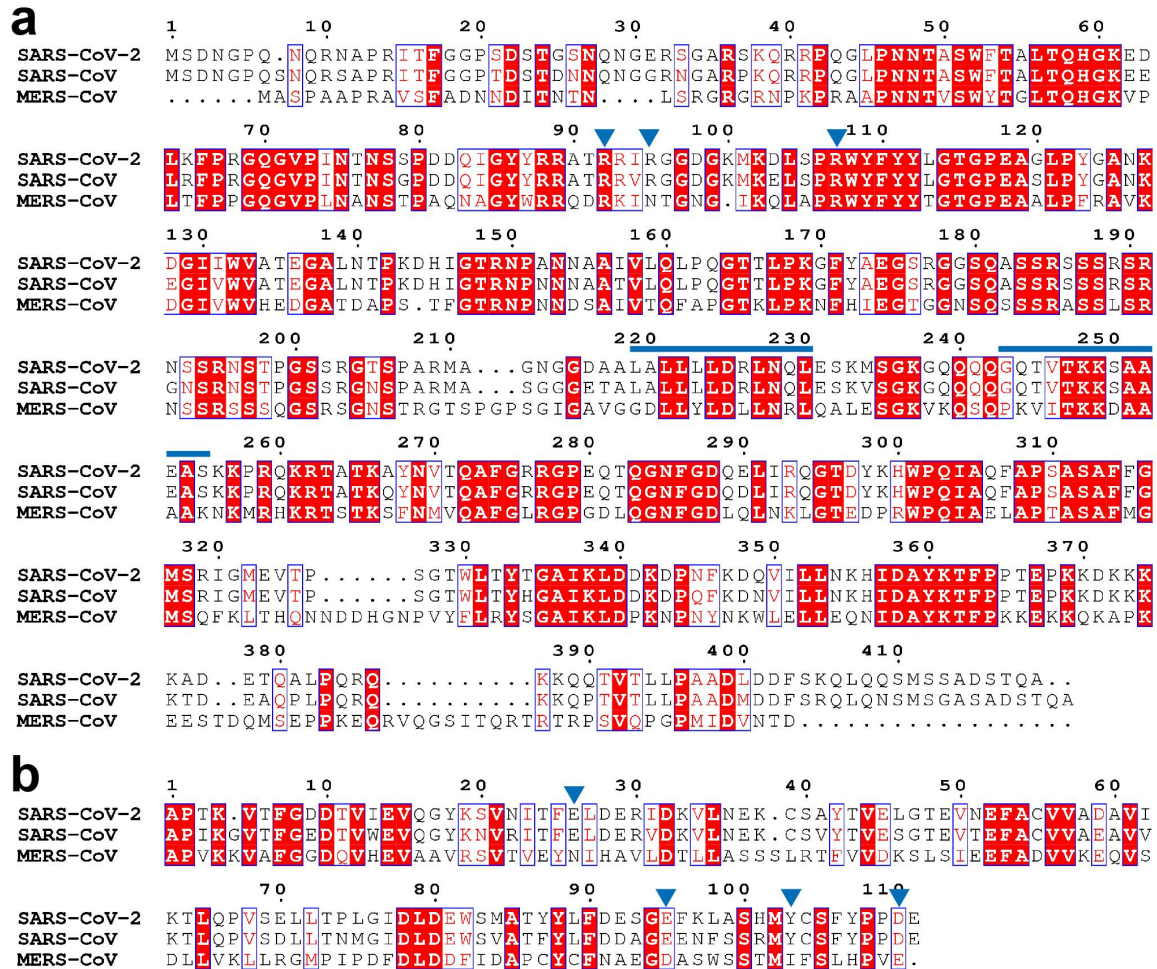|	
SARS-CoV-2	LQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE
B.1.1.7(Alpha)	LQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE
B.1.351(Beta)	LQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE
P.1(Gamma)	LQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE
B.1.617.2(Delta)	LQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE
B.1.1.529(Omicron)	LQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE

Supplementary Fig. 2. Structure-based multiple sequence alignment of NTD (a) and Ubl1 (b) in SARS-CoV-2 and its variant strains. The corresponding sequence accession numbers are: SARS-CoV-2, GenBank: NC_045512.2; B.1.1.7, GenBank: MZ622698; B.1.351, GenBank: MZ433432; P.1, GenBank: MZ477759; B.1.617.2, GenBank: OK091006; B.1.1.529, GenBank: OL672836. The secondary structures of SARS-CoV-2 NTD (PDB code: 7VNU) and Ubl1 (chain B) in NTD-Ubl1 complex (PDB code: 7WZO) are indicated. Residues R92, R95, R107 of NTD and *E26*, *E95*, *Y103*, *D110* of Ubl1 are labeled with black stars. Figures (**a** and **b**) were prepared by the program ESPript[1].

a

```
            1         10        20        30        40        50        60
SARS-CoV-2  MSDNGPQ.NQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALTQHGKED
SARS-CoV    MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNGARPKQRRPQGLPNNTASWFTALTQHGKEE
MERS-CoV    ......MASPAAPRAVSFADNNDITNTN....LSRGRGRNPKPRAAPNNTVSWYTGLTQHGKVP

            70        80        90        100       110       120
SARS-CoV-2  LKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANK
SARS-CoV    LRFPRGQGVPINTNSGPDDQIGYYRRATRRVRGGDGKMKELSPRWYFYYLGTGPEASLPYGANK
MERS-CoV    LTFPPGQGVPLNANSTPAQNAGYWRRQDRKINTGNG.IKQLAPRWYFYYTGTGPEAALPFRAVK

            130       140       150       160       170       180       190
SARS-CoV-2  DGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAEGSRGGSQASSRSSSRSR
SARS-CoV    EGIVWVATEGALNTPKDHIGTRNPNNNAATVLQLPQGTTLPKGFYAEGSRGGSQASSRSSSRSR
MERS-CoV    DGIVWVHEDGATDAPS.TFGTRNPNNDSAIVTQFAPGTKLPKNFHIEGTGGNSQSSSRASSLSR

            200       210       220       230       240       250
SARS-CoV-2  NSSRNSTPGSSRGTSPARMA...GNGGDAALALLLLDRLNQLESKMSGKGQQQQGQTVTKKSAA
SARS-CoV    GNSRNSTPGSSRGNSPARMA...SGGGETALALLLLDRLNQLESKVSGKGQQQQGQTVTKKSAA
MERS-CoV    NSSRSSSQGSRSGNSTRGTSPGPSGIGAVGGDLLYLDLLNRLQALESGKVKQSQPKVITKKDAA

            260       270       280       290       300       310
SARS-CoV-2  EASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFG
SARS-CoV    EASKKPRQKRTATKQYNVTQAFGRRGPDLQGNFGDQDLIRQGTDYKHWPQIAQFAPSASAFFG
MERS-CoV    AAKNKMRHKRTSTKSFNMVQAFGLRGPGDLQGNFGDLQLNKLGTEDPRWPQIAELAPTASAFMG

            320       330       340       350       360       370
SARS-CoV-2  MSRIGMEVTP......SGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKK
SARS-CoV    MSRIGMEVTP......SGTWLTYHGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDKKK
MERS-CoV    MSQFKLTHQNNDDHGNPVYFLRYSGAIKLDPKNPNYNKWLELLEQNIDAYKTFPKKEKKQKAPK

            380       390       400       410
SARS-CoV-2  KAD..ETQALPQRQ.........KKQQTVTLLPAADLDDFSKQLQQSMSSADSTQA..
SARS-CoV    KTD..EAQPLPQRQ.........KKQPTVTLLPAADMDDFSRQLQNSMSGASADSTQA
MERS-CoV    EESTDQMSEPPKEQRVQGSITQRTRTRPSVQPGPMIDVNTD.................
```

b

```
            1         10        20        30        40        50        60
SARS-CoV-2  APTK.VTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK.CSAYTVELGTEVNEFACVVADAVI
SARS-CoV    APIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNEK.CSVYTVESGTEVTEFACVVAEAVV
MERS-CoV    APVKKVAFGGDQVHEVAAVRSVTVEYNIHAVLDTLLASSSLRTFVVDKSLSIEEFADVVKEQVS

            70        80        90        100       110
SARS-CoV-2  KTLQPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDE
SARS-CoV    KTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGEENFSSRMYCSFYPPDE
MERS-CoV    DLLVKLLRGMPIPDFDLDDFIDAPCYCFNAEGDASWSSTMIFSLHPVE.
```

Supplementary Fig. 3. Multiple sequence alignment of N protein (**a**) and Ubl1 (**b**) among SASRS-CoV-2, SARS-CoV and MERS-CoV. The sequence accession numbers are: SARS-CoV-2, GenBank: NC_045512.2; SARS-CoV, GenBank: NC_004718.3; MERS-CoV, GenBank: NC_019843.3. Key residues for SARS-CoV-2 NTD-Ubl1 interaction are marked by blue-triangle arrows. Two Ubl1-binding motifs[2] ([219]LALLLLDRLNQL[230] and [243]GQTVTKKSAAEAS[255]) of SARS-CoV-2 N-LKR are indicated by the blue lines. Figures (**a** and **b**) were generated using the program ESPript[1].
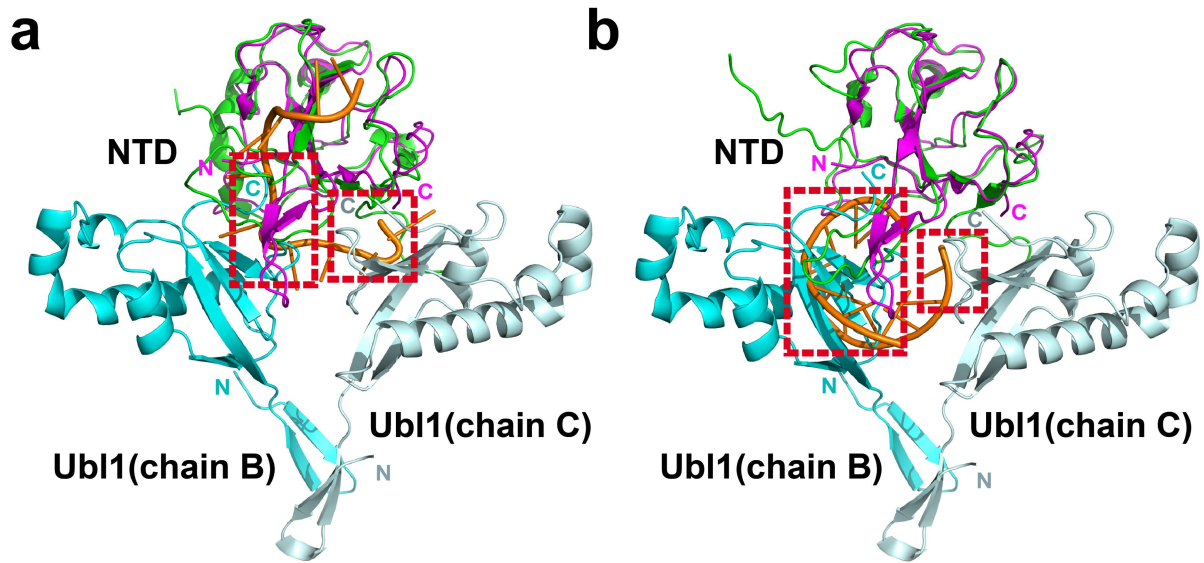
Supplementary Fig. 4. Overlapping region between Ubl1- and (ss/ds) RNA- binding site of SARS-CoV-2 N protein. a Superimposing the NTD-ssRNA complex structure[3] (PDB code: 7ACT) to our NTD-Ubl1 complex structure (PDB code: 7WZO). **b** Superimposing the NTD-dsRNA complex structure[3] (PDB code: 7ACS) to our crystal structure. NTD of the NTD-Ubl1 complex is shown in magenta. While NTD in the NTD-(ds/ss) RNA complex is displayed in green. Chain-B Ubl1 and chain-C Ubl1 are colored in cyan and light blue, respectively. The N- and C- termini of NTD and Ubl1s are labeled in the corresponding colors. Both the ssRNA and dsRNA are colored in orange. The overlapping regions between Ubl1s and (ss/ds) RNA are indicated by red dashed boxes. Figures (**a** and **b**) were prepared using PyMOL (https://pymol.org).
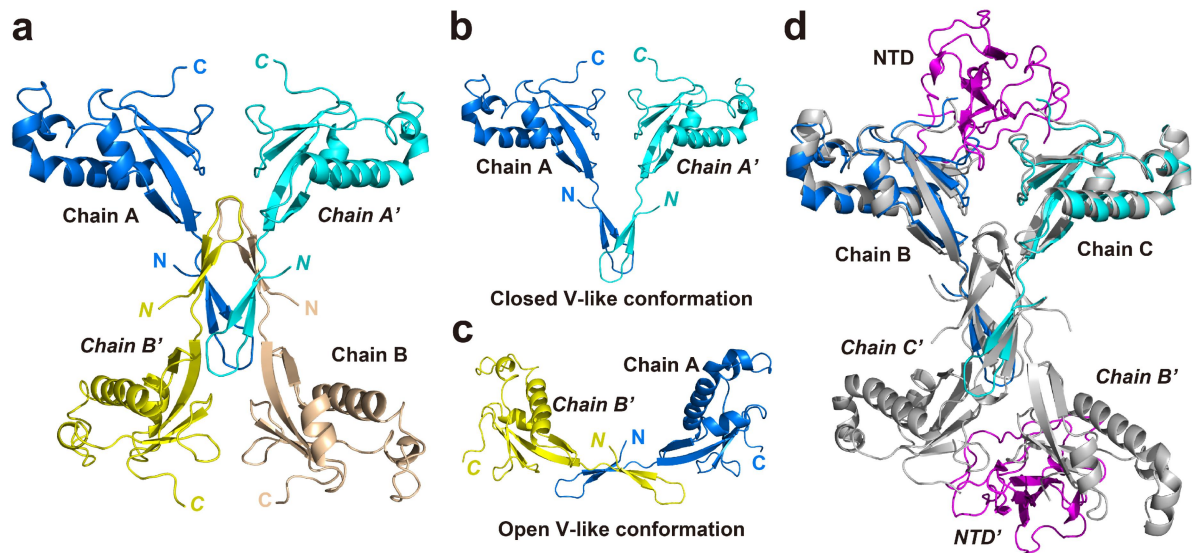
Supplementary Fig. 5. Two different conformations of the potential dimeric Ubl1 in the crystal structure. a A "closed V-like" (Chain A and *A′*) and an "open V-like" (Chain A and *B′*) conformations forming by the symmetry-related Ubl1s in the crystal structure of SARS-CoV-2 Ubl1 (PDB code: 7KAG). Ubl1 chain A, chain B, and their symmetrical protomers are colored in blue, brown, cyan, and yellow, respectively. **b** The "closed V-like" conformation of dimeric Ubl1. **c** The "open V-like" conformation of dimeric Ubl1. **d** Superimposing the dimeric "closed V-like" Ubl1 structure to our NTD-Ubl1 complex structure (PDB code: 7WZO). The Ubl1s in the NTD-Ubl1 complex and their symmetrical molecules are colored in gray. The NTD and its symmetrical protomer are displayed in magenta. Figures (**a-d**) were prepared using PyMOL (https://pymol.org).

a
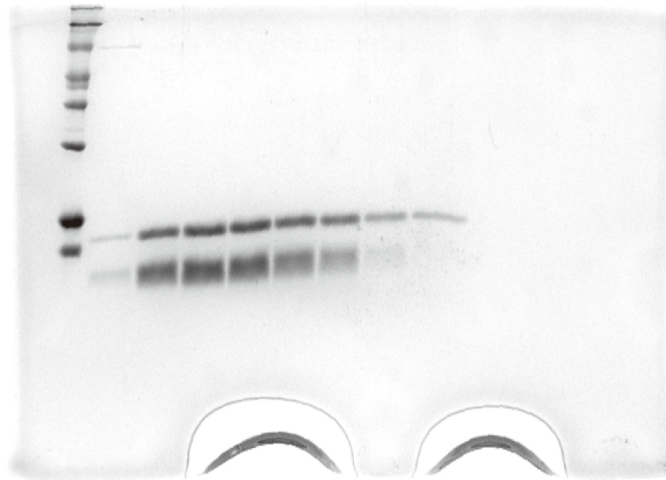
b
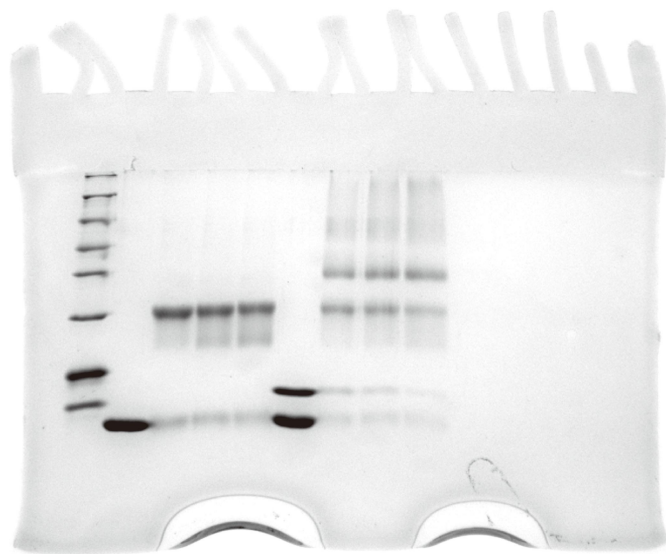
c
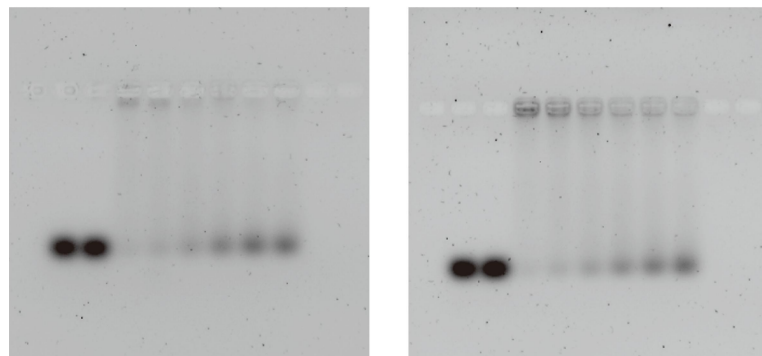
17mer ssRNA　　　　　**17mer ssDNA**

67　**Supplementary Fig. 6. Uncropped Gels. a** Unedited gel used in Figure 3a. **b** Unedited gel

68　used in Figure 3d. **c** Unedited gels used in Figure 6a.

69

70 **Supplementary Table 1. Primers used for SARS-CoV-2 *N*, *Nsp3* and *truncated Nsp3* in**

71 **the fluorescence assay.**

72

Gene	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)
N protein (1–419)	ACTCACTATAGGGAGACCCAT GTCTGATAATGGAC	CGTATGGGTACCCGGGCAT GGCCTGAGTTGAGTCAGCA
Nsp3.1 (1–654)	TCTATATAAGCAGAGCTCGTA TGGCGCCGACCAAAGTGA	CGGCACCTTCAGGCTACGC ATGTAACGAGCCGCTTCC
Nsp3.2 (649–1304)	GAGCTCCTTATGGCCTCTAGA CGTAGCCTGAAGGTGCCG	CAGGGTTTTCAGACCCAGC ACACGGCTCAGCTCGTT
Nsp3.3 (1299–1945)	CGAGCTCCTTATGGCCTCTAG ACTCGAGGGTGGAAAGCTTCT GGGTCTGAAAACCCTG	GGAATTCGGCCACCCTTCA GAGCGAT
Nsp3$^{\Delta 111}$ (112–645)	TCTATATAAGCAGAGCTCGTA TGGACGAGGAAGAGGGTGAC	CGGCACCTTCAGGCTACGC ATGTAACGAGCCGCTTCC
Nsp3$^{\Delta 168}$ (169–645)	TCTATATAAGCAGAGCTCGTA TGACCGTGGGCCAGCAGGAT	CGGCACCTTCAGGCTACGC ATGTAACGAGCCGCTTCC
Ubl1 (1–111)	TCTATATAAGCAGAGCTCGTA TGGCGCCGACCAAAGTGA	CTTGCTCACCAGAATTCGT TCATCCGGGTAGAAG

73

Supplementary Table 2. Primers used for SARS-CoV-2 *N*, *N truncations* and *Ubl1*.

Gene	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)
N protein (1–419)	CTAGCTAGCGAAAACCTGTATTTTCAGGGCATGTCTGATAATGGACCCCAAAATC	CGCGGATCCTTAGGCCTGAGTTGAGTCAGCACTGC
N-arm–NTD–LKR (1–247)	CTAGCTAGCATGTCTGATAATGGACCCCAAAATC	CGCGGATCCTTAAGTGACAGTTTGGCCTTGTTG
NTD–LKR (47–247)	CTAGCTAGCAATAATACTGCGTCTTGGTTCACCGC	CGCGGATCCTTAAGTGACAGTTTGGCCTTGTTGTTG
NTD (47–174)	CTAGCTAGCAATAATACTGCGTCTTGGTTCACC	CGCGGATCCTTATTCTGCGTAGAAGCCTTTTGG
CTD–C-tail (248–419)	CTAGCTAGCAAGAAATCTGCTGCTGAGGCTTC	CCGCTCGAGTTAGGCCTGAGTTGAGTCAGCAC
Ubl1 (1–111)	CGCCATATGGCGCCGACCAAAGTGACCTTCGGTG	CCGCTCGAGTTATTATTCATCCGGCGGGTAGAAGCTGCAGTAC

Supplementary Table 3. Primers used for SARS-CoV-2 *NTD* and *Ubl1* mutations.

Gene	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)
NTD R92A	TACCGAAGAGCTACCGCGCG AATTCGTG	CGCGGTAGCTCTTCGGTAGT AGCCAAT
Ubl1 E26A	GAGCGTGAACATCACCTTCGC GCTGGATGAACGTAT	TGTCGATACGTTCATCCAGC GCGAAGGTGATGTTCA
Ubl1 E95A	TTCGATGAGAGCGGCGCGTTC AAGCTGG	CGCGCCGCTCTCATCGAACA GGTAGTAG
Ubl1 Y103A	CAAGCTGGCGAGCCACATGG CCTGCAGCTTCTACCC	CCGGCGGGTAGAAGCTGCA GGCCATGTGGCTCGCCA
Ubl1 D110A	CGCCATATGGCGCCGACCAA AGTGACCTTCGGTG	CCGCTCGAGTTATTCCGCCG GCGGGTAGAAGCTGCAG

References

1. Gouet, P., Courcelle, E., Stuart, D. I., & Métoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308 (1999).

2. Bessa, L. M. et al. The intrinsically disordered SARS-CoV-2 nucleoprotein in dynamic complex with its viral partner nsp3a. *Sci. Adv.* **8**, eabm4034 (2022).

3. Dinesh, D. C. et al. Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. *PLoS Pathog.* **16**, e1009100 (2020).