Insights into the specificity for the interaction of the promiscuous SARS-CoV-2 nucleocapsid protein N-terminal domain with deoxyribonucleic acids

Icaro Putinhon Caruso[1,2,9,*], Vitor dos Santos Almeida[3,9], Mariana Juliani do Amaral[4,5,9], Guilherme Caldas de Andrade[1,3,9], Gabriela Rocha de Araújo[1,3,9], Talita Stelling de Araújo[1,5,9], Jéssica Moreira de Azevedo[1,5,9], Glauce Moreno Barbosa[1,3,9], Leonardo Bartkevihi[1,3,9], Peter Reis Bezerra[1,3,9], Katia Maria dos Santos Cabral[1,5,9], Isabella Otênio de Lourenço[2,9], Clara L. F. Malizia-Motta[6,9], Aline de Luna Marques[1,7,9], Nathane Cunha Mebus-Antunes[1,9], Thais Cristtina Neves-Martins[1,9], Jéssica Maróstica de Sá[2,9], Karoline Sanches[2,9], Marcos Caique Santana-Silva[1,7,9], Ariana Azevedo Vasconcelos[1,3,9], Marcius da Silva Almeida[1,5,9], Gisele Cardoso de Amorim[1,7,9], Cristiane Dinis Anobom[3,6,9], Andrea T. Da Poian[1,9], Francisco Gomes-Neto[3,8,9], Anderson S. Pinheiro[6,9], Fabio C. L. Almeida[1,3,9,*]

[1]Institute of Medical Biochemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil,

[2]Multiuser Center for Biomolecular Innovation (CMIB), Department of Physics, São Paulo State University (UNESP), São José do Rio Preto, Brazil;

[3]National Center of Nuclear Magnetic Resonance (CNRMN), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil;

[4]Faculty of Pharmacy, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil;

[5]Protein Advanced Biochemistry (PAB), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil;

[6]Department of Biochemistry, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil;

[7]Multidisciplinary Center for Research in Biology (NUMPEX), Campus Duque de Caxias Federal University of Rio de Janeiro, Duque de Caxias, Brazil.

[8]Laboratory of Toxinology, Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, Brazil

[9]Rio BioNMR Network, Rio de Janeiro, Brazil

***To whom correspondence should be addressed:**

e-mail: falmeida@bioqmed.ufrj.br

e-mail: icaro.caruso@unesp.br

SUPPLEMENTARY MATERIAL

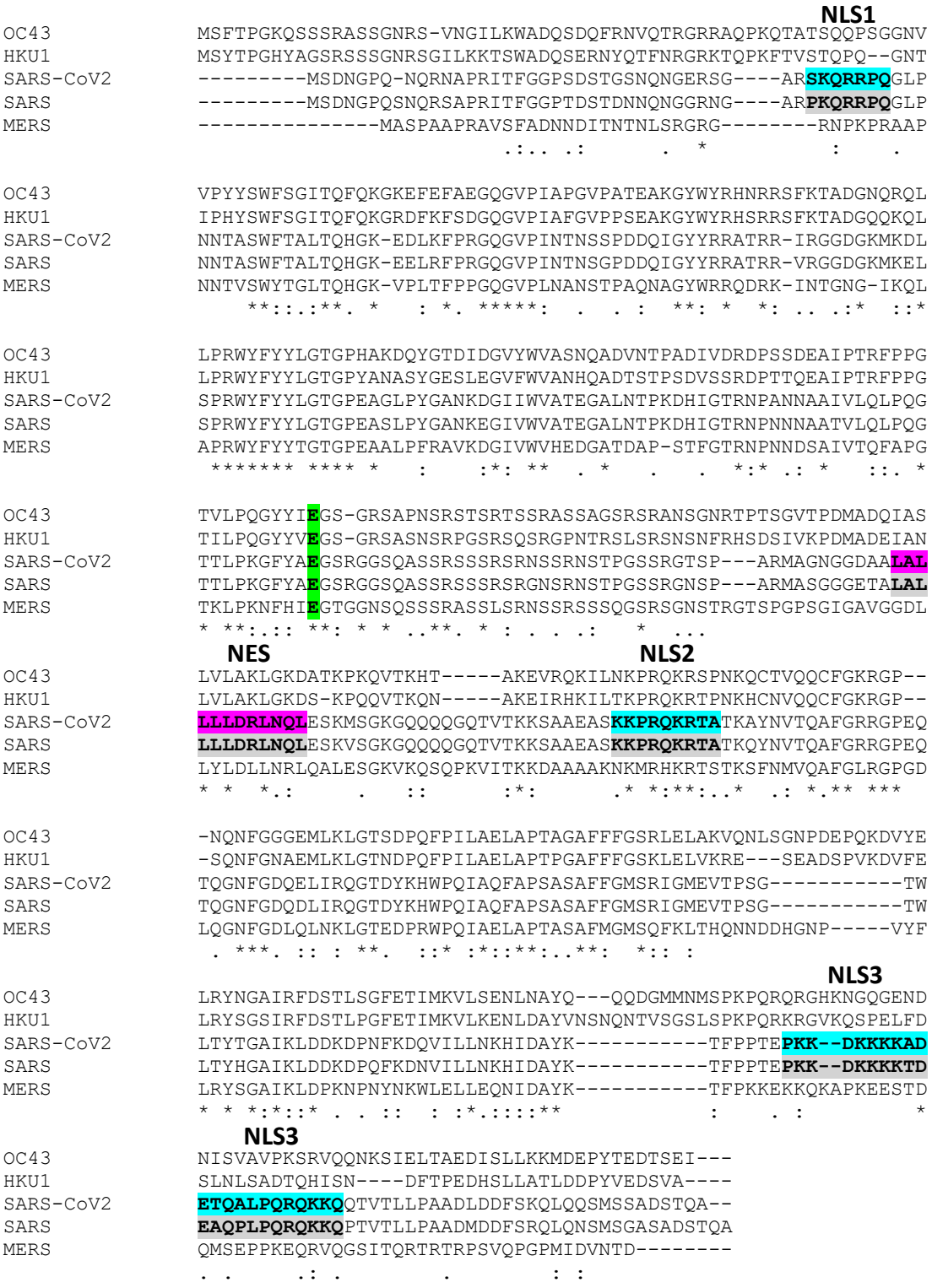

Figure S1. Sequence alignment of the N protein from SARS-CoV-2, SARS-CoV, CoV-MERS, CoV-HKU1, and CoV-OC43 performed using the webserver ClustalW [2]. Sequence homology between hCoV N protein is denoted by asterisk (*), colon (:), and period (.). The

asterisk indicates fully conserved residue, colon represents conservation between groups of strongly similar properties, and period indicates conservation between groups of weakly similar properties. The residue sequences of the SARS-CoV-2 N-NTD and its SR-motif are denoted in red and blue, respectively. The nuclear localization signals are indicated by NLS1, NLS2, and NLS3 with the residue sequence for SARS-CoV-2 and SARS-CoV highlighted in cyan and gray, respectively. The nuclear export signal is indicated by NES with the residue sequence for SARS-CoV-2 and SARS-CoV highlighted in magenta and gray, respectively. The residue E174 is denoted in green.

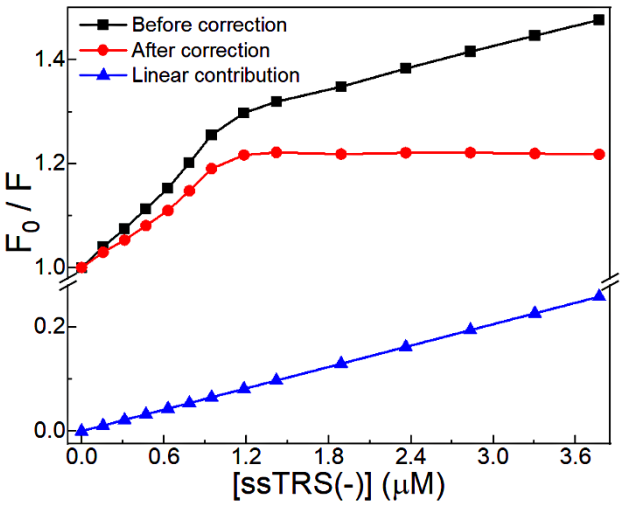

Figure S2. Removal of the upward linear contribution (blue triangle) from the fluorescence titration curves (black square) resulting in the corrected binding isotherm (red circle).

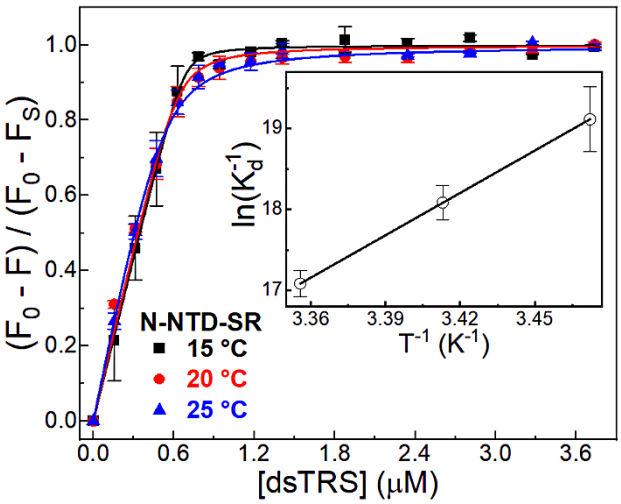

Figure S3. Fluorescence quenching changes of N-NTD as a function of the dsTRS concentration in 20 mM Bis-Tris buffer (pH 6.5) at 15, 20, and 25 °C. Each point on the binding isotherm represents the average and standard error calculated from duplicate measurements. The continuous lines denote the theoretical curves globally adjusted to the

experimental data. The inset shows the van't Hoff plot determined the enthalpy change value for the N-NTD/ssTRS(–) complex.

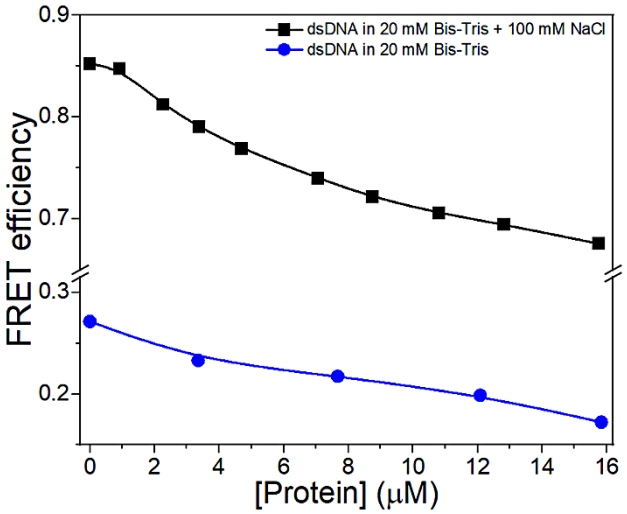

Figure S4. FRET efficiency as a function of the total concentration of N-NTD-SR in 20 mM Bis-Tris (pH 6.5) containing 0 mM (blue circles) and 100 mM NaCl (black squares) for dsDNA TRS.

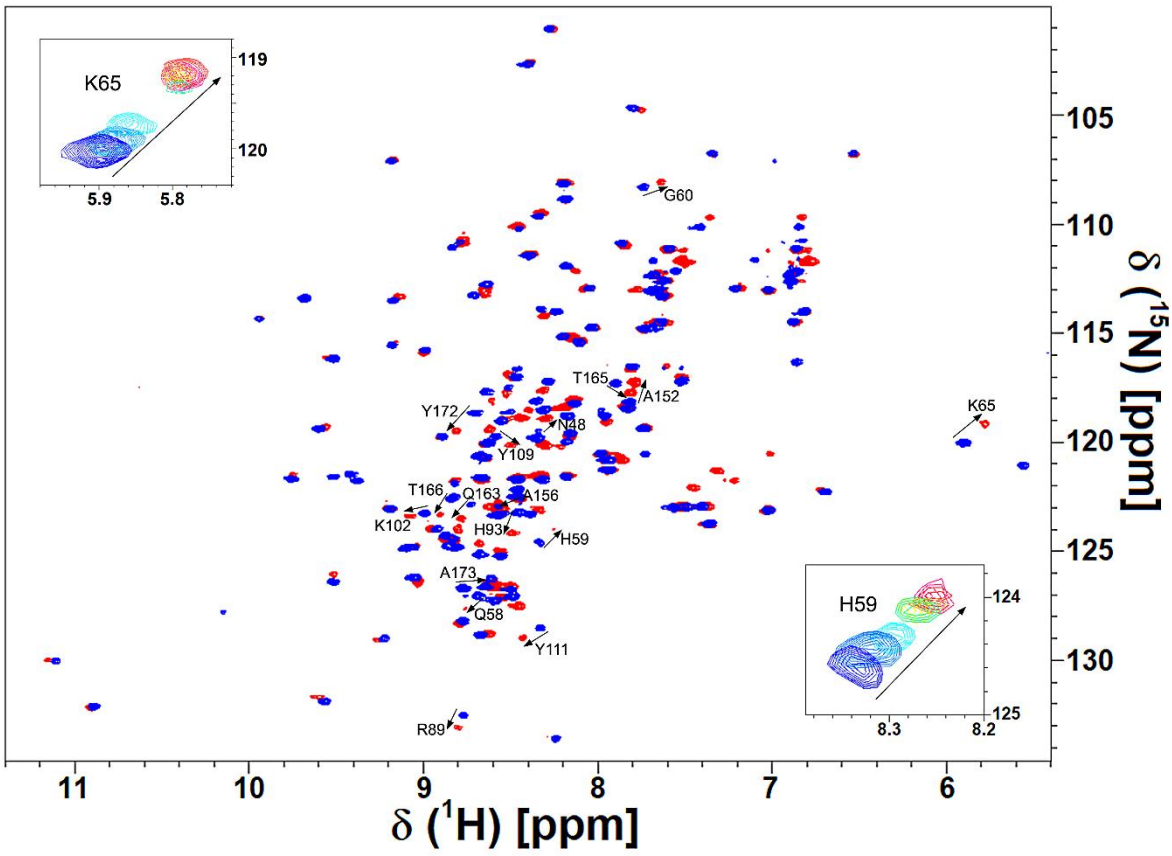

Figure S5. Two-dimensional [^{1}H-^{15}N] HSQC spectra of the free (blue) and ssTRS(–)-bound (red) [U-^{15}N]N-NTD collected by using an NMR spectrometer operating at ^{1}H frequency of

600.033 MHz at 20°C. The arrows indicate the residues that presented a chemical shift perturbation (CSP) upon DNA binding higher than $\Delta\delta_{ave}$ + SD. (Insets) Titration effect on the behavior of the exchange regime for H59 and K65 upon ssTRS(–) binding. The spectra were recorded at a protein concentration of 70 μM (dark blue) and at an DNA concentration of 9.8 μM (light blue), 23.8 μM (cyan), 53.9 μM (green), 70 μM (yellow), 109.9 μM (orange), 174.4 μM (pink), and 209.3 μM (red).

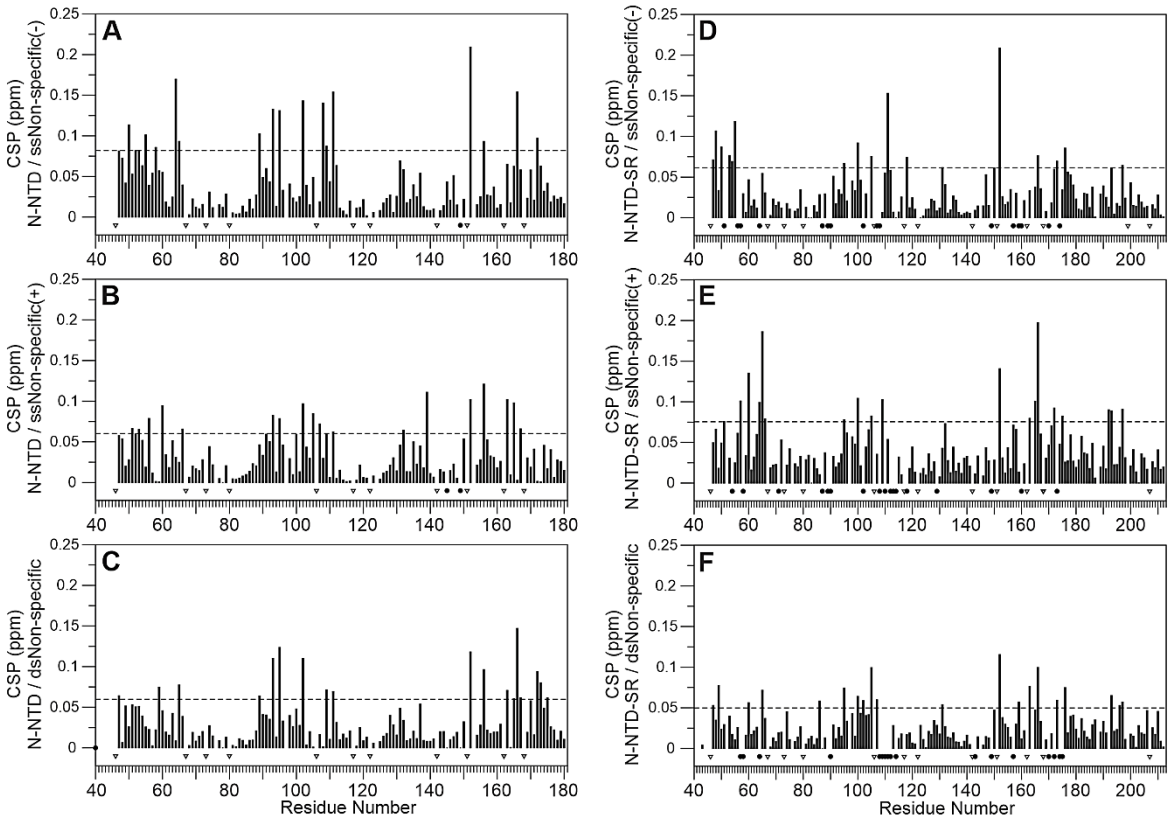

Figure S6. Chemical Shift Perturbation (CSP) of N-NTD and N-NTD-SR protein upon interaction with non-specific (NS) DNA. Chemical Shift Perturbation (CSP) of N-NTD (A, B, and C) and N-NTD-SR (D, E, and F) upon interaction with ssNS(–) (A, D), ssNS(+) (B, E), and dsNS (C, F). A chemical shift perturbation higher than the average CSP value ($\Delta\delta_{ave}$) plus one standard deviation (SD) (dotted line) was considered statistically significant and used as experimental constraints in subsequent docking calculations. The proline residues (46, 67, 73, 80, 106, 117, 122, 142, 151, 162, 168, 199 and, 207) are indicated by triangle-down. Resonances signals broadened beyond detection upon NS titration are represented by the filled black circles.

Figure S7. Alignment of the consensus residues with CSP values higher than Δδave + SD for the interaction of N-NTD and N-NTD-SR with ssDNAs and dsDNAs identified in the NMR titration experiments. The asterisk (*) and colon (:) indicate fully (three) and partial (two) consensus residues, respectively. Residue with resonance signal broadened beyond detection upon DNA binding is colored in red. The symbol § denotes the residues of N-NTD reported by Dinesh and cols. (2020) for the interaction with 7-nucleotide and 10-nucleotide ssTRS(+) RNA (5'–CUAAACG–3', 5'–UCUCUAAACG–3'), and 7-nucleotide non-specific RNA duplex (5'–CACUGAC–3' and 5'–GUCAGUG–3') [1]. Numbers represent the amino acid position in the protein sequence.

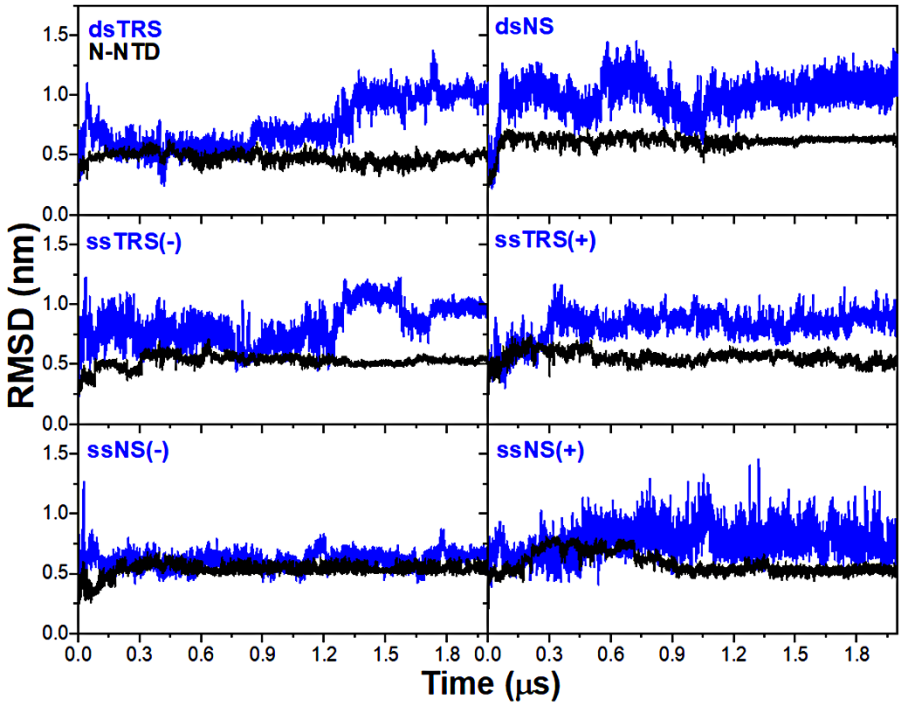

Figure S8. RMSD values of the backbone atoms of N-NTD (black) and protein-bound DNA (blue) along the 2 μs MD simulations.

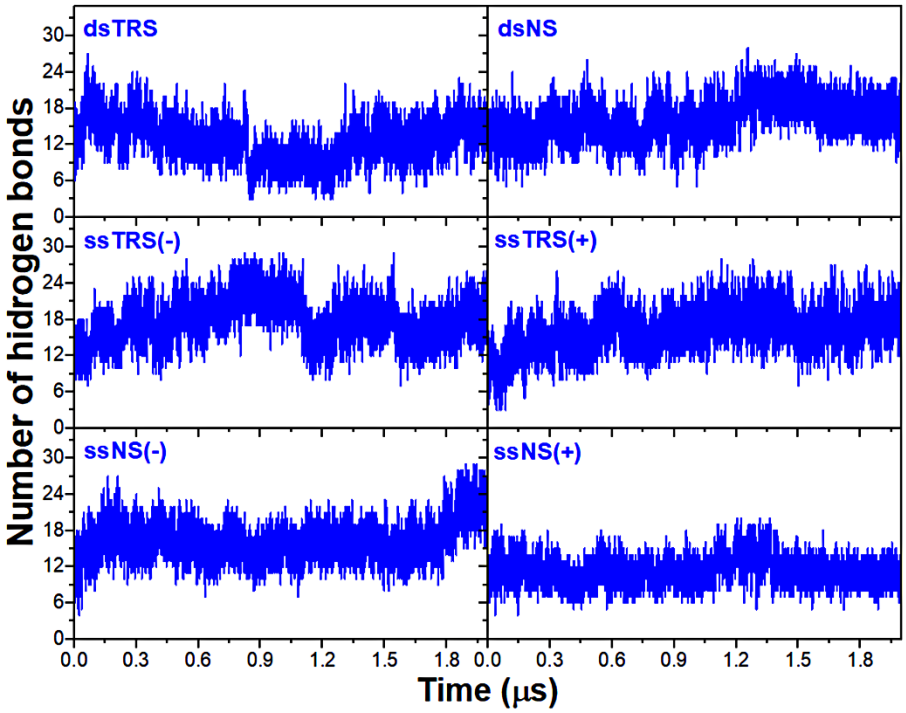

Figure S9. Number of protein-DNA hydrogen bonds formed between N-NTD and DNAs along the 2 μs MD simulations.

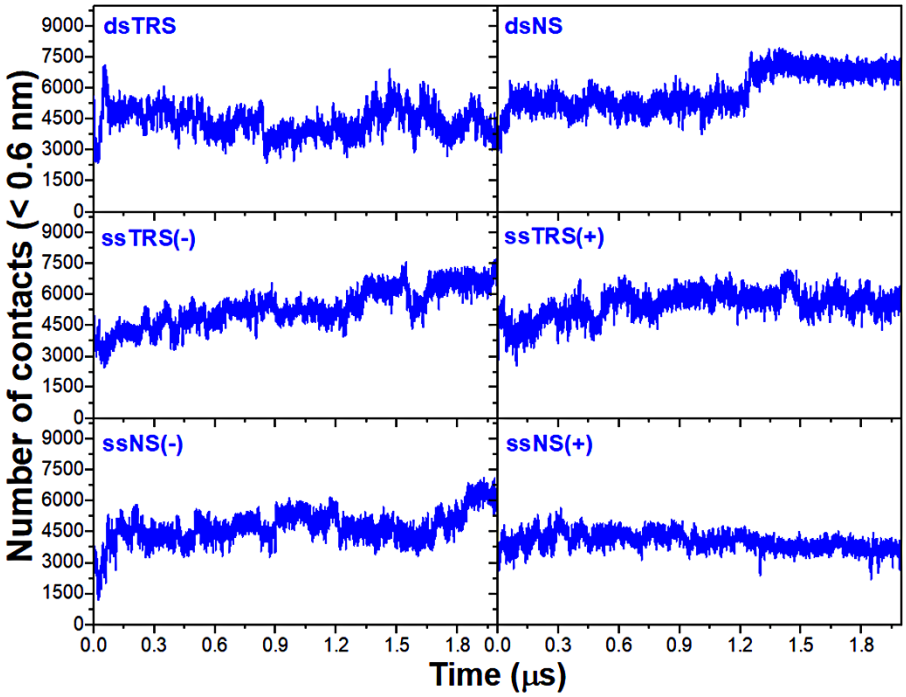

Figure S10. Number of contacts < 0.6 nm between the N-NTD and DNA atoms for the along the 2 μs MD simulations.

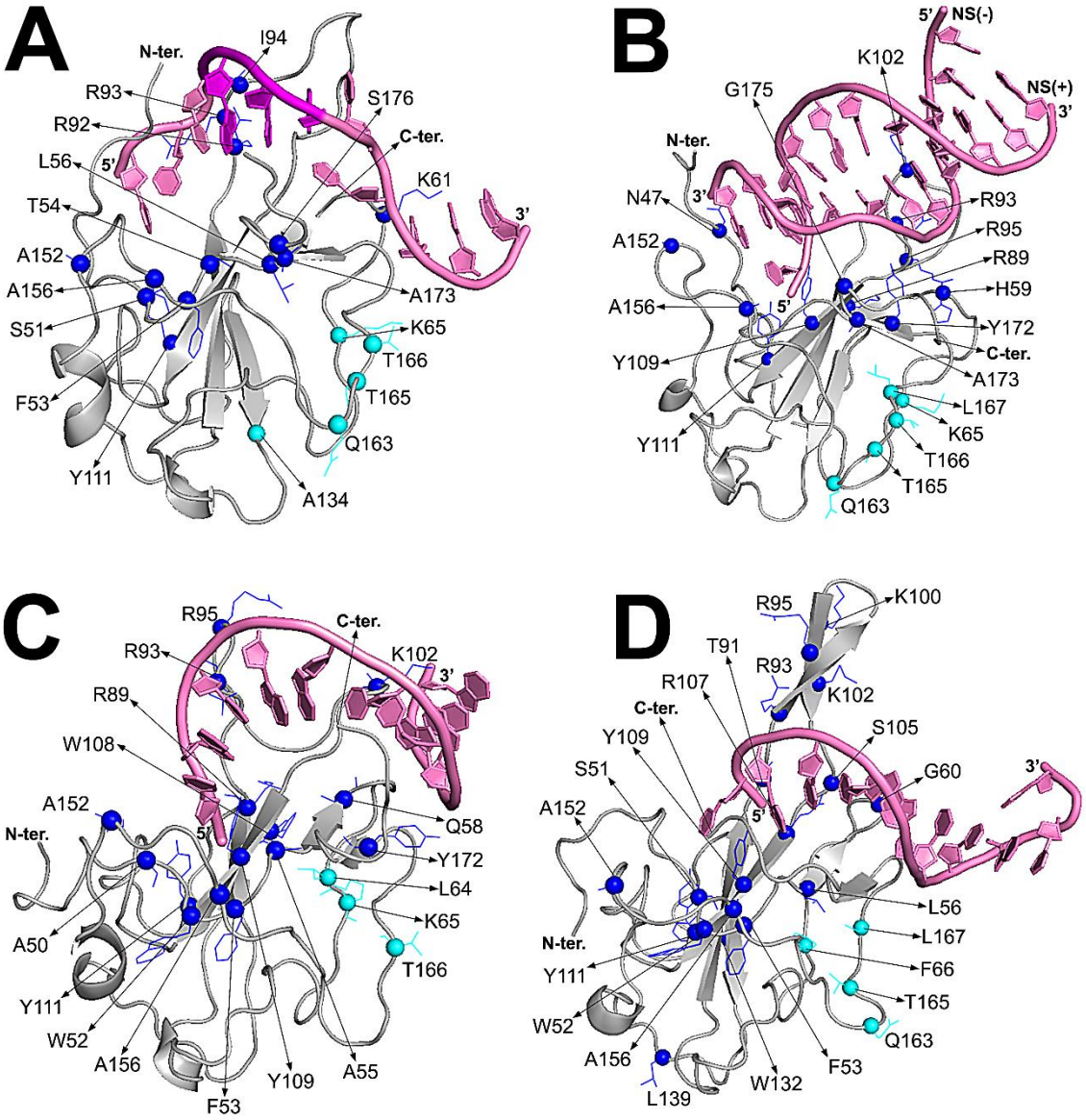

Figure S11. Representative structure models of the N-NTD complexed with (A) ssTRS(+), (B) dsNS, (C) ssNS(–), and (D) ssNS(+) obtained from the molecular docking and molecular dynamic simulations. The spheres denote the residues with CSP values higher than $\Delta\delta_{ave}$ + SD, participating directly in the binding interface (blue) and in the remote region (cyan). The proteins are shown as a cartoon model, DNAs are represented as a cartoon-ring model, and the TTT motif in TRS is colored in magenta.

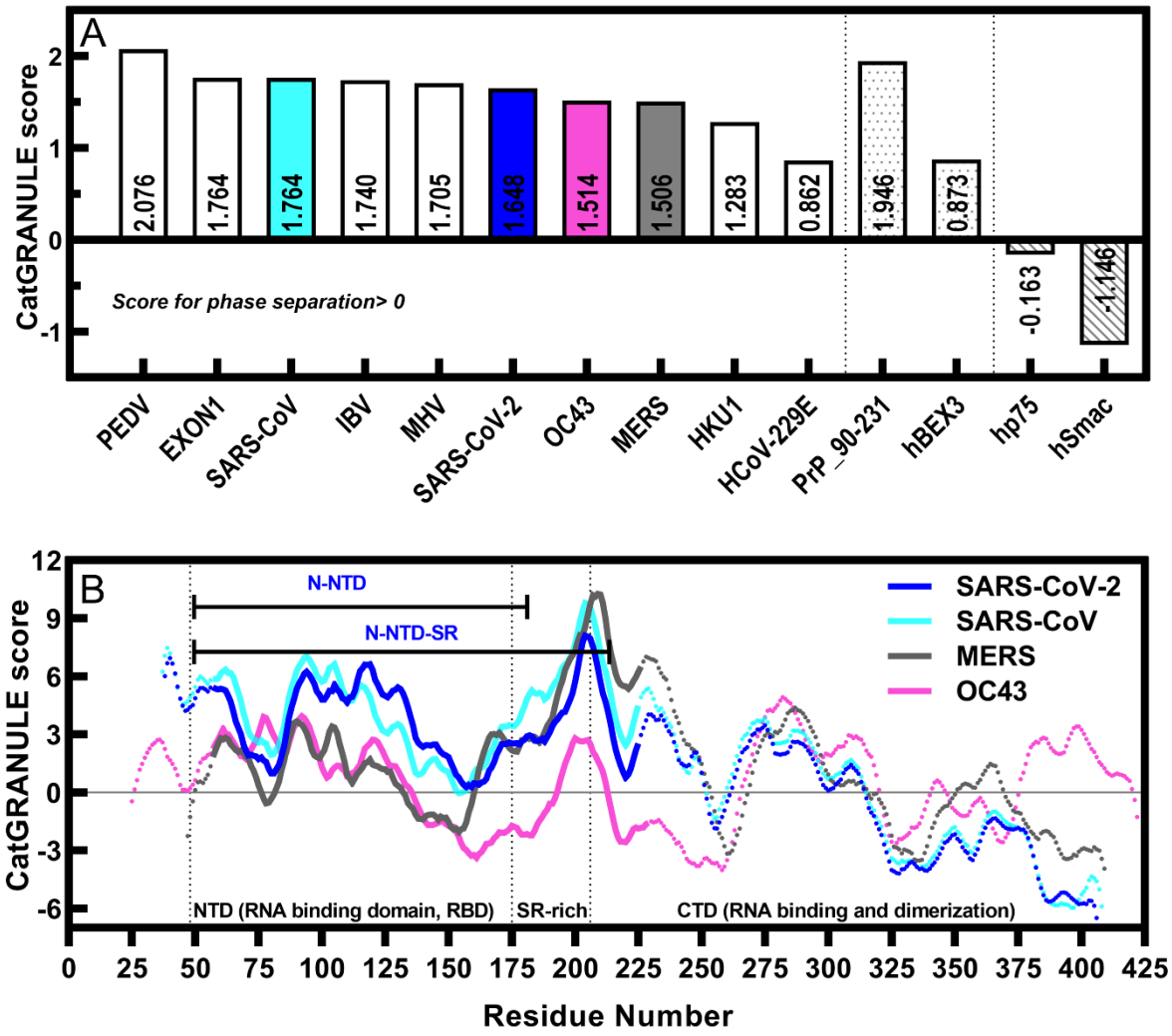

Figure S12. The liquid-liquid phase separation of N-protein is predicted to be conversed between coronaviruses and the SR-rich region shows the highest propensity. (A) Scores for liquid-liquid phase separation, based on catGRANULE algorithm [3] (score > 0 indicates granule formation; above 1.0 denotes strong propensity; available at http://service.tartaglialab.com/new_submission/catGRANULE). Scores of N-protein from SARS-Cov-2 (Uniprot ID P0DTC9); SARS-CoV (P59595); MERS (R9UM87); OC43 (P33469); HCoV-229E (P15130); PEDV (Q07499); IBV (P69596); HKU1 (Q5MQC6); EXON1 (D2E1E2) were compared with controls. Positive controls composed by proteins that were experimentally proved to undergo *in vitro* liquid-liquid phase separation modulated by nucleic acids, such as the prion protein (PrP_90-231; Uniprot ID P04925) [4] and Brain Expressed-X-linked 3 (BEX3; Uniprot ID Q00994) [5] represented by dotted bars. Negative controls composed by folded proteins such as hp75 (Uniprot ID P02768) and hSmac/Diablo (Uniprot ID Q9NR28) represented by stripped bars. Betacoronaviruses (colored bars) had phase separation propensity calculated along primary structure in **(B)** by catGRANULE. The dotted lines mark the domain organization of N-protein (1-48 is the IDR; 48-175 is the RBD; 176-206 is the SR-rich region followed by the C-terminus). The N-terminal constructs studied herein are highlighted in blue font. The SR-rich region, which is only present in the

N-NTD-SR construct, is likely to majorly contribute to N-protein liquid-liquid phase separation.

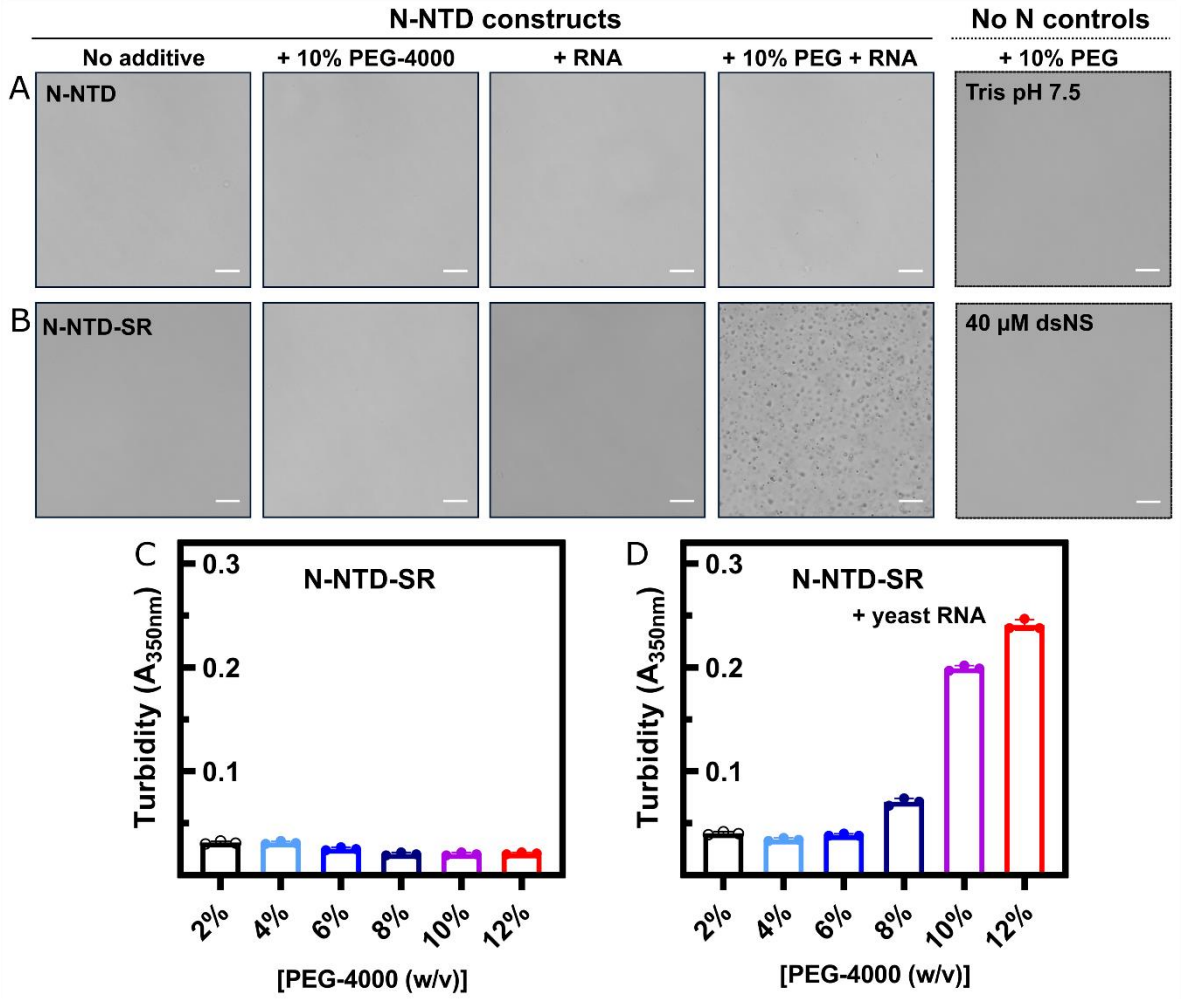

Figure S13. The N-protein SR-rich region is important for N-NTD demixing with RNA. Representative phase contrast micrographs of 20 µM N-NTD (A) and 20 µM N-NTD-SR (B) in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl. Left to right: only protein in buffer; protein in buffer containing the crowding agent 10% (w/v) PEG-4000; protein in buffer + Torula yeast RNA (1:2); protein in buffer + 10% (w/v) PEG-4000 + Torula yeast RNA (1:2). Controls are shown in dashed frames. Only vehicle consisting of buffer + 10% (w/v) PEG-4000. Only nucleic acid in the highest concentration used in the experiments (40 µM ds22-bp). Phase separation of 20 µM N-NTD-SR in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl monitored by absorbance measurements at 350 nm as a function of increasing concentrations of PEG-4000 (2% (black), 4% (light blue), 6% (marine blue), 8% (dark blue), 10% (purple), 12% (red)) in the absence (C) and presence (D) of Torula yeast RNA extract.

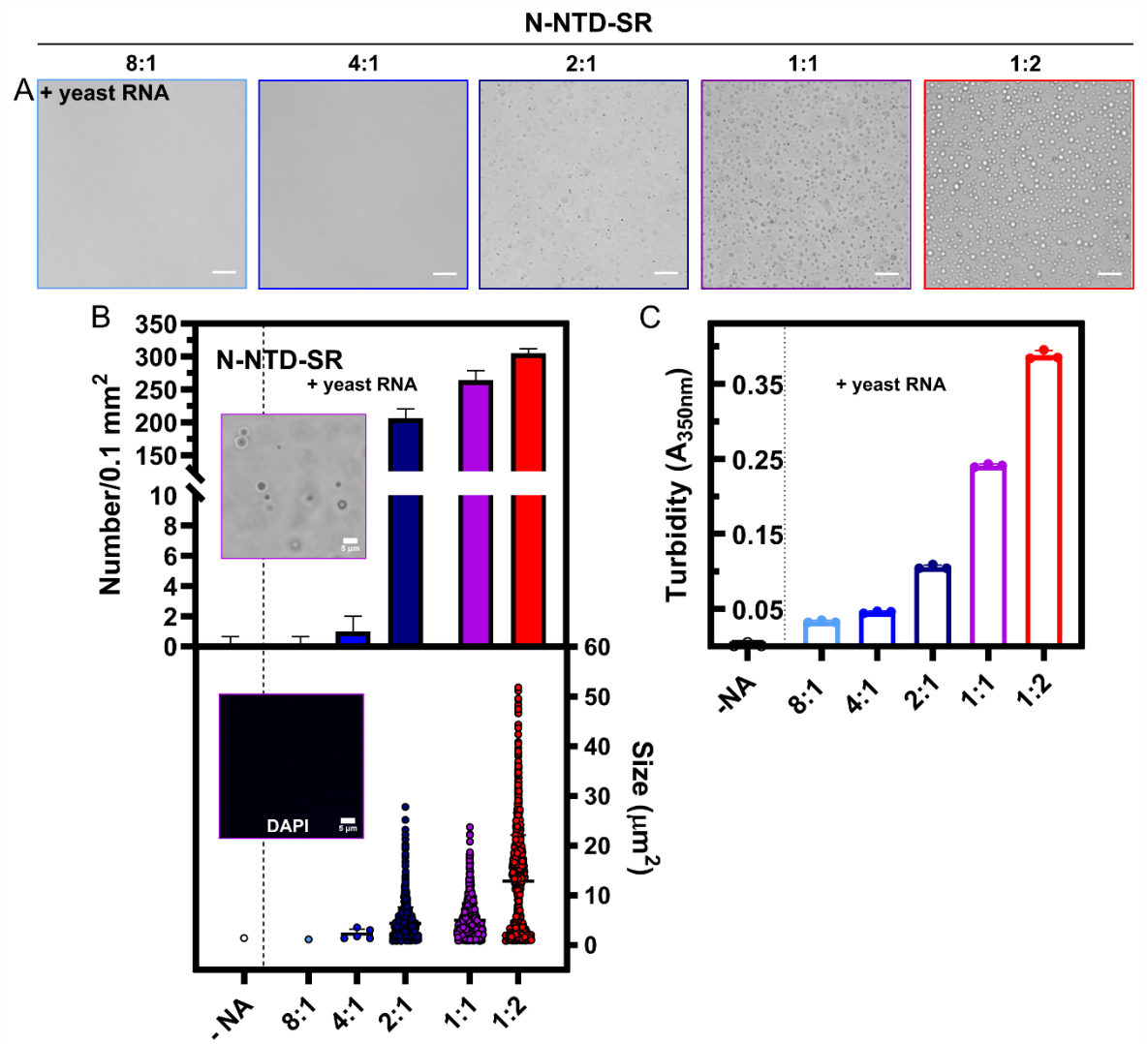

Figure S14. Phase separation is promoted by increasing concentrations of yeast RNA. LLPS assessment of 20 µM N-NTD-SR in the presence of increasing Torula yeast RNA concentrations: 8:1 (light blue); 4:1 (marine blue); 2:1 (dark blue); 1:1 (purple); 1:2 (red) protein:RNA molar ratios, respectively, in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl. (A) Representative phase contrast micrographs. (B) Top graph: Mean number of condensates ± S.D. per 100 µm^2 area (n= 5 images). Top inset: Phase contrast microscopy in the presence of DAPI for 1:1 protein:RNA stoichiometry. Bottom graph: Scatter plot from the size of condensates (represented as area in µm^2) obtained from micrographs analysis. The following number of condensates had their size measured: 8:1 (N= 1); 4:1 (N= 5); 2:1 (N= 1028); 1:1 (N= 1321) and 1:2 (N= 1753). Bottom inset: corresponding DAPI emission (0.25 µg/mL) from the top graph inset image (no fluorescence observed). All conditions contained 10% (w/v) PEG-4000. Scale bar, 20 µm apart from insets (5 µm). (C) Phase separation of 20 µM N-NTD-SR in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl) in the presence of 10% (w/v) PEG-4000 monitored by absorbance measurements at 350 nm as a function of increasing concentrations of Torula yeast RNA extract. The protein:RNA stoichiometries are the same as in (A).

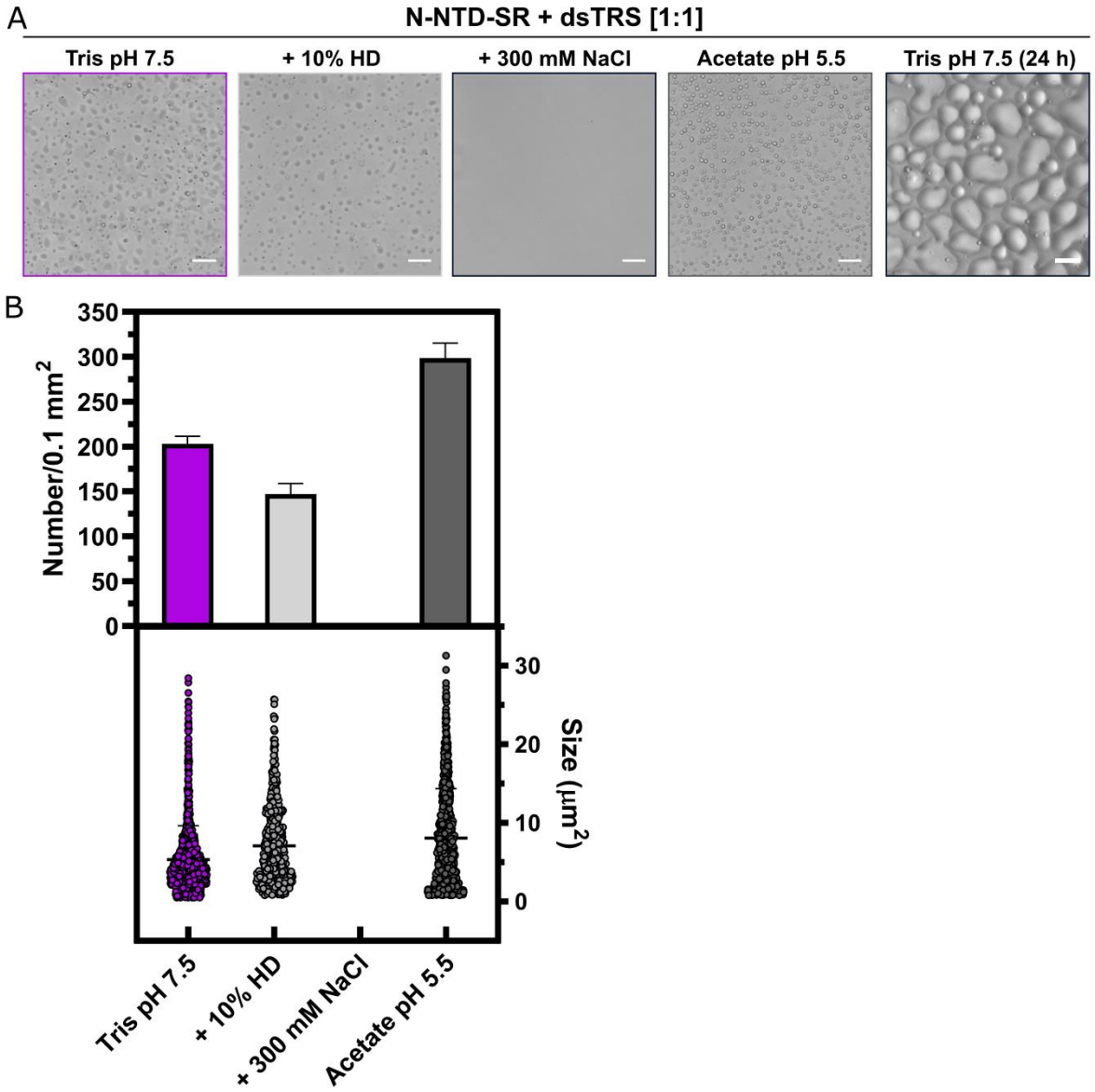

Figure S15. N-NTD-SR:dsTRS liquid-liquid phase separation is driven by electrostatic contacts and promoted by acidic pH. (A) Representative phase contrast micrographs from samples containing 20 µM N-NTD-SR in the presence of 20 µM dsTRS (1:1 molar ratio). Top left: sample in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl). Top right: with addition of 10% (w/v) 1,6-hexanediol (HD). Bottom left: with addition of 300 mM NaCl. Bottom right: sample in 20 mM sodium acetate buffer pH 5.5 with 30 mM NaCl. (B) Quantification of N-NTD-SR:dsTRS condensates in the presence of additives. Top graph: Mean number of condensates ± S.D. per 100 μm^2 area (n= 5 images). Bottom graph: Scatter plot from condensates' size (represented as area in μm^2) obtained from micrographs analysis. The following number of condensates' size were analyzed: sample in buffer (purple data, N= 1013); sample in buffer containing 10% HD (light gray data, N=598); no condensates visualized in presence of 300 mM NaCl (absence of column, N=0); sample in 20 mM sodium acetate pH 5.5, 30 mM NaCl (dark gray column, N=1174). All conditions contained 10% (w/v) PEG-4000. Scale bar, 20 µm.

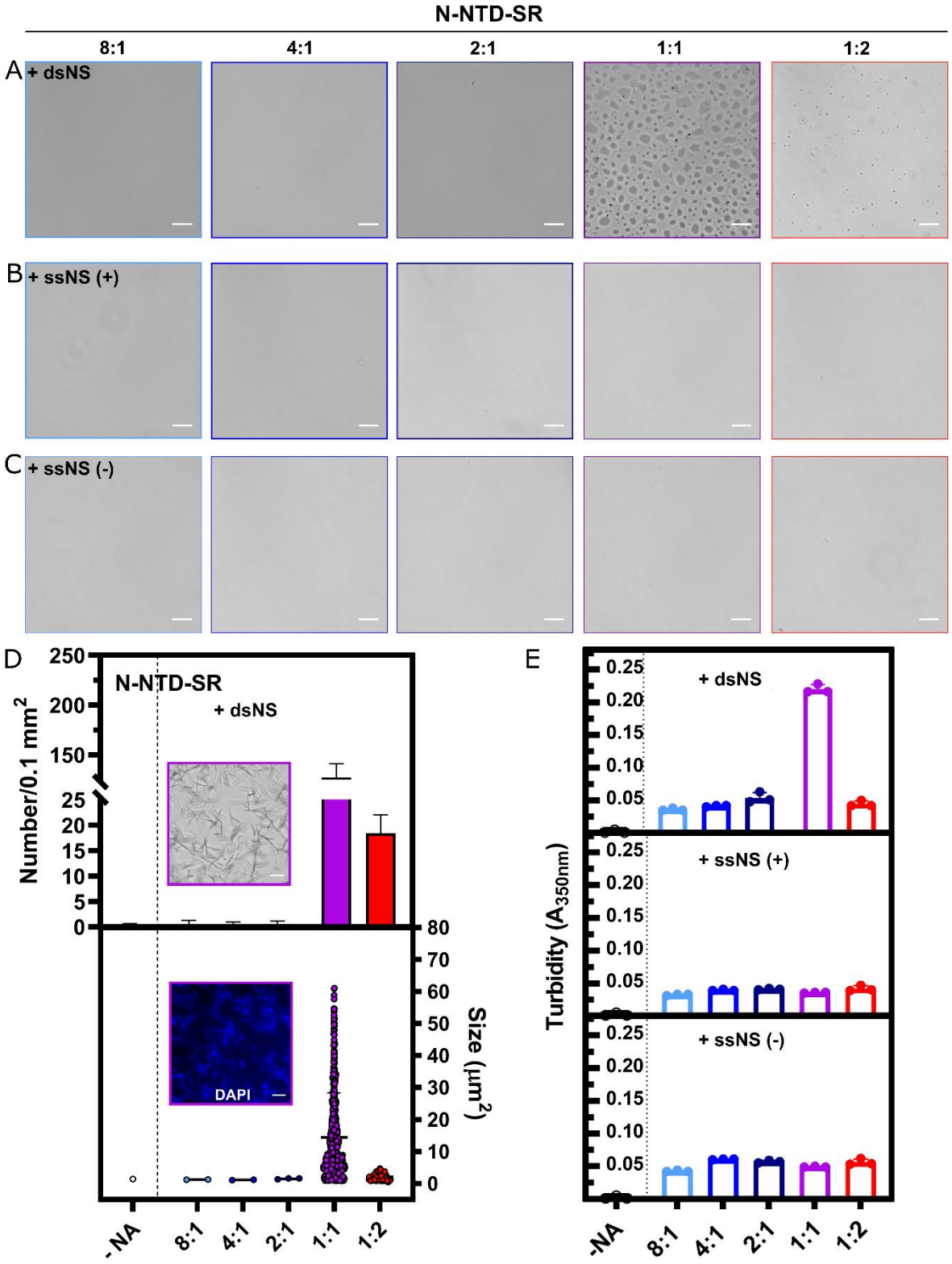

Figure S16. The non-specific DNA oligonucleotide needs secondary structure to induce N-NTD-SR liquid-liquid phase separation. Representative phase contrast micrographs of 20 μM N-NTD-SR in the presence of dsNS (A); ssNS(+) (B) or ssNS(–) (C) at the following protein:DNA stoichiometries: 8:1 (light blue); 4:1 (marine blue); 2:1 (dark blue); 1:1

(purple); 1:2 (red), respectively, in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl. (D) Top graph: Mean number of condensates ± S.D. per 100 μm^2 area (n= 5 images). Top inset: Phase contrast microscopy in the presence of DAPI for 1:1 N-NTD-SR:dsNS). Bottom graph: Scatter plot from the size of condensates (represented as area in μm^2) obtained from micrographs analysis. Since no phase separation was observed for the single-stranded oligonucleotides, the following number of condensates' size were measured for the dsNS: -NA (1); at 8:1 (N= 2); 4:1 (N= 2); 2:1 (N= 3); 1:1 (N= 632) and 1:2 (N= 81). Bottom inset: corresponding DAPI emission from the top graph inset image acquired after 2 h incubation evidencing needle-like crystals formed by N-NTD-SR:dsNS at 1:1 stoichiometry. All conditions contained 10% (w/v) PEG-4000. Scale bar, 20 µm including insets. (E) Phase separation of 20 µM N-NTD-SR in 20 mM Tris-HCl buffer (pH 7.5) containing 30 mM NaCl in the presence of 10% (w/v) PEG-4000 monitored by absorbance measurements at 350 nm as a function of increasing concentrations of the non-specific DNA oligonucleotides dsNS (top graph), ssNS(+) (middle graph), and ssNS(–) (bottom graph). The protein:DNA stoichiometries are the same as in (A).

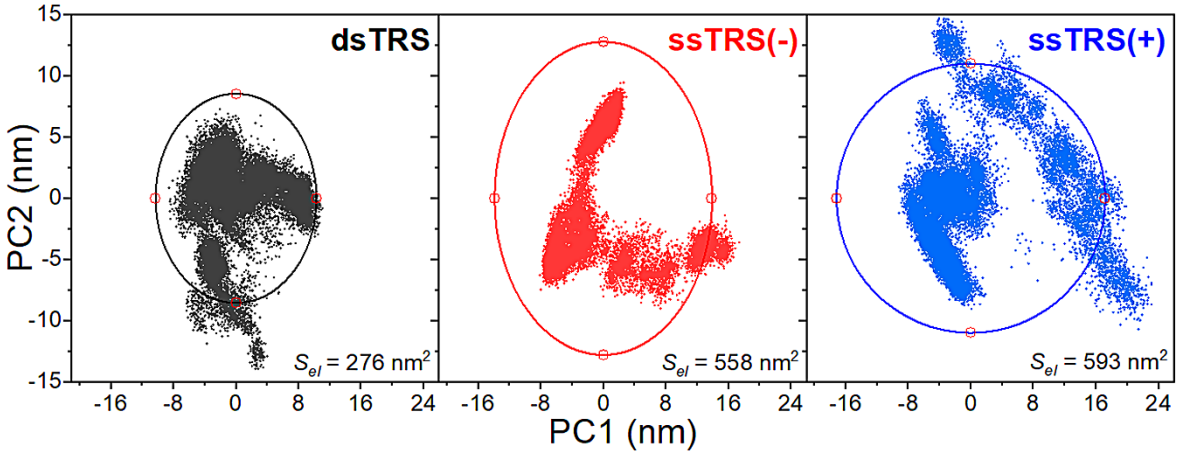

Figure S17. PCA scatter plots PC1 and PC2 for N-NTD complexed with dsTRS (black dots, *left*), ssTRS(–) (red dots, *middle*), and ssTRS(+) (blue dots, *right*). The extent of the conformational space for each scatter plot was measured by fitting an elliptical shell (solid lines) that contains 95% (confidence) of the density. The area of the ellipse (S_{el}) is a measure of the conformational space for the free and bound protein. The N-NTD complexed with dsTRS explored a smaller conformational space along the 2 µs MD simulation, meaning a more restricted motion of the protein, which may be a result of a better-defined position and orientation of the DNA duplex.

Table S1. Persistent hydrogen bonds (and salt bridges) between N-NTD and dsTRS predicted from the MD simulation.

dsTRS				
Donor	**Donor Atom**	**Acceptor**	**Acceptor atom**	**Persistence percentage**
T1(+)	O5'	ASN48	OD1	14.579
ASN48	ND2	T1(+)	O4'	14.199
ARG88	NH1	C2(+)	O2P	56.627
ASN48	ND2	C2(+)	O1P	42.968
ARG88	NH2	C2(+)	O2P	20.094
ARG88	NH2	C2(+)	O1P	18.484
C2(+)	N4	GLU174	OE2	15.709
C2(+)	N4	GLU174	OE1	12.979
ASN47	N	C2(+)	O1P	10.274
ARG107	NH1	T3(+)	O4	49.028
ARG92	NE	T3(+)	O2P	12.009
ARG107	NH2	T3(+)	O4	10.324
ARG92	NH2	A4(+)	O2P	10.174
ARG95	NH1	A5(+)	O2P	11.854
GLY97	N	G4(-)	O2P	31.763
LYS102	NZ	G4(-)	O2P	26.994
LYS102	NZ	G4(-)	O1P	18.974
LYS102	NZ	G4(-)	N7	12.384
GLY97	N	G4(-)	O1P	11.399
LYS61	NZ	G5(-)	O2P	40.308
LYS102	NZ	G5(-)	O6	11.439
GLY60	N	T6(-)	O2P	32.713
GLY60	N	T6(-)	O1P	30.848
SER105	OG	T6(-)	O2P	22.159
HIS59	NE2	T7(-)	O1P	38.308
ARG177	NH2	T7(-)	O2P	22.799
TYR172	OH	T7(-)	O2P	21.344
ARG177	NH1	T7(-)	O1P	11.399
ARG177	NE	T7(-)	O1P	10.509
ARG107	NH2	T8(-)	O4	36.593
ARG177	NH2	T8(-)	O2P	16.424
ARG177	NH1	T8(-)	O2P	13.154
GLY178	N	T8(-)	O2P	11.609
ARG177	NE	T8(-)	O1P	11.309
A9(-)	N6	GLU174	OE2	14.734
G9(-)	N2	ALA42	O	16.694

Table S2. Persistent hydrogen bonds (and salt bridges) between N-NTD and dsNS predicted from the MD simulation.

dsNS				
Donor	**Donor Atom**	**Acceptor**	**Acceptor atom**	**Persistence percentage**
TYR111	OH	C1(+)	O5'	87.506
ASN48	ND2	C1(+)	O2	74.066

ASN48	ND2	A2(+)	O4'	79.506
A2(+)	N6	TYR109	OH	46.253
ARG88	NH2	A2(+)	O1P	44.448
ARG88	NH1	A2(+)	O1P	40.648
ARG88	NH1	A2(+)	O2P	34.933
LEU45	N	A2(+)	N3	27.964
ASN47	ND2	A2(+)	N3	17.499
ARG92	NH2	C3(+)	O2P	19.569
ARG92	NH1	C3(+)	O2P	16.679
ARG92	NE	C3(+)	O2P	10.604
ALA42	N	T4(+)	O2	19.659
LYS102	N	G4(-)	O2P	55.752
LYS102	NZ	G4(-)	N7	17.869
ARG93	NH2	G5(-)	O2P	80.041
ARG93	NH1	G5(-)	O2P	67.972
LYS102	NZ	G5(-)	O6	48.863
MET101	N	G5(-)	O2P	44.918
LYS100	NZ	G5(-)	O1P	22.094
ARG93	NH2	G5(-)	O5'	21.444
GLY99	N	G5(-)	O1P	17.599
SER105	OG	T6(-)	O2P	47.303
SER105	N	T6(-)	O2P	42.653
ARG177	NH2	T6(-)	O1P	27.444
ARG93	NH1	G5(-)	O1P	19.864
ARG177	NH1	T6(-)	O1P	17.459
LYS102	NZ	T6(-)	O4	11.899
ARG177	NH1	C7(-)	O2P	34.583
SER105	OG	C7(-)	N4	22.519
ARG177	NH2	C7(-)	O2P	15.014
TYR172	OH	C7(-)	O2P	10.309
ARG107	NH1	G9(-)	O6	55.462
G9(-)	N2	ALA42	O	34.598
ARG107	NH2	G9(-)	O6	17.359
GLY44	N	T10(-)	O2	32.723
G11(-)	N2	LEU45	O	36.778
G11(-)	O3'	GLY44	O	24.269
G11(-)	N2	ASN47	OD1	21.884

Table S3. Persistent hydrogen bonds (and salt bridges) between N-NTD and ssTRS(–) predicted from the MD simulation.

ssTRS(-)				
Donor	**Donor Atom**	**Acceptor**	**Acceptor atom**	**Persistence percentage**
ASN47	ND2	C1	O2	39.168
C1	N4	SER180	OC2	37.868
C1	N4	SER180	OC1	32.798
C1	O5'	THR49	O	20.004
THR49	OG1	C1	O5'	18.069

TYR111	OH	C1	O5'	14.929
TYR109	OH	G2	O2P	81.581
ALA42	N	G2	O1P	34.533
G2	N1	SER180	OC2	24.599
G2	N1	SER180	OC1	22.829
SER180	OG	G2	N1	20.089
G2	N2	ARG177	O	19.794
LEU45	N	G2	O1P	18.994
G2	N2	SER180	OC1	18.649
G2	N2	SER180	OC2	15.734
G2	N2	GLY178	O	12.084
ARG149	NH2	G2	O6	11.744
ARG92	NH2	G2	O2P	11.049
ARG107	NH2	C3	O2P	48.448
ARG177	N	C3	O4'	31.698
ARG107	NH1	C3	O2P	30.493
ARG93	NH2	C3	O1P	25.624
ARG177	NE	C3	O2	17.514
ARG177	NH2	C3	O2P	14.899
ARG93	NE	C3	O1P	14.064
ARG177	NH1	C3	O5'	13.159
ARG95	NH1	G4	O2P	35.548
ARG95	NH2	G4	O2P	33.258
ARG93	NH2	G4	O1P	13.449
ARG107	NH2	G4	N7	11.914
ARG93	NH2	G4	O2P	11.859
ARG93	NE	G4	O1P	11.014
ARG92	NH2	G5	O6	31.083
ARG95	NE	G5	O2P	30.638
G5	N1	GLU174	OE1	23.599
G5	N1	GLU174	OE2	23.094
ARG92	NH1	G5	O6	23.054
G5	N2	GLU174	OE1	20.434
ARG95	NH2	G5	O5'	20.164
G5	N2	GLU174	OE2	19.564
ARG95	NH2	G5	O2P	14.729
SER105	N	T6	O4	29.894
SER105	OG	T6	N3	17.864
TYR172	OH	T6	O2	10.819
TYR172	OH	T7	O2	34.293
GLY170	N	A9	O1P	49.623
LYS169	NZ	G10	O1P	47.848
LYS61	N	A11	N1	14.344

Table S4. Persistent hydrogen bonds (and salt bridges) between N-NTD and ssTRS(+) predicted from the MD simulation.

ssTRS(+)				
Donor	**Donor Atom**	**Acceptor**	**Acceptor atom**	**Persistence percentage**
GLY44	N	T1	O3'	24.984
T1	N3	ASN153	OD1	11.039
ARG107	NH1	C2	O2	55.107
GLY44	N	C2	O1P	49.063
ARG107	NH2	C2	O2	46.193
ARG93	N	C2	O3'	30.173
THR91	N	C2	N3	17.054
ARG93	NH1	C2	O2P	15.149
ARG93	NH2	C2	O2P	11.189
ARG93	NH2	C2	O1P	10.979
ARG92	NE	T3	O2P	85.696
ILE94	N	T3	O2P	80.171
ARG93	N	T3	O2P	59.827
ARG92	NH2	T3	O2P	53.822
ARG93	NE	T3	O1P	39.638
ARG93	NH2	T3	O1P	29.774
ARG92	NH2	T3	O5'	25.129
MET43	N	T3	O2	13.219
GLY41	N	T3	O3'	11.714
ARG92	NH1	A4	N7	20.259
GLY41	N	A4	O1P	16.654
A4	N6	GLU174	OE2	15.519
A4	N6	GLU174	OE1	13.904
ARG92	NH2	A4	O2P	10.359
ARG177	NH2	A6	O3'	18.094
ARG177	NH1	C7	O1P	26.609
ARG177	NH2	C7	O1P	20.904
TYR172	OH	C7	O2	10.769
HIS59	NE2	C8	O2P	21.889
LYS61	N	G9	O6	42.953
GLY170	N	G9	O4'	33.788
HIS59	NE2	G9	O5'	29.959
G9	N1	GLU62	OE2	13.409
G9	N2	GLU62	OE2	12.664
G9	N2	GLU62	OE1	12.489
G9	N1	GLU62	OE1	12.439
G9	N2	GLY60	O	10.804
LYS169	NZ	C10	O1P	36.403
GLY170	N	C10	O1P	23.959
LYS169	NZ	G11	O1P	16.139

Table S5. Persistent hydrogen bonds (and salt bridges) between N-NTD and ssNS(–) predicted from the MD simulation.

ssNS(-)				
Donor	**Donor Atom**	**Acceptor**	**Acceptor atom**	**Persistence percentage**
C1	O5'	ASN154	O	24.759
ASN154	ND2	C1	O2	12.644
C1	O5'	PHE53	O	11.774
TYR109	OH	G2	O2P	45.728
THR91	OG1	G2	O1P	18.169
THR91	N	G2	O1P	14.094
ARG88	NH1	G2	O1P	12.874
ARG93	NH2	C3	O2P	30.998
ARG95	NH2	C3	O1P	16.854
ARG93	NH1	C3	O2P	13.529
ARG95	NH1	C3	O1P	12.884
ARG93	NE	C3	O2P	11.774
ARG107	NH1	G4	O6	74.326
ARG107	NH2	G4	O6	68.737
G4	N2	SER180	OC1	42.148
ARG93	NH2	G4	O2P	25.424
G4	N1	SER180	OC1	20.654
ARG93	NH1	G4	O2P	19.449
ARG93	NH1	G4	N7	15.744
G4	N2	SER180	OC2	13.974
ARG93	NE	G4	O2P	13.334
G4	N1	SER180	OC2	12.244
G5	N2	SER180	OC2	64.952
G5	N2	SER180	OC1	30.078
SER180	N	T6	O2	80.641
SER180	OG	T6	O4'	80.456
ARG177	NH1	T6	O2	11.934
HIS59	NE2	C7	O2	83.621
ARG177	NH1	C7	O2	68.262
SER105	OG	C7	N4	43.693
C7	N4	ASP103	O	17.694
A8	N6	ASP103	O	16.174
TYR172	OH	G9	O1P	86.146
G9	N2	ASP103	OD2	36.903
G9	N2	ASP103	OD1	32.158
LYS61	N	T10	O1P	50.137
LYS102	NZ	T10	O2	17.034
LYS61	NZ	G11	O1P	37.243
LYS61	NZ	G11	O2P	30.938
ASP103	N	G11	O3'	12.754

Table S6. Persistent hydrogen bonds (and salt bridges) between N-NTD and ssNS(+) predicted from the MD simulation.

ssNS(+)				
Donor	**Donor Atom**	**Acceptor**	**Acceptor atom**	**Persistence percentage**
TYR111	OH	C1	N4	55.397
C1	N4	ASN47	O	48.013
TYR109	OH	C1	O5'	38.443
C1	N4	THR49	O	25.709
SER176	OG	A2	N1	93.46
A2	N6	SER176	O	92.245
ARG92	NE	A2	O2P	24.794
ARG92	NE	A2	O1P	20.869
ARG92	NH1	A2	O2P	19.324
ARG92	NH2	A2	O2P	17.249
ARG92	NH2	A2	O1P	15.019
SER176	OG	A2	N6	13.449
ARG92	NH1	A2	O1P	11.304
C3	N4	GLU174	OE2	48.983
C3	N4	GLU174	OE1	45.353
ARG92	NH1	C3	O2P	17.549
ARG92	NH2	C3	O2P	15.644
ARG107	NH2	T4	O4	43.378
TYR172	OH	G5	N2	68.177
TYR172	OH	A6	N3	50.672
GLY170	N	C8	O1P	77.686
LYS169	NZ	G9	O1P	47.033

Table S7. Binding constants and thermodynamic parameters of the interaction of N-NTD and N-NTD-SR with different DNAs in 20 mM Bis-Tris pH 6.5. Data obtained from the fluorescence quenching experiments. The statistical significance between the observed means was tested using Pearson t student test with a significance of $P < 0.05$ and 95 % confidence interval of the reported difference. The values compared are: ssTRS(+) & ssTRS(–) *, ssNS(+) & ssNS(–) **, ssTRS(+) & ssNS(+) ***, ssTRS(–) & ssNS(–) ****, and dsTRS & dsNS *****.

Protein	DNA	T (°C)	Kd (nM)	P < 0.05	ΔH (kJ/mol)	ΔG (kJ/mol)	T·ΔS (kJ/mol)
N-NTD	ssTRS(–)	15	9 ± 4	*		–44 ± 1	100 ± 2
		25	4 ± 2	*	56 ± 1	–47 ± 1	103 ± 2
		35	2 ± 1			–51 ± 1	107 ± 2
	ssTRS(+)	15	45 ± 12	*,***		–41 ± 1	103 ± 1
		25	19 ± 7	*	62 ± 1	–44 ± 1	106 ± 1
		35	8 ± 7	***		–48 ± 2	110 ± 2
	dsTRS	15	6 ± 2	*****		–45 ± 1	–15 ± 2
		20	9 ± 3	*****	–60 ± 2	–45 ± 1	–15 ± 2
		25	14 ± 4	*****		–45 ± 1	–15 ± 2
	ssNS(–)	15	16 ± 8			–43 ± 1	68 ± 1
		25	11 ± 10		25 ± 1	–45 ± 2	70 ± 2
		35	8 ± 5			–48 ± 2	73 ± 2
	ssNS(+)	15	17 ± 4	***		–43 ± 1	29 ± 2
		25	20 ± 4		–14 ± 2	–44 ± 1	30 ± 2
		35	25 ± 6	**,***		–45 ± 1	31 ± 2
	dsNS	15	30 ± 8	*****		–41 ± 1	1 ± 1
		20	40 ± 11	*****	–42 ± 1	–41 ± 1	1 ± 1
		25	54 ± 12	*****		–41 ± 1	1 ± 1
N-NTD-SR	ssTRS(–)	15	2 ± 1	*,****		–48 ± 1	109 ± 10
		25	1 ± 2	*,****	61 ± 10	–51 ± 5	112 ± 11
		35	0.3 ± 1	****		–56 ± 8	117 ± 13
	ssTRS(+)	15	31 ± 6	*,***		–41 ± 1	87 ± 1
		25	16 ± 7	*	46 ± 1	–44 ± 1	90 ± 1
		35	9 ± 7	***		–47 ± 2	93 ± 2
	dsTRS	15	5 ± 2	*****		–46 ± 1	–99 ± 1
		20	14 ± 3		–145± 1	–44 ± 1	–101 ± 1
		25	38 ± 6	*****		–42 ± 1	–103 ± 1
	ssNS(–)	15	36 ± 13	**,****		–41 ± 1	32 ± 1
		25	41 ± 6	****	–9 ± 1	–42 ± 1	33 ± 1
		35	46 ± 6	**,****		–43 ± 1	34 ± 1
	ssNS(+)	15	5 ± 4	**,***		–46 ± 2	–61 ± 2
		25	24 ± 10		–106 ± 2	–43 ± 1	–63 ± 2
		35	94 ± 19	**,***		–41 ± 1	–65 ± 2
	dsNS	15	89 ± 16	*****		–39 ± 1	171 ± 1
		20	35 ± 14		132 ± 1	–42 ± 1	174 ± 1
		25	14 ± 4	*****		–45 ± 1	177 ± 1

Table S8. Affinity of interaction between N-NTD and DNAs estimated by CSP in 20 mM sodium phosphate buffer (pH 6.5) containing 50 mM NaCl. The statistical significance between the observed means was tested using Pearson t student test with a significance of P < 0.05 and 95 % confidence interval of the reported difference. The values compared are: ssTRS(+) & ssTRS(–) *, ssNS(+) & ssNS(–) **, ssTRS(+) & ssNS(+) ***, and ssTRS(–) & ssNS(–) ****.

Residue/Region	**Kd (μM)**				**P < 0.05**
	ssTRS(+)	ssTRS(–)	ssNS(+)	ssNS(–)	
N-terminal					
N47	8±12	4±5	19±11	5±3	
N48	4±3	3±1	16±6	6±2	**,***
β-sheet II (β1/β5)					
T54	1±2	0.1±0.8	11±10	11±7	****
L56	4±3	5±5	22±9	1±4	**,***
Remote region					
D63	4±5	2±3	9±8	2±7	
K65	5±2	1.1±0.6	7±2	4±2	*
F66	7±11	1±2	6±5	6±7	
Q163	1±2	3±2	18±6	12±7	***
T166	4±1	-	11±2	6±2	**,***
L167	5±4	0.3±0.6	21±7	1±1	**,***
Finger					
F90	13±7	7±5	18±13	1±2	
R95	1±4	0.1±0.5	1±2	4±3	
G96	3±9	1±5	24±18	5±6	
K102	1±3	-	21±11	3±2	**,***
Palm (β2/β3/β4)					
Y111	2±2	0.4±0.5	1±1	2.3±0.9	****
Thumb					
A152	0.7±0.7	0.4±0.4	8±3	2.5±0.6	**,***,****
C-terminal					
S176	3±2	1±5	6±7	1±3	
G178	2±2	0.1±0.5	5±4	9±12	
G179	2±4	2±5	3±5	2±5	
All regions					
Average	4±3	2±2	12±7	4±3	

REFERENCES

[1] D.C. Dinesh, D. Chalupska, J. Silhan, E. Koutna, R. Nencka, V. Veverka, E. Boura, Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein., PLoS Pathog. 16 (2020) e1009100. https://doi.org/10.1371/journal.ppat.1009100.

[2] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. Mcgettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0, Bioinformatics. 23 (2007) 2947–2948.

https://doi.org/10.1093/bioinformatics/btm404.

[3] B. Bolognesi, N.L. Gotor, R. Dhar, D. Cirillo, M. Baldrighi, G.G. Tartaglia, B. Lehner, A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression, Cell Rep. 16 (2016) 222–231. https://doi.org/10.1016/j.celrep.2016.05.076.

[4] C.O. Matos, Y.M. Passos, M.J. do Amaral, B. Macedo, M.H. Tempone, O.C.L. Bezerra, M.O. Moraes, M.S. Almeida, G. Weber, S. Missailidis, J.L. Silva, V.N. Uversky, A.S. Pinheiro, Y. Cordeiro, Liquid-liquid phase separation and fibrillation of the prion protein modulated by a high-affinity DNA aptamer, FASEB J. 34 (2020) 365–385. https://doi.org/10.1096/fj.201901897R.

[5] M.J. do Amaral, T.S. Araujo, N.C. Díaz, F. Accornero, C.R. Polycarpo, Y. Cordeiro, K.M.S. Cabral, M.S. Almeida, Phase Separation and Disorder-to-Order Transition of Human Brain Expressed X-Linked 3 (hBEX3) in the Presence of Small Fragments of tRNA, J. Mol. Biol. 432 (2020) 2319–2348. https://doi.org/10.1016/j.jmb.2020.02.030.