Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Biological Macromolecules 203 (2022) 466–480

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

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

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

[a] Institute of Medical Biochemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[b] Multiuser Center for Biomolecular Innovation (CMIB), Department of Physics, São Paulo State University (UNESP), São José do Rio Preto, Brazil
[c] National Center of Nuclear Magnetic Resonance (CNRMN), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[d] Faculty of Pharmacy, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[e] Protein Advanced Biochemistry (PAB), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[f] Department of Biochemistry, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[g] Multidisciplinary Center for Research in Biology (NUMPEX), Campus Duque de Caxias Federal University of Rio de Janeiro, Duque de Caxias, Brazil
[h] Laboratory of Toxinology, Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, Brazil
[i] Rio BioNMR Network, Rio de Janeiro, Brazil

ARTICLE INFO

Keywords:
SARS-CoV-2 nucleocapsid protein
DNA/RNA binding protein
Binding specificity

ABSTRACT

The SARS-CoV-2 nucleocapsid protein (N) is a multifunctional promiscuous nucleic acid-binding protein, which plays a major role in nucleocapsid assembly and discontinuous RNA transcription, facilitating the template switch of transcriptional regulatory sequences (TRS). Here, we dissect the structural features of the N protein N-terminal domain (N-NTD) and N-NTD plus the SR-rich motif (N-NTD-SR) upon binding to single and double-stranded TRS DNA, as well as their activities for dsTRS melting and TRS-induced liquid-liquid phase separation (LLPS). Our study gives insights on the specificity for N-NTD(-SR) interaction with TRS. We observed an approximation of the triple-thymidine (TTT) motif of the TRS to β-sheet II, giving rise to an orientation difference of ~25° between dsTRS and non-specific sequence (dsNS). It led to a local unfavorable energetic contribution that might trigger the melting activity. The thermodynamic parameters of binding of ssTRSs and dsTRS suggested that the duplex dissociation of the dsTRS in the binding cleft is entropically favorable. We showed a preference for TRS in the formation of liquid condensates when compared to NS. Moreover, our results on DNA binding may serve as a starting point for the design of inhibitors, including aptamers, against N, a possible therapeutic target essential for the virus infectivity.

* Correspondence to: I.P. Caruso, Multiuser Center for Biomolecular Innovation (CMIB), Department of Physics, São Paulo State University (UNESP), São José do Rio Preto, Brazil.
** Correspondence to: F.C.L. Almeida, National Center of Nuclear Magnetic Resonance (CNRMN), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil.
E-mail addresses: icaro.caruso@unesp.br (I.P. Caruso), falmeida@bioqmed.ufrj.br (F.C.L. Almeida).

https://doi.org/10.1016/j.ijbiomac.2022.01.121
Received 31 October 2021; Received in revised form 17 January 2022; Accepted 18 January 2022
Available online 22 January 2022

1. Introduction

The pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of Covid-19, is still a global health emergency. SARS-CoV-2 is an airborne virus that has rapidly spread around the world, infecting more than 300 million people, and causing more 5.5 million deaths by January 2022. The infection in humans is mainly through inhalation [1,2] and the typical clinical symptoms include fever, cough, sore throat, fatigue, headache, myalgia, and breathlessness. It can progress to severe pneumonia, respiratory failure, and other complications that occasionally leads to death [1,3].

Coronaviruses (CoVs) are enveloped viruses containing a single-stranded positive-sense large genome composed of up to 32 kb. CoV genome replication is a continuous process, whereas transcription requires a complex discontinuous mechanism that culminates in the synthesis of a nested set of subgenomic mRNAs (sgmRNAs) [4]. This discontinuous transcription process is controlled by a specific sequence motif called transcriptional regulatory sequences (TRS) [5]. The common leader TRS (TRS-L) is present at the 5′-untranslated region of the genome, while other TRSs containing conserved core sequences are located upstream of each ORF, named body TRSs (TRS-B). During RNA synthesis, the negative-sense RNA strand containing TRS-B is transferred to and hybridizes with TRS-L. This allows the RNA-dependent RNA polymerase to switch template to TRS-L, generating a set of short negative-sense RNAs that serve as a template for the transcription of positive-sense sgmRNAs [4]. sgmRNAs code for several accessory and structural proteins, including spike (S), membrane (M), envelope (E), and nucleocapsid (N). Recently, subgenomic RNA transcription has been validated for SARS-CoV-2 [6].

N protein is encoded by the most abundant sgmRNA during infection. It is a multifunctional RNA-binding protein that plays critical roles in nucleocapsid assembly as well as in the regulation of RNA replication, transcription, and translation [7]. N is the only viral structural protein found at the replicase-transcriptase complex [8]. N is a protein dimer composed of two globular domains, an N-terminal domain (NTD) and a C-terminal domain (CTD), which are flanked by flexible regions, including the N-terminal arm, the central Ser/Arg rich (SR) intrinsically disordered linker named LKR, and the C-terminal tail (7). The CTD mediates N protein dimerization [9,10]. In addition, the NTD, the CTD, and the intrinsically disordered regions all engage in RNA binding [11]. N protein has been shown to participate in CoVs' subgenomic RNA transcription. Deletion of the N protein-encoding sequence from the viral replicon significantly reduces the levels of sgmRNA [12]. Furthermore, N-NTD specifically interacts with TRS and efficiently melts double-stranded TRS RNA. This RNA chaperone activity is suggested to facilitate template switch and thus regulate discontinuous transcription [10,13–15]. However, the structural basis for TRS binding specificity and the molecular mechanism by which N-NTD melts dsRNA are still poorly understood.

Multivalent RNA-binding proteins, especially those containing intrinsically disordered regions and low complexity domains, have been shown to undergo liquid-liquid phase separation (LLPS) [16]. LLPS is the physical process underlying the formation of membraneless organelles, dynamic micron-sized compartments that provide spatiotemporal control of cellular reactions and organization of biomolecules [17,18]. Interestingly, LLPS has been regarded as a universal nucleation step for the assembly of several viruses, including Measles [19]. Previous reports have demonstrated that SARS-CoV-2 N forms liquid condensates upon RNA interaction [20–23]. RNA-driven LLPS of N is mainly mediated by electrostatic interactions [21] and is regulated by RNA sequence and structure [22]. Remarkably, N liquid droplets recruit and concentrate components of the replication machinery [21]. Despite these recent findings, the role played by TRS binding on N condensation remains unexplored.

Betacoronaviruses N proteins are known to be exported to the nucleus and nucleoli, having conserved nuclear localization signals (NLS) and a nuclear export signal (NES). The N protein of many human coronaviruses has been spotted in the nucleus, where it plays a role in the regulation of sgmRNA synthesis and it is found to associate with ribosomal subunits and nucleolar proteins, such as nucleolin and fibrillarin [11]. The N protein of SARS-CoV and SARS-CoV-2 contains conserved NLS and NES (Fig. S1) [24], but unexpectedly SARS-CoV has never been observed inside the nucleus [25,26] and no study has been reported for SARS-CoV-2. Recently, SARS-CoV-2 N has been located in the perinulear space [27]. In addition, a natural compound with a putative binding site encompassing one NLS of N resulted in a decreased extracellular SARS-CoV-2 titer in cell media [28]. One important feature of N is the promiscuity toward the binding target. It is important for its function of mediating the nucleocapsid assembly. Stringency toward specific dsRNA sequences would prevent an efficient interaction with the virus genome. N binds to single and double-stranded DNA as RNA mimetics [29,30]. Takeda and cols. (2008) and Zhou and cols. (2020) characterized the DNA interaction with the C-terminal domain of N (N-CTD) from SARS-CoV [31] and SARS-CoV-2 [32], respectively, demonstrating the promiscuity of the N protein. N-NTD has specific functions in the recognition of L- and B-TRSs and in template switching. However, it is not yet well understood how the promiscuous N protein can exert such specific functions.

Here, we report on the ability of N-NTD/N-NTD-SR to melt dsTRS DNA and to undergo LLPS upon TRS binding. Critically, N-NTD-SR droplet formation was dependent on the intrinsically disordered SR-rich region. Based on experimental NMR data, we calculated structural models for the interaction of N-NTD with single and double-stranded TRS DNAs. DNA oligonucleotides bind at the positively charged cleft between the palm and the finger, similarly to RNA [15]. Electrostatic interactions play a major role in complex stabilization, specifically by protein-DNA hydrogen bonds and salt bridges. dsTRS binding is enthalpically favorable and entropically unfavorable, while the opposite is observed for ssTRSs. We propose that TRS binding leads to a well-defined positioning of the nucleic acid duplex with a geometry favorable for melting activity. This positioning finely tunes favorable and unfavorable contributions to the binding free energy that may balance the biological activity and explain specificity.

2. Material and methods

2.1. Protein expression and purification

The gene sequences encoding the SARS-CoV-2 (NC_045512.2) nucleocapsid protein N-terminal domain (N-NTD; residues 44–180) and the N-NTD containing the C-terminal SR-rich motif (N-NTD-SR; residues 44–212) were codon-optimized, synthesized, and subcloned into pET28a by GenScript (Piscataway, USA). Expression and purification of these constructs were previously described by the international Covid19-NMR consortium (https://covid19-nmr.de/) [33]. Briefly, pET28a-N-NTD and pET28a-N-NTD-SR were transformed into *Escherichia coli* BL21 (DE3). Cells were grown at 37 °C until optical density ~0.7 at 600 nm. Protein expression was induced with 0.2 mM IPTG for 16 h at 16 °C. For the production of ^{15}N and $^{15}N/^{13}C$-labeled protein, expression was induced in M9 minimal medium containing either $^{15}NH_4Cl$ (1.0 g/L) or $^{15}NH_4Cl$ (1.0 g/L) and ^{13}C-glucose (3.0 g/L). All media contained 50 μg/mL of kanamycin. Cells were harvested by centrifugation (10,000 rpm, 4 °C, 15 min), resuspended in 30 mL of lysis buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole, 10% glycerol, 0.01 mg/mL DNAase, 5 mM MgCl, and 0.1 mM PMSF, and lysed by sonication. The supernatants were clarified by centrifugation (10,000 rpm, 4 °C, 15 min) and subsequently applied to a HisTrap FF column (Cytiva) and purified by nickel-affinity chromatography. The washing buffer contained 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole, and 10% glycerol. The bound proteins were eluted with a linear 20–500 mM imidazole gradient. Fractions containing the proteins of interest were pooled and dialyzed overnight at

4 °C against buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 1.0 mM DTT. Simultaneously, the constructs were cleaved with TEV protease (1:30 TEV:protein molar ratio) to remove the His_6 tag. After dialysis, samples were reapplied to a HisTrap FF column, using the same purification buffers. Fractions containing the proteins of interest were pooled, dialyzed against buffer containing 20 mM sodium phosphate (pH 6.5) and 50 mM NaCl, and concentrated. Protein concentration was determined spectrophotometrically using the molar extinction coefficient of 26,930 $M^{-1}{\cdot}cm^{-1}$ at 280 nm.

2.2. Nucleic acid fragments

In this study, we used a 11-nucleotide DNA sequence corresponding to a 7-nucleotide conserved sequence of the positive-sense ssTRS followed by a CGCG generic segment (ssTRS(+), 5′–TCTAAACCGCG–3′). This CGCG sequence increased the melting temperature of the corresponding TRS duplex and thus enabled us to work at room temperature. The negative-sense ssTRS contained the complementary strand (ssTRS (−), 5′–CGCGGTTTAGA–3′). As a control, we used the single and double-stranded non-specific (NS) oligonucleotides containing the 7-nucleotide DNA sequences used by Dinesh and cols. (2020) [15] followed by the CGCG generic segment (ssNS(+), 5′–CACTGACCGCG–3′; and ssNS(−), 5′–CGCGGTCAGTG–3′). All DNA oligonucleotides were purchased from GenScript (Piscataway, USA). Equimolar amounts of positive and negative sense ssDNAs were dissolved in buffer and annealed by heating to 50 °C for 10 min and slowly cooling to room temperature. The concentrations of ssTRS(+), ssTRS(−), ssNS(+), and ssNS(−) were determined spectrophotometrically using the molar extinction coefficients of 103,000, 107,500, 99,100, and 103,000 $M^{-1}{\cdot}cm^{-1}$ at 260 nm, respectively. For measurements of melting activity, dsTRS RNA (sense strand 5′–UCUAAACCGCG–3′ and antisense strand 5′–CGCGGUUUAGA–3′) was also used.

2.3. Fluorescence spectroscopy

Protein intrinsic fluorescence quenching measurements were performed using a PC1 steady-state spectrofluorimeter (ISS, Champaign, IL, USA) equipped with a quartz cell of 1.0 cm optical path length and a Neslab RTE-221 thermostat bath. Excitation and emission bandwidths were set to 1.0 and 2.0 mm, respectively. To solely excite tryptophan, the excitation wavelength (λ_{ex}) was set to 295 nm. The emission spectrum was collected in the range of 305–500 nm with an increment of 1.0 nm. Each point in the emission spectrum is the average of 10 accumulations. Increasing concentrations of DNA oligonucleotides were added to a 1.0 μM protein solution in 20 mM Bis-Tris buffer (pH 6.5). The DNA concentration varied from 0 to 3.8 μM, and experiments were performed in either duplicate or triplicate. Titration measurements were performed at 15, 25, and 35 °C for ssDNAs, and 15, 20, and 25 °C for dsDNAs. The effect of salt (NaCl) and inorganic phosphate (Pi) on the N-NTD:ssDNA complex formation was investigated at 25 °C. The binding isotherms were fitted using the following equation after data treatment [34]:

$$\frac{F_0 - F}{F_0 - F_S} = \frac{(K_d + [L_T] + [P_T]) - \sqrt{(K_d + [L_T] + [P_T])^2 - 4[L_T][P_T]}}{2[P_T]} \quad (1)$$

where F is the measured fluorescence intensity in the presence of DNA, F_0 is the fluorescence in the absence of DNA, F_S is the fluorescence emission of the protein saturated with DNA, K_d is the dissociation constant, $[L_T]$ is the total ligand concentration (DNA), and $[P_T]$ is the total protein concentration (N-NTD or N-NTD-SR). The K_d values were calculated from the fitting process by nonlinear least-squares optimization using Levenberg–Marquardt interactions in the software OriginPro 2021. The global fitting was performed for the protein:DNA binding isotherms at three temperatures using as a constraint the linearity of the van't Hoff plot, $\ln(K_d)$ *versus* T^{-1}, where T is the temperature in Kelvin. The linear constraint is based on the non-significant variation of the enthalpy change (ΔH) in the temperature range studied (15–35 °C for ssDNAs, and 15–25 °C for dsDNAs). Next, ΔH values were determined from the van't Hoff equation:

$$\frac{d}{d\frac{1}{T}} ln\left(\frac{1}{K_d}\right) = -\frac{\Delta H}{R} \quad (2)$$

where R is the universal gas constant. The values of the Gibbs free energy changes (ΔG) and entropy change (ΔS) were determined from $\Delta G = RTln(K_d)$ and $T\Delta S = \Delta H - \Delta G$.

Prior to fitting, the fluorescence quenching data were corrected for the background buffer fluorescence, the inner filter effects [35], and for removing an upward linear contribution from the fluorescence titration curves at the largest DNA concentrations (Fig. S2). This near-linear contribution arises due to the difference in the accessibility of the fluorophores (W52, W108, and W132) by the quencher (DNAs). Two of the three fluorophores (W52 and W108) are located at the putative nucleic acid-binding site in N-NTD, while the third (W132) is oppositely oriented to this binding site and thus presents a non-specific linear contribution for the fluorescence titration curves. To remove the non-specific contribution, F_0/F *versus* $[L_T]$ curves for each protein:DNA titration (duplicate or triplicate) was fitted using Eq. (1) plus a linear term ($constant{\cdot}[L_T]$). After that, F_0/F linear contributions were calculated and removed manually (Fig. S2).

Fluorescence resonance energy transfer (FRET) experiments were performed on a Cary Eclipse spectrofluorometer (Agilent Technologies, Santa Clara, CA, USA) equipped with a quartz cell of 1.0 cm optical path length and Peltier temperature controller at 20 °C. A bandwidth of 5 nm was set for both excitation and emission. FRET measurements were carried out with DNA and RNA oligonucleotides labeled at the 5′-end of the positive-sense TRS strand with the donor probe Q570 (Q570-ssTRS (+)) and at the 3′-end of the negative-sense TRS strand with the acceptor probe Q670 (Q670-ssTRS(−)). The donor Q570 was excited at 535 nm and the emission spectrum was collected in the 550–800 nm range, observing the FRET efficiency (E) of the acceptor Q670 at maximum emission at 670 nm. The oligonucleotide concentration was set to 50 nM, while the final concentrations of protein was in the range of 9–17 μM. Fluorescence measurements were recorded in 20 mM Bis-Tris buffer (pH 6.5) in the absence or presence of 100 mM NaCl, and in 20 mM sodium phosphate buffer (pH 6.5) containing 50 and 100 mM NaCl. The FRET efficiency was calculated from $E = 1 - F_{DA}/F_D$, where F_D and F_{DA} are the fluorescence intensities of the donor Q570 in the absence and presence of the acceptor Q670, respectively.

2.4. NMR spectroscopy

Chemical shift perturbations (CSP) were monitored by a series of 2D [^{1}H,^{15}N] HSQC spectra, recorded on a Bruker 600 MHz spectrometer at 20 °C, after the addition of increasing concentrations of ssDNAs (TRS and NS). N-NTD and N-NTD-SR were dissolved in 20 mM sodium phosphate buffer (pH 6.5), 50 mM NaCl, 500 μM PMSF, 3 mM sodium azide, 3 mM EDTA, and 5.5% (v/v) of D_2O, at a final concentration of 70 μM. The ssDNAs were titrated into N-NTD and N-NTD-SR for the following ssDNA:protein molar ratios: 0.14, 0.34, 0.77, 1.0, 1.6, 2.5, and 3.0. For dsDNAs (280 μM), CSPs were calculated from a protein sample containing 4× molar excess of dsDNA. Spectra were processed with NMRPipe [36], and analyzed with CCPNMR Analysis [37] using the NMR resources platform NMRBox [38]. The chemical shift of water was used as an internal reference for ^{1}H, while the ^{15}N chemical shift was referenced indirectly to water (Wishart et al., 1995). Three [^{1}H,^{15}N] HSQC spectra were collected for the free protein and used to determine the experimental error. The chemical shift perturbations (CSP, $\Delta\delta$) of all assigned amides were calculated by:

$$\Delta\delta = \sqrt{\frac{\Delta\delta_{HN}^2 + (0.15{\cdot}\Delta\delta_N)^2}{2}} \quad (3)$$

where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the chemical shift differences of ^{1}H and ^{15}N, respectively, recorded in the absence and presence of DNA. The standard deviation of the average weighted $\Delta\delta$ values for all assigned amides was used as a cutoff to identify the protein residues involved in DNA binding. Residues showing fast exchange in the presence of DNA were used for the calculation of K_d from Eq. (1).

2.5. Computation simulations

The HADDOCK (version 2.4) server [39] was used to construct a structural model of the N-NTD complex with TRS and NS DNAs. The protein structural coordinates used as input were obtained from the Protein Data Bank (PDB) under access code 6YI3 [15]. The DNA structures used in the docking calculations were obtained from the 3DNA server (http://web.x3dna.org/). The ^{1}H, ^{15}N chemical shifts of residues forming the DNA-binding interface (identified by the CSP analysis) were used as ambiguous restrains for the docking. Residues with chemical shift perturbation $\Delta\delta > \Delta\delta_{av} + SD$ were classified as active and the passive residues were automatically selected by the HADDOCK server. In total, 2000 complex structures of rigid-body docking were calculated by using the standard HADDOCK protocol with an optimized potential for liquid simulation (OPLSX) parameters [40]. The final 200 lowest-energy structures were selected for subsequent explicit solvent (water) and semi-flexible simulated annealing refinement, to optimize side chain constants. The final structures were clustered using the fraction of common contacts (FCC) with a cutoff of 0.6 [41].

Molecular dynamics (MD) calculations for the N-NTD:DNA complexes were performed using GROMACS (version 5.1.4) [42]. The molecular systems were modeled with the corrected AMBER14-OL15 package, including the ff14sb protein [43] and ff99bsc0OL15 DNA [44] force fields, as well as the TIP3P water model [45]. The structural models of the N-NTD:DNA complexes (from molecular docking) were placed in the center of a cubic box solvated by a solution of 50 mM NaCl in water. The protonation state of ionizable residues at pH 7.0 was set according to the PROPKA server [46]. Periodic boundary conditions were used, and all simulations were performed in NPT ensemble, keeping the system at 25 °C and 1.0 bar using Nose-Hoover thermostat ($\tau_T = 2$ ps) and Parrinello-Rahman barostat ($\tau_P = 2$ ps and compressibility $= 4.5 \times 10^{-5}$·bar^{-1}). A cutoff of 12 Å for both Lennard-Jones and Coulomb potentials was used. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) algorithm. In every MD simulation, a time step of 2.0 fs was used and all covalent bonds involving hydrogen atoms were constrained to their equilibrium distance. A conjugate gradient minimization algorithm was used to relax the superposition of atoms generated in the box construction process. Energy minimizations were carried out with the steepest descent integrator and conjugate gradient algorithm, using 1000 kJ·mol^{-1}·nm^{-1} as maximum force criterion. One hundred thousand steps of molecular dynamics were performed for each NVT and NPT equilibration, applying force constants of 1000 kJ·mol^{-1}·nm^{-2} to all heavy atoms of N-NTD:DNA complexes. At the end of preparation, 2 μs MD simulations of each molecular system (N-NTD bound to ssTRSs, dsTRS, ssNS, and dsNS) were carried out for data acquisition. Following dynamics, the trajectories of each molecular system were firstly concatenated and analyzed according to the root mean square deviation (RMSD) for the backbone atoms of protein and DNA, number of contacts for distances lower than 0.6 nm between pairs of atoms of N-NTD and DNA, and number of protein-DNA hydrogen bonds with cutoff distance (heavy atoms) of 3.5 Å and maximum angle of 30°. The percentages of protein-DNA hydrogen bond persistence were obtained from *plot_hbmap_generic.pl* script [47]. The number of protein-DNA hydrogen bonds with persistence greater than 10% was counted with respect to amino acid and nucleotide residues for each molecular system. The contributions of N-NTD residues to the DNA-binding energy were calculated from the MD trajectories using the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method implemented in the *g_mmpbsa* program, along with the *MmPbSaDecomp.py* script [48,49]. The tool *g_cluster* of GROMACS package [50] was used to perform a cluster analysis in the 2.0 μs MD trajectories of the N-NTD:DNA complexes. The *gromos* algorithm [51] was used to generate the clusters with cutoff of 3 and 4 Å for ssDNA and dsDNA, respectively. The RMSD was calculated between all the structures and superimposed considering only non-hydrogen atoms. The principal component analysis (PCA) was applied to investigated conformational space of the protein in the interaction with DNA. PCA scatter plots were generated, and conformational motions were filtered from the eigenvectors of the first and second principal components (PC1 and PC2, respectively). The conformational space was quantified by fitting an elliptical shell with 95% (confidence) of the density for each scatter plot and making its extent proportional to the area (S_{el}) of this shell. The structural representations of the constructed models were displayed using PyMOL [52].

2.6. Microscopy and turbidity for LLPS study

Cleaning of microscopy glasses followed by coating of Knittel coverslips with 0.5% (*w/v*) bovine serum albumin (BSA, Sigma-Aldrich A2153) was carried out to provide a hydrophilic surface. Specifically, coverslips were incubated for 1 h with the BSA solution and rinsed twice with Milli-Q water. Subsequently, 5 × 22 mm imaging chambers were mounted onto glass slides using double-sided tapes to stick coverslips as reported in [53,54]. The protein samples of N-NTD and N-NTD-SR at 20 μM were prepared in low-binding microtubes (Eppendorf® LoBind) containing 20 mM Tris-HCl buffer (pH 7.5) with 30 mM NaCl, specified PEG-4000 concentrations (Thermo Scientific, USA), and/or reactants such as 10% (w/v) 1,6-hexanediol (Merck, Germany), 300 mM NaCl. The following nucleic acids were used: a pool of RNA consisting of 5000–8000 Da from Torula yeast (Sigma-Aldrich, R6625), ssTRS(+), ssTRS(−), ssNS(+), and ssNS(−) with concentrations ranging from 2.5 to 40 μM according to the protein:nucleic acid molar ratios (8:1, 4:1, 2:1, 1:1, and 1:2). The dsTRS and dsNS were produced as mentioned above. Immediately following nucleic acid addition to N-NTD solution, a 15 μL aliquot was pipetted through one opening of the rectangular chamber and settle to stand for 30 min with coverslip surface facing down to let protein condensates stand by gravity followed by imaging and concomitant turbidity measurement. All samples were imaged at time 30 min at room temperature. The working concentration of the fluorescent dye DAPI (BioRad, #135–1303, stock in Milli-Q water) was 0.25 μg/mL and the emission recorded using a LED light fluorescent cube (excitation centered at 357/44 and emission at 447/60 nm). Samples described above were imaged in an inverted phase contrast microscope (EVOS® FL Cell Imaging System, Thermo Fisher Scientific, USA) with a 40× apo objective. To quantify the number of condensates, five areas of 100 μm^2 were analyzed for each sample followed by Fourier bandpass filtering. Next, images had their threshold adjusted for liquid droplets correct recognition, followed by mask creation and the "fill holes" (fills droplets that were poorly delimited) and "watershed" operations (separate droplets undergoing fusion). The condensates number (above 0.5 μm condensates were counted) and size (area in μm^2) were determined using the "analyze particles" plugin. All quantification and image processing steps were performed in Fiji. The DAPI fluorescence images were obtained with 1% excitation intensity, apart from the crystal sample formed by 1:1 N-NTD-SR:dsNS which was excited with 20% laser power. Representative micrographs from three independent experiments had their brightness/contrast corrected by histogram stretching. Data were analyzed with GraphPad Prism (version 8.1.1; GraphPad Software).

Turbidity (detected as absorbance at 350 nm) was performed concomitant to imaging time (30 min). Samples of 20 μM N-NTD or N-NTD-SR consisting of varied protein:nucleic acid molar ratios (8:1, 4:1, 2:1, 1:1, 1:2) were prepared in low-binding microtubes (Eppendorf® LoBind) in the presence of 10% PEG-4000 (w/v) in 20 mM Tris-HCl

buffer (pH 7.5) with 30 mM NaCl. Three technical replicates were measured using 5 μL of samples in a μLite sample port (0.5 mm optical path) of a BioDrop Duo spectrophotometer (Biochrom, UK). Protein: nucleic acid sample readings were subtracted from samples containing all components apart from protein. Data were analyzed with GraphPad Prism (version 8.1.1; GraphPad Software).

3. Results

3.1. N-NTD(-SR) melts double-stranded DNA and RNA with similar efficiency

We first measured the dsDNA and dsRNA melting activity of N-NTD, as well as N-NTD containing the SR-rich motif (N-NTD-SR), using fluorescence resonance energy transfer (FRET) from the 5′ Q570-labeled positive-sense TRS (Q570-ssTRS(+)) to the 3′ Q670-labeled negative-sense TRS (Q670-ssTRS(−)) (Fig. 1A). In the absence of the protein, the FRET pair Q570/Q670 is in proximity, displaying FRET efficiency for dsTRS. Addition of N-NTD or N-NTD-SR results in an increase in fluorescence emission at 570 nm with a concomitant decrease at 670 nm, leading to a decay in FRET efficiency and, consequently, indicating DNA and RNA duplex melting (Fig. 1B). Unexpectedly, the N-NTD-SR melting activity toward dsTRS RNA and DNA presented similar efficiency (Fig. 1C), suggesting that this domain has a promiscuous melting activity. Studies of the interaction of N-NTD or N-NTD-SR with DNA could provide insights on the structural features responsible for the promiscuity of the N protein, which is important for nucleocapsid assembly, and yet, on the specific TRS recognition, important for the discontinuous transcripconstion through the template switch. The similar melting activity suggests that despite of the structural differences between duplex DNA (B-double helix) and duplex RNA (A-double helix), the nucleic acid/protein interaction displays similar features. The study with DNA is complementary to the available data on the N protein interaction with RNA [7,15,20]. We also compared the dsTRS melting activity of N-NTD and N-NTD-SR. Fig. 1C shows that the N-NTD and N-NTD-SR of SARS-CoV-2 has similar dsTRS melting activity. For Mouse Hepatitis Virus (MHV), a virus most closely related to SARS-CoVs, the melting activity of N-NTD-SR is higher than that observed for N-NTD [7].

3.2. N-NTD and N-NTD-SR bind tighter to ssTRS(−) than ssTRS(+) DNA

We investigated the thermodynamics of N-NTD/N-NTD-SR binding to the specific (ssTRS(+), ssTRS(−), and dsTRS) and non-specific DNA sequences (ssNS(+), ssNS(−), and dsNS), using intrinsic tryptophan fluorescence. Using 20 mM Bis-Tris buffer, titration of DNA oligonucleotides into a protein solution resulted in fluorescence quenching until saturation, indicating protein:DNA association. After data treatment, the binding isotherms exhibited a hyperbolic shape (Figs. 2A and S3). Formation of N-NTD:DNA complexes showed salt dependence. We observed a fluorescence recovery upon the increase of inorganic phosphate (Pi) and sodium chloride (NaCl) concentrations (Fig. 2B and C). The salt and phosphate dependence suggests a large contribution of electrostatic interactions to the binding, being essential for complex formation. Despite the high salt and phosphate dependence, N-NTD-SR remains active, independently of the buffer (20 mM Bis-Tris or 20 mM NaPi) and salt concentration (0, 50, or 100 mM NaCl) (Figs. 1C and S4).

The protein:DNA binding affinities are at the nanomolar range (Table S7). The affinity for ssTRS(−) is higher than that for ssTRS(+), being 4–5-fold for N-NTD and 15–16-fold for N-NTD-SR. This tendency was also observed for the NS DNA, albeit with a smaller magnitude. This observation is remarkable since ssTRS(−) is the strand transferred to TRS-L during template switch [4]. The thermodynamic parameters are similar for both N-NTD and N-NTD-SR binding to TRS. For ssTRSs, the interaction is enthalpically unfavorable ($\Delta H > 0$) and entropically driven ($\Delta S > 0$), while for dsTRS, it is enthalpically driven ($\Delta H < 0$) and entropically unfavorable ($\Delta S < 0$). The analysis of these enthalpic and entropic contributions reveals that hydrophobic contacts are important for N-NTD/N-NTD-SR binding to ssTRSs, while hydrogen bonds (and salt bridges) and van der Waals interactions play a major role in stabilizing the interaction with dsTRS [55]. Based on this observation, we hypothesize that the destabilization of dsTRS Watson-Crick hydrogen bonds induced by N-NTD or N-NTD-SR binding would be entropically driven. It would also be enthalpically unfavorable, which might contribute to further dissociation of ssTRSs.

In general, we observed a higher affinity for TRS than for NS in the interaction with N-NTD and N-NTD-SR, especially for the duplex DNA (Table S7). For N-NTD-SR, the affinity is even higher, especially for ssTRS(−), suggesting that the arginine residues in the SR-rich motif may contribute to binding with non-specific electrostatic interactions. Unlike the thermodynamic pattern observed for N-NTD/N-NTD-SR binding to the specific sequences (TRSs), we did not observe a regular pattern for the enthalpic and entropic contributions to the interaction with NS DNAs (Fig. 2D and Table S1). Although a thermodynamic profile similar to that of TRS was observed for N-NTD binding to NS DNAs, a different energetic profile was observed for N-NTD-SR.

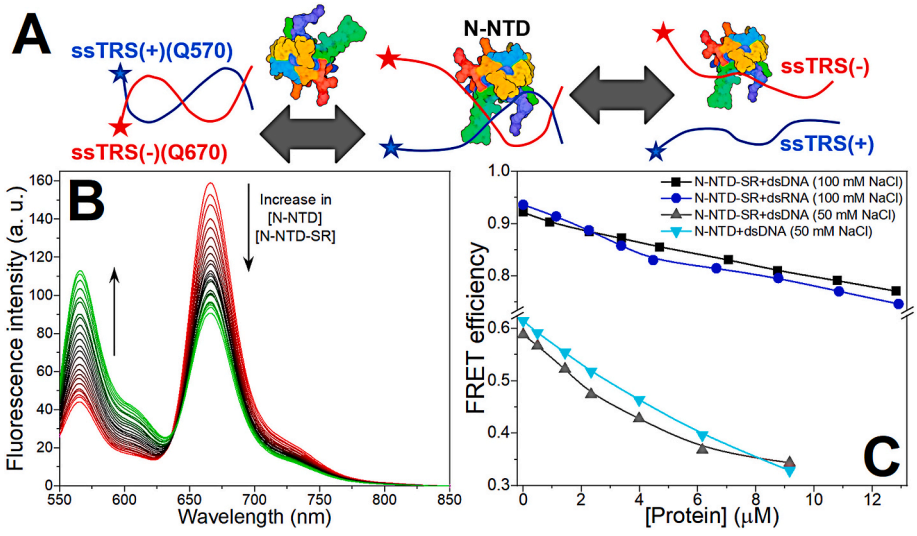

Fig. 1. dsDNA/dsRNA melting activity of N-NTD and N-NTD-SR. (A) Cartoon model of the dsTRS melting activity of N-NTD(-SR). The ssTRSs are presented as curved lines and the fluorescent probes (Q570 and Q670) as stars. ssTRS(+) and ssTRS(−) are colored in blue and red, respectively. The protein is denoted as a colored surface. (B) Fluorescence spectra of the Q570-ssTRS(+)/Q670-ssTRS(−) DNA duplex (50 nM) at absence and presence of N-NTD or N-NTD-SR (λ_{exc} = 535 nm and temperature at 20 °C). The arrows denote the changes in the spectra as protein is added. (C) FRET efficiency as a function of the total concentration of N-NTD-SR for dsDNA (black squares) and dsRNA (blue circles) melting activity in 20 mM sodium phosphate buffer (pH 6.5) containing 100 mM NaCl, and as a function of total concentration of N-NTD-SR (gray triangles up) and N-NTD (cyan triangle down) for dsDNA melting activity in 20 mM sodium phosphate buffer (pH 6.5) containing 50 mM NaCl.

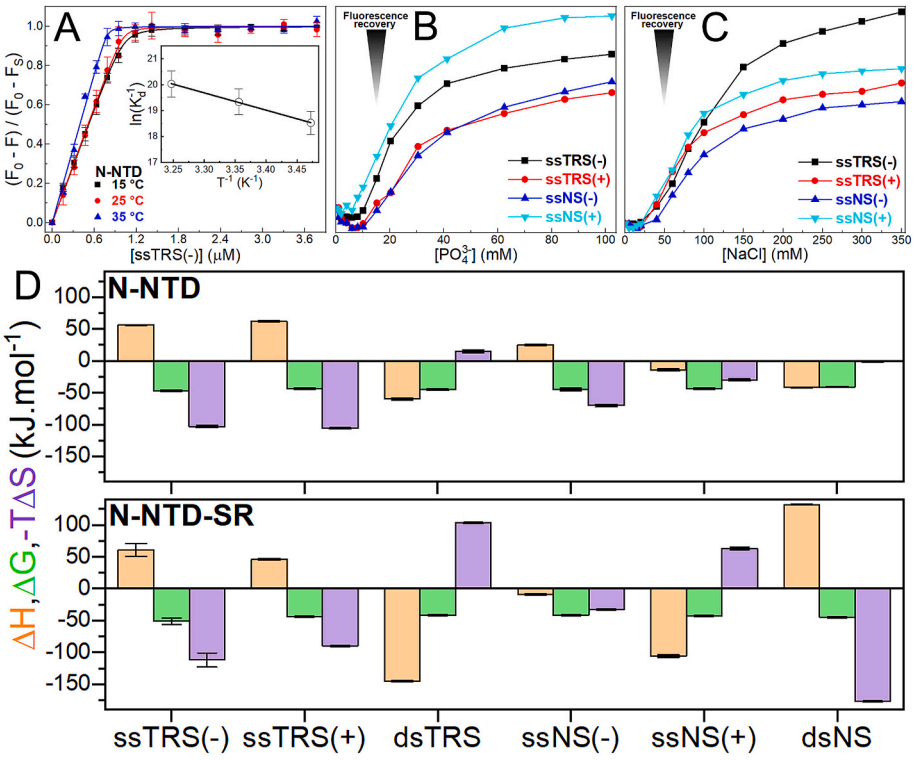

Fig. 2. Thermodynamic analysis of the binding of TRS and NS DNAs to N-NTD and N-NTD-SR. (A) Protein intrinsic fluorescence quenching changes of N-NTD as a function of the ssTRS(−) concentration in 20 mM Bis-Tris buffer (pH 6.5) at 15, 25, and 35 °C. Each point on the binding isotherm represents the average and standard error calculated from triplicate 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. (B, C) Fluorescence recovery as a function of inorganic phosphate (Pi) and sodium chloride (NaCl) concentration for the formation of the N-NTD:ssDNA complexes. (D) Thermodynamic parameters for the interaction of TRS and NS DNAs with the N-NTD and N-NTD-SR at 25 °C. For each DNA, the orange bar on the left denotes the enthalpy change (ΔH); the green bar in the middle, the Gibbs free energy change (ΔG); and the purple on the right, the entropic term ($T\Delta S$). The error bars represent the standard deviation calculated from duplicate or triplicate measurements.

3.3. Mapping the N-NTD and N-NTD-SR interaction with DNA

To map the residues involved in N-NTD and N-NTD-SR interaction with DNA, single and double-stranded DNA oligonucleotides were titrated into ^{15}N-labeled protein samples and amide chemical shifts were measured from 2D [^{1}H,^{15}N] HSQC spectra (Fig. S5). Residues displaying statistically significant chemical shift perturbation (CSP) values (higher than the average plus one standard deviation (SD)) are located in well-defined regions (Figs. 3 and S6). Using the experimental CSP data, we modeled the interaction of N-NTD with the DNAs (Fig. 3G–I). The structural models were fundamental to discriminate the interacting amino acid residues from the ones that seem to undergo allosteric effects upon DNA binding. The binding site includes: (*i*) the flexible N-terminal region (A42, N47, N48, T49), (*ii*) the palm formed by β-sheet I (β2/β3/β4) (Y86, Y87, R89, R107, W108, Y109, F110, Y111, Y112, G129, I131, W132, A134) and the N-terminal residues (A50, S51, W52, F53, T54, A55), (*iii*) the finger (β2/β3 loop) (A90, T91, R92, R93, I94, R95, G96, K100, K102, D103, L104, S105), (*iv*) the β-sheet II (β1/β5) (L56, T57, Q58, H59, G170, Y172, A173, E174), (*v*) the thumb (α2/β5 loop) (R149, N150, A152, A156, I157, L159, Q160), β3/β4 loop (L113, G114, E118), and (*vi*) the C-terminal region (G175, S176, R177). We also observed significant CSPs for residues L64, K65, F66, G71, Q163, T165, T166, and L167, which are located in a remote region from the principal binding site (Fig. 3G and H), suggesting that these residues form either a secondary binding site, interacting with DNA directly, or an allosteric site, undergoing indirect conformational changes due to DNA binding. In addition, residues at the SR motif (N192, S193, N196, and S197) displayed significant CSPs, suggesting that this region engages in DNA binding. The observed CSPs are strikingly similar (even for the remote region) to those observed by Dinesh and cols. (2020) for a 7-mer and a 10-mer ssTRS(+) RNA (5′–CUAAACG–3′, 5′–UCUCUAAACG–3′), and a 7-mer non-specific RNA duplex (5′–CACUGAC–3′ and 5′–GUCAGUG–3′) [15], suggesting that DNA and RNA bind at the same interface on N-NTD.

To find a consensus amino acid sequence for the interaction of N-NTD and N-NTD-SR with ssDNAs and dsDNAs, we aligned the residues with CSPs larger than the average plus one SD and with resonance signals broadened beyond detection upon DNA binding (Fig. S7). To identify which residues are common to all interactions and which are unique, we grouped the different titration results and plotted them as an intersecting set of residues represented by circles containing the CSP information for each titration (Fig. 4). When we analyzed the set of intersections for N-NTD and N-NTD-SR with single and duplex TRS, as well as with the non-specific (NS) sequences (Fig. 4A, B, C, and D), the following residues stood out: (*i*) A152 in the thumb is present in all 4 intersections; (*ii*) K102 in the finger and T166 in the remote region from the main binding site are present in 3 intersections; (*iii*) T57 in the β-sheet II, L64 and K65 in the remote region, R95 and K100 in the finger, Y111 in the interface between the palm and the β-sheet II, A156 in the thumb, and A173 in the β-sheet II are present in 2 intersections. The regions mapped by the residues at the intersections involve the finger, the two elements of the palm, and the flexible thumb of N-NTD (Fig. 4E). These residues may be key for N-NTD binding to several nucleic acids, independent of the sequence specificity. Furthermore, they comprise not only positively charged residues, responsible for electrostatic interactions, but also hydrophobic residues (L64, T56, Y111, A152, A156, T166, and A173).

We used CSP as a function of DNA concentration to estimate the K_d for the interaction with N-NTD (Table S8). In contrast to the nanomolar affinities measured in the absence of NaCl and inorganic phosphate (Table S7), we observed apparent dissociation constants in the order of micromolar, and yet, despite the difference in affinities, the protein is active (melting activity) in all tested conditions (Figs. 1 and S4). The error of the K_d measurements from CSP are in general high, making it more difficult to infer statistically significant differences. Nevertheless, the data suggested a similar tendency of K_d variations for the samples at high salt concentration (Table S8) as the observed for those at low salt concentration (Table S7). In general, the affinities for ssDNA(−) are higher than those for ssDNA(+), with statistical significance for just few of the measurements. The same is valid for the observation that ssTRSs display greater affinities than those observed for ssNSs.

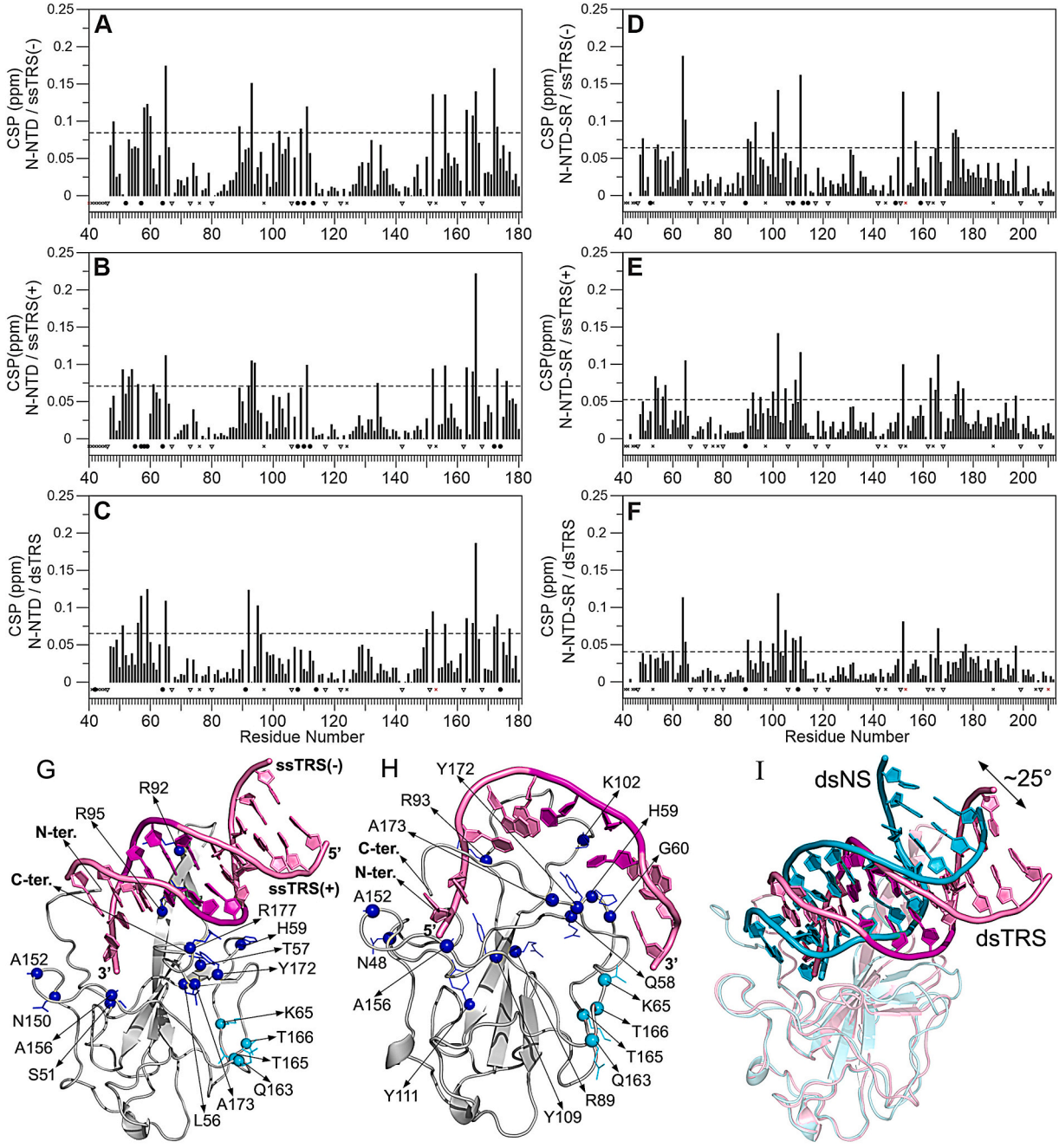

Fig. 3. Mapping the residues involved in the protein/TRS interaction. Chemical shift perturbation (CSP) for the interaction of the N-NTD with (A) ssTRS(−), (B) ssTRS(+), and (C) dsTRS, and of the N-NTD-SR with (D) ssTRS(−), (E) ssTRS(+), and (F) dsTRS. The dotted line denotes the average CSP value ($\Delta\delta_{ave}$) plus one standard deviation (SD), which is the cutoff used to identify the most significant residues involved in the binding to DNA. The proline residues (46, 67, 73, 80, 106, 117, 122, 142, 151, 162, 168, 199, and 207) are indicated by triangle. Resonances signals broadened beyond detection upon TRS titration are represented by the filled black circles. Representative structural model of the N-NTD:DNA complexes for (G) dsTRS and (H) ssTRS(−) 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). (I) Comparison of the position and orientation for dsTRS (light pink) and dsNS (cyan) with a difference of ~25° between them. 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.

3.4. Hydrogen bonds and energy contributions to N-NTD:DNA binding

To better describe the interactions, we used the experimentally CSP-mapped binding interface to generate structural models of the N-NTD:DNA complexes using molecular docking followed by 2.0 μs molecular dynamics (MD) simulations. It is noteworthy that there is no experimental structure available for the complex between N-NTD and nucleic acids. Dinesh and cols. [15] presented structural models for the interaction of N-NTD with ds and ssRNAs, which are used throughout this paper as a reference, but are also NMR-derived docking models followed by restrained MD simulation. Structural models of the complexes generated by docking calculations correspond to the lowest energy structure from the cluster with the lowest HADDOCK score (fraction of common contacts >0.8, interface-RMSD <0.2 Å, and ligand-RMSD <5.0 Å). Next, we performed 2 μs MD simulations for the docking models. This simulation time is long enough to probe the structural stability of the models, but not to explore all the conformational space. With that in mind, we showed that N-NTD complexes with both ssDNA and dsDNA

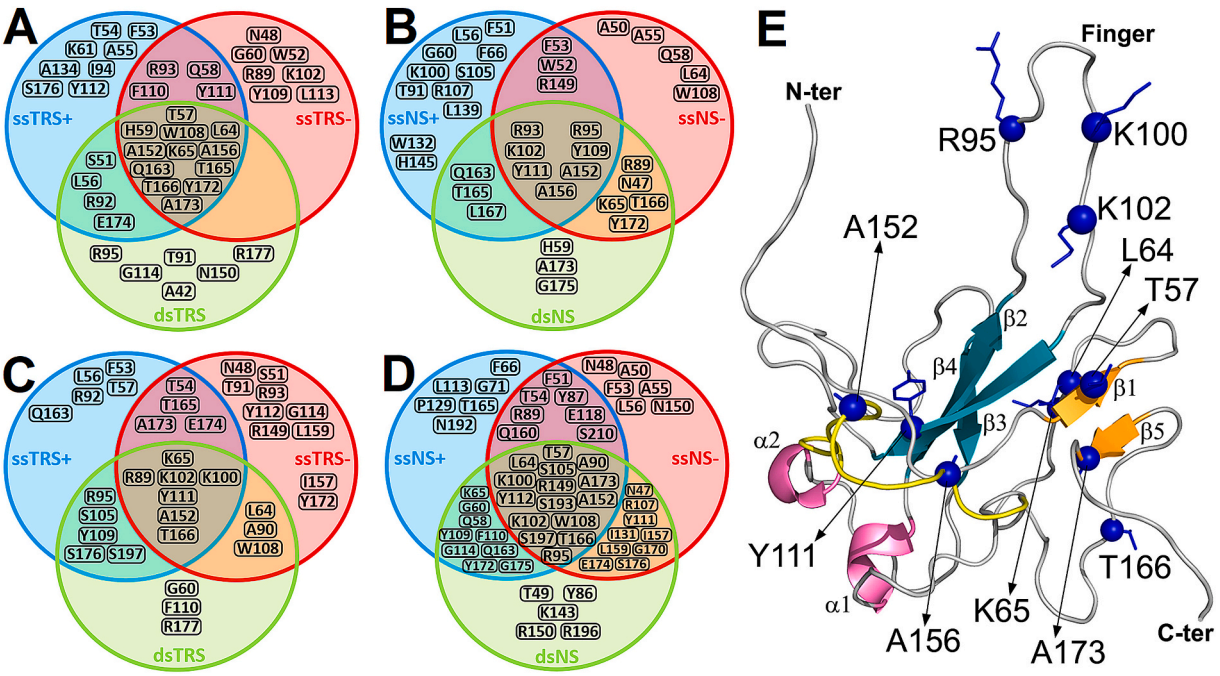

Fig. 4. Consensus of residues involved in the protein:DNA binding. Intersecting set of residues represented by circles containing the significant CSP information (higher than $\Delta\delta_{ave}$ + SD or resonances signals broadened beyond detection upon DNA binding) for the interaction of the TRS and NS with (A and B) N-NTD and (C and D) N-NTD-SR. The blue, red, and green sphere denotes the CSP set for ssDNA(+), ssDNA(−), and ds DNA, respectively. (E) The DNA-binding residues most recurrent in the intersections for the protein:DNA binding are indicated as blue spheres on the N-NTD structure. The protein is shown as a cartoon model with the β-sheet I (β2/β3/β4), β-sheet II (β1/β5), and α-helices (α1 and α2) colored in dark green, light orange, and magenta, respectively. The thumb (residues I146–V158 in α2/β5) is colored in yellow.

reach stability at the end of the 2 μs MD simulations, exhibiting stable values of RMSD, number of protein-DNA contacts, and number of protein-DNA hydrogen bonds (Figs. S8, S9, and S10). Thus, we determined representative structural models of the N-NTD:DNA complexes from the clustering of the MD trajectories (Figs. 3G, H, and S11). The representative models are consensual in many aspects. For dsDNA, the major groove is recognized by the finger, and the 5′-end of the positive-sense strand is oriented toward the palm. We observed a difference of ~25° in orientation of dsTRS and dsNS in the representative structure models (Fig. 3I), positioning the dsNS close to the finger and dsTRS buried into the palm. The β-sheet II (β1/β5) interacts with the negative-sense DNA strand, especially with the TTT motif, suggesting that it might be involved in TRS recognition. For ssDNA, the 5′-end is oriented toward the palm and the finger interacts with a stretch of nucleotides spanning 5 residues from the 3rd to the 7th register. At least the last 3 nucleotides are free and do not interact with N-NTD. It is worth mentioning that the consensual structures of N-NTD:dsDNA complexes observed here are similar in ligand position and orientation compared to the consensual structure with the RNA duplex (PDB id 7ACS) [15].

The MD trajectories may report, not all, but important interactions related to the binding stability and specificity. Thus, we determined protein-DNA hydrogen bonds from the 2 μs MD trajectories and calculated their percentages of persistence along the simulations (Fig. S9). Protein-DNA hydrogen bonds with persistence higher that 10% were counted with respect to amino acid and nucleotide residues (Fig. 5, Tables S1–S6). It is noteworthy that protein-DNA hydrogen bonds also report the presence of salt bridges between arginine or lysine residues and DNA phosphate groups (Fig. 5G and H). For both dsDNA and ssDNA complexes, the presence of persistent protein-DNA hydrogen bonds mirrored the CSP results for the following N-NTD regions: the flexible N-terminal, the palm formed by β-sheet I, the finger, the β-sheet II, and the C-terminal region. The remote region and the thumb displayed significant CSPs but no persistent protein-DNA hydrogen bonds or salt bridges (Fig. 5). This feature reveals that non-polar interactions contribute to binding of the thumb to DNA and that the remote region is not directly involved in the recognition interface. Binding of DNA to the cleft formed between the finger and the palm is mostly mediated by positively charged residues. Residues R92, R107, Y172, and R177 are hydrogen bonded to the DNA oligonucleotide in most MD runs (Fig. 5I). Remarkably, mutations R92E and R107E almost abolished N-NTD interaction with RNA [15]. Residues E174 and K61 make hydrogen bonds almost exclusively with TRS. The negatively charged E174 (conserved in human CoVs, Fig. S1) is hydrogen bonded to the nitrogenous bases, positioned approximately in the middle of the two strands. By using molecular dynamic simulations, we suggested previously that this residue may contribute to the dsRNA melting activity of N-NTD [14].

The positive-sense strands of dsDNA (*middle*, Fig. 5A and B) and ssDNA (Fig. 5C and D) are mainly hydrogen bonded to N-NTD through their 5′ termini (nucleotide 1 to 4). Conversely, the negative-sense strands of dsDNA are mainly recognized by N-NTD through hydrogen bonds with their 3′ termini (nucleotides 4 to 9) (*bottom*, Fig. 5A and B), which include the specific TTT motif for TRS. It is interesting to note that binding of N-NTD to positions 1 to 5 of the positive-sense strand and 4 to 9 of the negative-sense strand is maintained independently of the nucleotide sequence (TRS or NS), suggesting that the orientation of dsDNA with respect to the positively charged cleft is conserved and sequence independent. We observed consistently that the total number of protein-DNA persistent hydrogen bonds with ssDNA(−) (Fig. 5E and F) is significantly larger than that with ssDNA(+) (Fig. 5C and D). These results are in agreement with the observation that ssDNA(−) shows higher affinity to N-NTD and N-NTD-SR than the ssDNA(+) (Table S7). The same tendency is observed for the negative-sense strands of dsDNAs (Fig. 5A and B).

Due to the high content of charged residues and the above-mentioned electrostatic contribution (salt and phosphate dependence, Fig. 2B and C) to DNA recognition, we decided to compute the theoretical Gibbs free energy of binding ($\Delta G_b{}^t$) using Poisson-Boltzmann

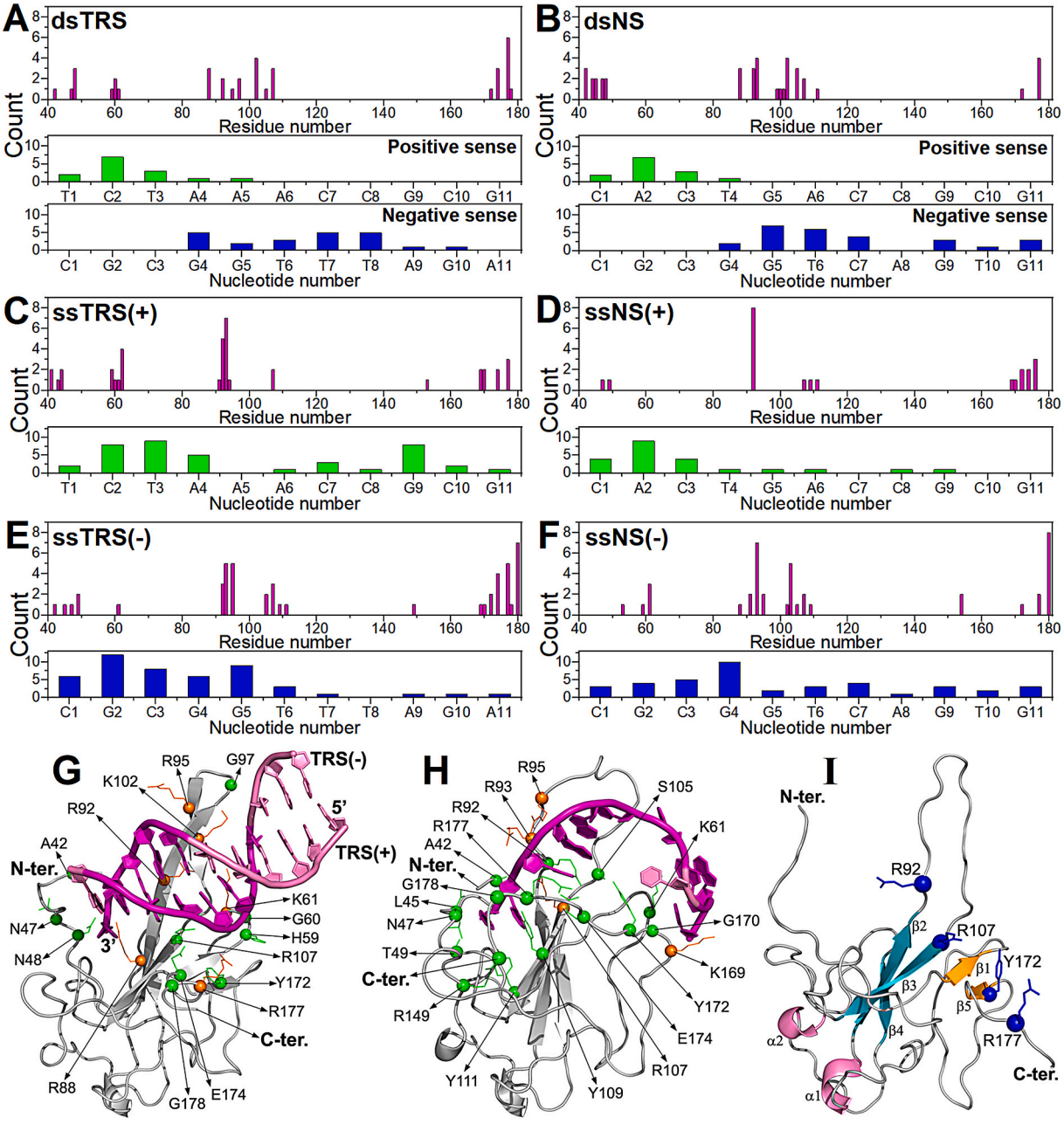

Fig. 5. Protein-DNA hydrogens bonds involved in the structure models of the N-NTD:DNA complex. Count of protein-DNA hydrogen bonds with persistency higher that 10% were with respect to amino acid (magenta bars) and nucleotide residues (green and blue bars) for the interaction of the N-NTD with (A) dsTRS, (B) dsNS, (C) ssTRS(+), (D) ssNS(+), (E) ssTRS(−), and (F) ssNS(−). The residues involved in the persistent protein-DNA hydrogen bonds (green sphere) and salt bridges (orange sphere) are indicated on the representative structural models of the (G) N-NTD:dsTRS and (H) N-NTD:ssTRS(−) complex. The protein is shown as a cartoon model in gray and DNA as a cartoon-ring model in light pink with TTT motif in magenta. (I) The most recurrent residues (R92, R107, Y172, and R177) involved in protein-DNA hydrogen bonds are indicated as blue spheres on the N-NTD structure. The protein is shown as a cartoon model with the β-sheet I (β2/β3/β4), β-sheet II (β1/β5), and α-helices (α1 and α2) colored in cyan, orange, and magenta, respectively.

Surface Area (PBSA). This method enables to evaluate and discriminate the main protein residues and nucleotides that contribute to $\Delta G_b{}^t$. The most significant contributions come from the charged residues (Fig. 6A to F) distributed throughout the protein (Fig. 6G and H). Positively charged arginine and lysine residues contribute favorably, while negatively charged aspartate and glutamate residues contribute unfavorably. It is important to consider the unfavorable contributions because they may be responsible for the DNA/RNA duplex melting activity of N-NTD.

It is also important to analyze the contributions to $\Delta G_b{}^t$ coming from the nucleic acid. For dsDNA, the contribution is mostly favorable for the positive-sense strand. For the negative-sense strand, the contribution is favorable at the 5′ and 3′ termini, but for positions 4 to 8 it varies according to the nucleotide sequence, being less favorable or near-zero for dsNS and unfavorable for dsTRS. For dsTRS, the stretch of nucleotides at positions 4 to 8, which contains the specific TTT motif, is hydrogen bonded to the β-sheet II (β1/β5) (Fig. 5A and G). The same is not observed for dsNS, because it is tilted ~25° away from β-sheet II (β1/β5) (Fig. 3I). In this protein region, we observed charged residues that contribute unfavorably to $\Delta G_b{}^t$, such as E174, which was identified in our previous study with dsRNA [14].

For ssDNAs, both positive and negative-sense strands contributed similarly to $\Delta G_b{}^t$, being slightly more favorable for the positive-sense

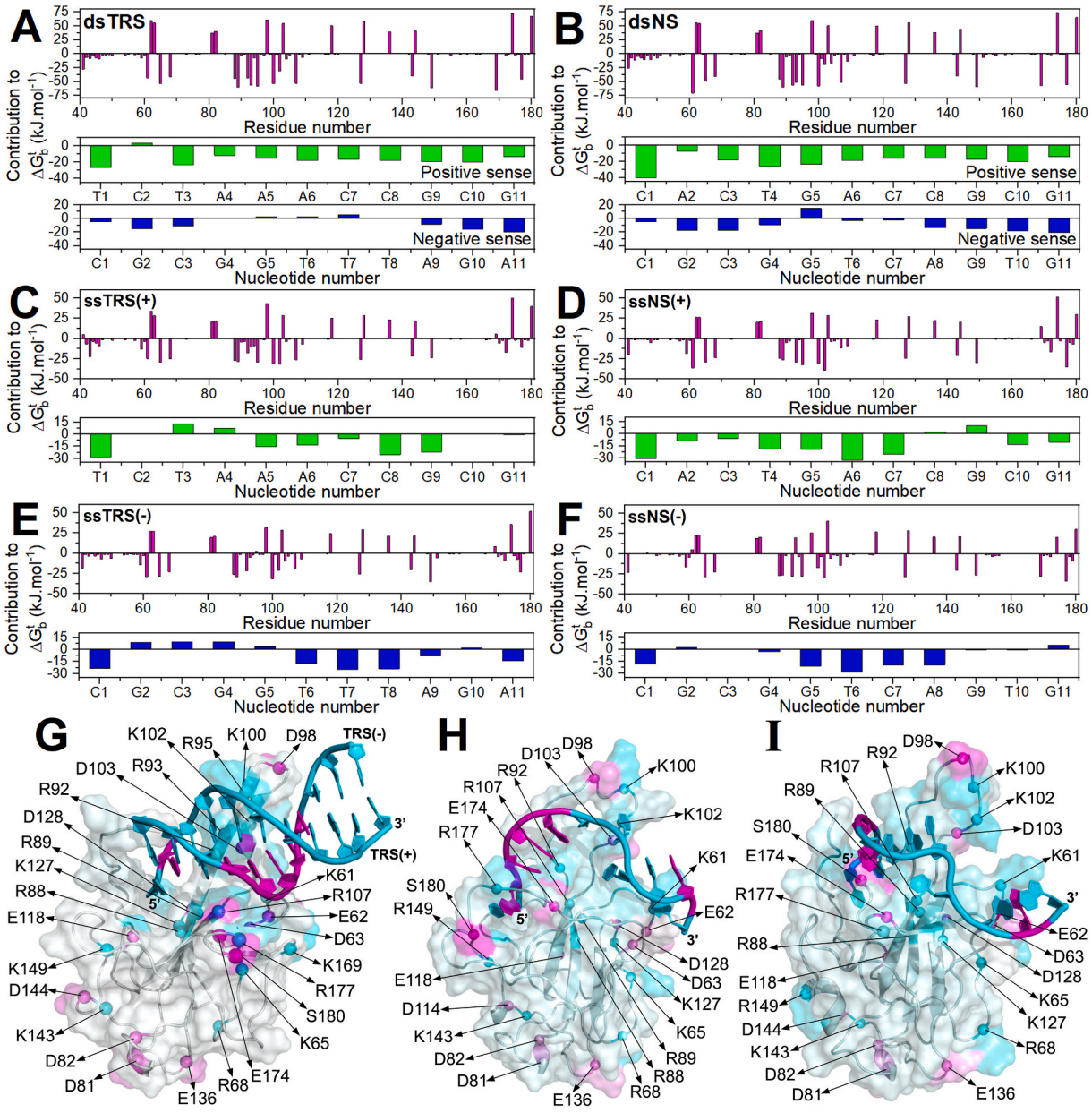

Fig. 6. Energy contribution to Gibbs free energy change ($\Delta G_b{}^t$) of N-NTD:DNA binding. Energy contribution of the amino acid (magenta bars) and nucleotide (green and blue bars) residues to $\Delta G_b{}^t$ (in kJ·mol^{-1}) for the interaction of the N-NTD with (A) dsTRS, (B) dsNS, (C) ssTRS(+), (D) ssNS(+), (E) ssTRS(−), and (F) ssNS(−). Charged residues with most significant energy contribution to $\Delta G_b{}^t$ are indicated as spheres on the representative structural models of the complexes of N-NTD with (G) dsTRS, (H) ssTRS(−), and (I) ssTRS(+). The protein is shown as a cartoon model and translucid surface in gray with the favorable and unfavorable energy contribution of the residues colored in cyan and magenta gradient, respectively. The DNA molecule is displayed as a cartoon-ring model with the favorable and unfavorable energy contribution of the nucleotides colored in cyan and magenta, respectively.

strands. Interacting as a duplex, the positive-sense strands (TRS and NS) showed favorable contribution all along its sequence (Fig. 6A and B); however, when interacting as a ssDNA, the contribution is less favorable (Fig. 6C and D). For the negative-sense strands (TRS and NS), the nucleotides at positions 4 to 8 contribute unfavorably when interacting as a duplex DNA (Fig. 6A and B), while favorably as a single strand (Fig. 6E and F). This is remarkable and agrees with the hypothesis that these unfavorable contributions (TTT motif and β-sheet II) may play a key role in N-NTD melting activity. The contrasting behavior for each strand in the single or double-stranded DNA may be attributed to the fact that the dsDNA has a better-defined positioning when compared to the ssDNA, due to the presence of secondary structure in the duplex. The difference of ~25° in orientation of dsTRS and dsNS (Fig. 3I) may serve as an indicative of specificity of the TRS, contributing to the positioning β-sheet II, E174 for instance, close to TTT motif. This difference in positioning seems to be more relevant for the specificity than the difference in binding affinities between dsTRS and dsNS.

3.5. N-NTD-SR phase separation is triggered by the SR motif and modulated by TRS binding

We investigated the ability of N-NTD and N-NTD-SR to form biomolecular condensates by liquid-liquid phase separation (LLPS) as a consequence of their interaction with specific or non-specific nucleic acids. Indeed, full-length N has been shown to form liquid droplets induced by RNA interaction [20–23]. Interestingly, N from several coronaviruses is predicted to undergo LLPS by the catGRANULE algorithm (Fig. S12A), while N-NTD-SR exhibits the highest LLPS propensity,

most strikingly the SR-rich motif (Fig. S12B). These characteristics prompted us to investigate whether the N-terminal domain contributes to the nucleic acid-driven LLPS of N, the role played by the SR-rich in phase separation, and the influence of nucleic acid binding specificity and stoichiometry on LLPS modulation. Notably, in the simultaneous presence of 10% PEG-4000 and a long non-specific RNA, we observed numerous spherical micron-sized condensates for N-NTD-SR only (Fig. S13B). The same was not observed for N-NTD (Fig. S13A), suggesting that the SR-rich region is important for inducing RNA-driven LLPS under macromolecular crowding. Neither N-NTD nor N-NTD-SR alone formed condensates in aqueous buffer (20 mM Tris-HCl pH 7.5, 30 mM NaCl) even with 10% PEG-4000. We also did not observe condensates in the presence of the long non-specific RNA without the crowding agent (Fig. S13). The dependence of PEG-4000 for N-NTD-SR:RNA condensation was confirmed quantitatively by turbidity measurements (Fig. S13C and D). To further dissect the impact of RNA concentration in promoting N-NTD-SR LLPS, we quantified the number of condensates per 100 μm^2 area and their respective sizes (in μm^2) at five protein:RNA molar ratios. We observed a progressive increase in the number of liquid droplets visible by microscopy accompanied by an increase in turbidity upon the increase in RNA concentration, starting at 2:1 N-NTD-SR:RNA molar ratio (Fig. S14A and C). Specifically, an excess of RNA (1:2 N-NTD-SR:RNA) resulted in 305 ± 7 condensates in a 100 μm^2 area (*top*, Fig. S14B). This number was only 1.5 times higher than that observed at critical concentration for RNA-driven N-NTD-SR LLPS, *i.e.* at the 2:1 N-NTD-SR:RNA stoichiometry. The 1:2 N-NTD-SR:RNA ratio showed the highest dispersion on condensates size, consonant to large condensates formed by fusion (Fig. S14B, bottom graph).

Since nucleic acid structure and sequence can finely tune the phase behavior, we followed this phenomenon in the presence of ssTRS and dsTRS DNAs (Fig. 7). dsTRS (Fig. 7A and D) triggered significant N-NTD-SR LLPS at equimolar concentration (1:1 N-NTD-SR:dsTRS), whereas at 1:2 M ratio, the condensates dissolved, a characteristic behavior of reentrant liquid condensation as reported for RNA-binding proteins containing disordered regions rich in arginine [56,57]. The appearance of condensates is coupled with an increase in turbidity (after 30 min

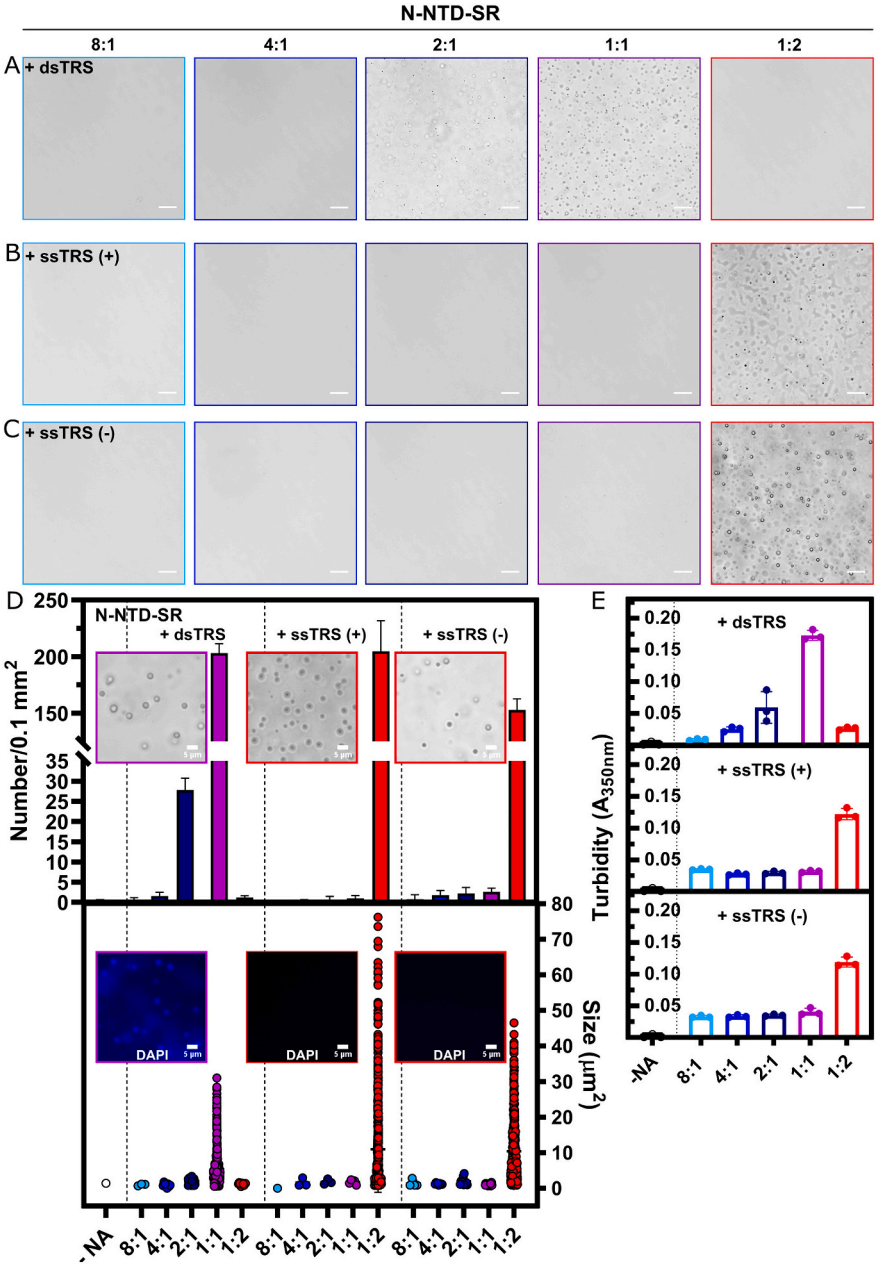

Fig. 7. The high-affinity duplex TRS triggers N-NTD-SR phase separation at a lower protein:DNA stoichiometry. Representative phase contrast micrographs of 20 μM N-NTD-SR in presence of dsTRS (A); ssTRS(+) (B) or ssTRS (−) (C) at the following protein:DNA stoichiometries: 8:1 (light blue); 4:1 (marine blue); 2:1 (dark blue); 1:1 (purple); 1:2 (red), respectively, in 20 mM Tris-HCl buffer (pH 7.5) with 30 mM NaCl. (D) Top graph: Mean number of condensates ± S.D. per 100 μm^2 area (n = 5 images). Top insets: Phase contrast microscopy in the presence of DAPI for conditions with the highest number of liquid droplets (1:1 for the dsTRS; 1:2 for ssTRS(+), and 1:2 for ssTRS(−)). Bottom graph: Scatter plot from the size of condensates (represented as area in μm^2) obtained from micrographs analysis. The following number of condensates were measured: -NA (1); for dsTRS at 8:1 (N = 3); 4:1 (N = 8); 2:1 (N = 132); 1:1 (N = 1014) and 1:2 (N = 11); for ssTRS(+) at 8:1 (N = 1); 4:1 (N = 3); 2:1 (N = 3); 1:1 (N = 5) and 1:2 (N = 1019); for ssTRS(−) at 8:1 (N = 4); 4:1 (N = 9); 2:1 (N = 11); 1:1 (N = 21) and 1:2 (N = 759). Bottom inset: corresponding DAPI emission from the top graph insets images (DAPI stains condensates in presence of dsTRS and no fluorescence were observed for the ssTRSs). All conditions contained 10% (w/v) PEG-4000. Scale bar, 20 μm apart from insets (5 μm). (E) Phase separation of 20 μM N-NTD-SR in 20 mM Tris-HCl buffer (pH 7.5) with 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 specific DNA oligonucleotides dsTRS (top graph), ssTRS(+) (middle graph), and ssTRS(−) (bottom graph). The protein:DNA stoichiometries are the same as in (A).

incubation) at 1:2 stoichiometry, followed by a steep decrease in the presence of molar excess of DNA, confirming the microscopy data (Fig. 7E). We also investigated the effect of the single-stranded counterparts of the duplex TRS. Incubation of N-NTD-SR with either ssTRS (+) (Fig. 7B) or ssTRS(−) (Fig. 7C) only resulted in the appearance of numerous condensates observed by microscopy as well as a significant increase in turbidity at 1:2 protein:DNA stoichiometry (Fig. 7E). Additionally, condensates formed by N-NTD-SR:dsTRS displayed a smaller and more homogenous size as evidenced by their narrower size distribution in comparison to the ssTRSs (at 1:1 with dsTRS 5.4 ± 4.4 μm^2; at 1:2 with ssTRS(+) 11 ± 12.1 μm^2 and 1:2 with ssTRS(−) 10.4 ± 8.2 μm^2). To confirm that the base-paired dsTRS was demixed together with N-NTD-SR, we used the probe DAPI, which shows about 20 times higher fluorescence quantum yield when bound to dsDNA [58]. Indeed, just N-NTD-SR:dsTRS samples showed the extrinsic blue fluorescence emission (Fig. 7D, insets). Fluorescence micrographs acquired with the same settings were obtained for N-NTD-SR in the presence of the long non-specific RNA (Fig. S14B, insets) or with addition of ssTRS(+) or ssTRS (−) (Fig. 7D, insets) and did not show DAPI fluorescence. To unveil the nature of the interactions that drive N-NTD-SR:dsTRS condensation at equimolar stoichiometry, we changed the solution composition (Fig. S15). The addition of a high concentration of 1,6-hexanediol (10% (w/v) HD) decreased the number of droplets by about 30% (203.2 ± 8.2 droplets in absence of HD *versus* 146.8 ± 12.1 droplets with 10% HD) (Fig. S15). This aliphatic alcohol is known to disrupt the weak hydrophobic contacts involved in the liquid condensates, whereas solid-like aggregates [59] and lipid membranes [60] are resistant to HD treatment. Proteins such as hBex3 [54], PrP$^{90–231}$ [53], and RNA polymerase [61], to name a few scaffolds of LLPS, have phase separation abrogated by 10% HD, thus, electrostatic contacts should play a dominant role for N-NTD-SR:DNA LLPS. Indeed, upon incubation with 300 mM NaCl, N-NTD-SR:dsTRS LLPS was totally disassembled (Fig. S15A). Interestingly, in buffer containing 20 mM sodium acetate at pH 5.5, droplet formation was higher, and condensates had marked circular morphology (Fig. S15). Specifically, the droplet number increased by 32% compared to pH 7.5 (203.2 ± 8.2 droplets at pH 7.5 *versus* 298.6 ± 16.8 droplets at pH 5.5), suggesting that acidic pH induces nucleic acid-driven N-NTD-SR LLPS.

To investigate the role of sequence specificity in promoting the N-NTD-SR LLPS, we followed the condensation process using the NS DNAs (Fig. S16). In agreement with dsTRS, the relevant condition for LLPS was at 1:1 protein:DNA stoichiometry. In addition, excess of dsNS dissolved the condensates (127.2 ± 13.9 condensates/μm^2 at 1:1 *versus* 18.4 ± 3.6 condensates/μm^2 at 1:2). This behavior was confirmed by turbidity measurements (Fig. S16E). However, condensates were not homogeneously spherical, as the ones formed with specific dsTRS, and most of them wetted the coverslip surface. Curiously, when observing the entire cover of the glass slide from 1:1 N-NTD-SR:dsNS samples, we observed a few crystals. Since all images were obtained after 30 min incubation, we sought to understand whether crystal formation would be enhanced with prolonged incubation. After 2 h of incubation, we observed crystals presenting DAPI staining (Fig. S16D, inset). Condensates are supersaturated and estimated to be 10 to 300 times enriched in macromolecules compared to the diffuse phase [62], thus crystallization is thermodynamically favored. The buffer with addition of 10% PEG-4000 did not show any artifact (Fig. S13B, top right image), as well as isolated oligonucleotides used in the experiments were excluded to give rise to condensates as shown for dsNS (Fig. S13B, bottom right). The N-NTD-SR:dsNS transition from liquid-like condensates to crystals suggests that the structural properties of the separated phase are different from the N-NTD-SR:dsTRS, once again highlighting the specificity of interaction. Since both dsTRS and dsNS have secondary structure, the higher affinity of N-NTD-SR toward dsTRS might play a role in maintaining the liquidity and consequent dynamics of the protein:nucleic acid condensates. Additionally, we did not observe either condensate formation or increase in turbidity in the presence of the ssNS(+) or ssNS(−) at the five different molar ratios examined (Fig. S16B, C, and E). Therefore, an indicative of sequence specificity may play a pivotal role for DNA-driven N-NTD-SR LLPS since only ssTRS(+) and ssTRS(−) were able to trigger LLPS and not the ssNSs.

4. Discussion

The nucleocapsid N protein is well characterized for its promiscuous interaction with RNA, which is essential to understand two important biological processes for the viral cycle: (i) the assembly of the helical ribonucleoprotein (RNP) complex, and (ii) the discontinuous transcriptional mechanism. N's dsRNA melting activity enables template switch during discontinuous transcription [7,10]. In addition, its ability to phase separate creates a membraneless compartment (liquid-like condensates) that regulates transcription and replication. Previous studies have shown that N binds to single and double-stranded DNA as RNA mimetics [29,30]. Takeda and cols. (2008) and Zhou and cols. (2020) characterized the DNA interaction with the C-terminal domain of N (N-CTD) from SARS-CoV [31] and SARS-CoV-2 [32], respectively. Here, we describe at molecular detail the binding and activity (melting and formation of liquid condensates) of N-NTD and N-NTD-SR toward specific and non-specific DNAs, which are important to understand key interactions with nucleic acids since we showed that N-NTD-(SR) has similar melting activity toward dsRNA and dsDNA (TRS). This observation raised the question of the specificity/promiscuity of N-proteins of coronaviruses, giving structural insights of how a promiscuous protein can exert specific functions. We showed that N-NTD is most likely the domain responsible for the specificity toward TRS. For the melting activity, N-NTD and N-NTD-SR are very similar. This is different for MHV [7], used as a model to study betacoronaviruses. The presence of the SR motif led to a subtle increase in TRS and NS binding affinities and did not change the thermodynamic pattern observed for the binding of N-NTD and TRS. Moreover, the SR motif was essential for LLPS, as previously showed [21], and the formation of liquid condensates was specific for TRS.

Structural and binding studies revealed that solvent-exposed charged residues and electrostatic interactions are the main driving forces for N: nucleic acid complex formation [10,15,29,31,32]. In line with that, we observed that formation of the N-NTD:DNA complex is NaCl and inorganic phosphate dependent (Fig. 2B and C). In addition, dissociation constant increases from nanomolar in 20 mM Bis-Tris buffer (no salt and pH 6.5, Table S7) to the micromolar range in 20 mM sodium phosphate buffer (pH 6.5, Table S8) containing 50 mM NaCl, while maintaining melting activity (Figs. 1 and S4). This is in agreements with kinetic simulations of the melting activity that suggests that the protein is active so long $K_d < 10^{-1}$ M^{-1} [14]. It is worth mentioning that the low micromolar K_d values are similar to those described for RNA binding to N-NTD of SARS coronaviruses [15,63] and Mouse Hepatitis Virus (MHV) [7], a virus closely related to SARS-CoVs. The MD simulations, based on the experimental binding data obtained by NMR, revealed that electrostatic interactions constitute the main energetic contribution for N-NTD:DNA complex stabilization, which happens through the formation of hydrogen bonds between opposite charged residues, *i.e.* salt bridges. Despite the positive electrostatic potential on the binding interface [14], MD simulations also showed the importance of unfavorable contributions from negatively charged residues to $\Delta G_b{}^t$. Accordingly, Dinesh and cols. (2020) showed that R92E and R107E mutations, located at the binding cleft, lead to a decrease in N-NTD-RNA binding affinity, while E174R promotes an increase [15]. We also observed from PBSA analysis that short and long-range contributions of charged residues to $\Delta G_b{}^t$ play a key role in protein-DNA binding affinity and perhaps in dsDNA melting activity.

We provide the first evidence that N-NTD and N-NTD-SR are active upon binding to DNA. Both constructs show dsDNA melting activity, and N-NTD-SR forms liquid condensates in the presence of DNA in a crowded physiological buffer. It is noteworthy that these activities were reported

for RNA only [7,10,20,21,64], and that the ability to melt dsRNA and to form condensates in the presence of RNA are important biological functions of N to spatiotemporally regulate the transcription and replication machinery. We hypothesize that, despite the structural differences between A-dsRNA and B-dsDNA, the similar positioning of DNA and RNA in the binding cleft (between the finger and the palm) is the determining factor for the activities of N-NTD and N-NTD-SR, making the interaction with DNA also a trigger for LLPS. This proposal is based on the similarity between the CSP profiles obtained for N-NTD/N-NTD-SR when interacting with DNAs (Figs. 3, S6, and S7) and RNAs [15], both single and double-stranded. As a consequence, the structural models for N-NTD:DNA (Figs. 3G, H, and S11) and N-NTD:RNA complexes [15] are remarkably similar in position and orientation, especially for the double-stranded nucleic acids. We speculate that the conformational space explored by the single-stranded nucleic acid dynamics might compensate for the occupancy of the double helix in the binding cleft. Taking this interpretation together with the intrinsic protein dynamics, especially at the binding site where a tweezer-like motion between the finger and α2/β5 loop takes place [14], we speculate that the entropically driven ssTRS interaction (Fig. 2D) may be due to the conformational freedom of both protein and ssDNAs. In contrast, the better-defined position and orientation of dsTRS in its enthalpically driven interaction (Fig. 2D) would have the protein conformational dynamics as a major contribution since DNA duplex destabilization is an entropically driven process. The PCA analysis of the MD simulations showed that the conformational space probed by the ssTRS-bound N-NTD is significantly higher than by the dsTRS-bound N-NTD (Fig. S17). Our results suggest that the dsDNA better-defined positioning impacts both the formation of liquid droplets and the melting activity. For LLPS, the presence of nucleic acid secondary structure (dsDNA/dsRNA) favors adequate positioning and perhaps this is linked to the fact that dsTRS induces LLPS at lower DNA concentration when compared to ssTRS (2:1 for N-NTD-SR:dsTRS and 1:2 for N-NTD-SR:ssTRSs). For the melting activity, the defined positioning is also important. We observed dsDNA melting activity at low and high salt and phosphate concentrations (Figs. 1 and S4), for which we showed changes in affinity ranging from nano- to micromolar, suggesting that it is not the difference in affinity that leads to the melting activity, but a property of the interaction itself. Protein dynamics and the defined positioning of dsDNA/dsRNA may be the key for this N protein biological function.

Unlike homeodomains found in eukaryotic transcription factors that possess well-characterized sequence-specific DNA binding [65], CoV N and its CTD are described as non-specific nucleic acid-binding proteins as they bind to RNA and DNA, both single and double-stranded [10,29,63]. This characteristic corroborates the multifunctional role played by N during the viral cycle [11]. Our results for N-NTD/N-NTD-SR binding to DNA reinforce the importance of non-specific charged interactions, such as those observed for RNA [7,10,66]. In general, we noticed modest differences comparing K_d values for TRS and NS DNAs, as well as for ssDNAs and dsDNA, which might suggest a modest sequence-specific recognition of TRS with respect to NS and an affinity preference for ssDNA(−) over ssDNA(+) and dsDNA (Tables S7 and S8). However, considering the difference between specific and non-specific dsDNA binding for homeodomains, there are a ratio of hundreds of times of K_d values [67] that defeats our suggestion of sequence-specific recognition by N-NTD or N-NTD-SR. According to these observations, we cannot explain the specificity based on differences in the affinities, thus we propose the existence of the following indicatives of sequence-specific recognition for TRS: (i) the unique encounter of unfavorable energetic contributions to $\Delta G_b{}^t$ along with protein-dsTRS hydrogen bonds between the TTT motif and β-sheet II that might trigger dsRNA/dsDNA melting; (ii) a difference of ~25° in the orientation between dsTRS and dsNS, putting the TTT motif closer to β-sheet II; (iii) a well-defined pattern for the thermodynamic profile (ΔH and $\Delta S > 0$ for ssTRSs, and ΔH and $\Delta S < 0$ for dsTRS, Fig. 2D) for TRS interaction with N-NTD/N-NTD-SR; and (iv) the preference for TRS over NS DNAs in the formation of liquid condensates. We observed LLPS for ssTRSs but not for ssNS, which might be explained by the difference in affinity to ssNSs (Tables S7 and S8), and crystalline aggregates for dsNS, despite the occurrence of LLPS for both dsTRS and dsNS.

In conclusion, our results contribute to a detailed understanding of the molecular details of binding and activities of the SARS-CoV-2 N protein toward nucleic acids. We described the interaction of N-NTD/N-NTD-SR with single and double-stranded DNA, as well as its dsDNA melting activity and DNA-induced LLPS. In agreement with RNA studies, we characterized the importance of electrostatic interactions for N-NTD (-SR):DNA complex formation and stabilization. Our study also shed light on indicatives of specificity in a protein that is widely described as nonspecific. We suggest that interaction of dsDNA/dsRNA in the binding cleft presents a defined position and orientation of the double helix. For TRS, the melting activity of N-NTD/N-NTD-SR does not seem to be related to the differences in binding affinity to DNA (or RNA), but to differences in positioning of the TTT(UUU) motif in the duplex close to β-sheet II (Fig. 3I). It is also important to further explore the fact that the melting activity was similar for A-dsRNA and B-dsDNA, suggesting that the specificity may depend more on the nucleotide sequence than on the structure of the duplex nucleic acid. The activity toward DNA may serve as a starting point for the design of inhibitors and/or aptamers against a therapeutic target essential for the virus infectivity [7]. In addition, it opens a new avenue of investigation arguing if DNA is indeed a cellular target important for the virus cycle.

Funding

Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ, Brazil: Grant 255.940/2020, 202.279/2018, 239.229/2018, 210.361/2015, and 204.432/2014. Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Brazil: 309564/2017-4 and 439306/2018-3.

CRediT authorship contribution statement

Icaro Putinhon Caruso: Conceptualization, Investigation, Formal analysis, Supervision, Writing - Original Draft, Writing - Review & Editing; **Vitor dos Santos Almeida:** Investigation; **Mariana Juliani do Amaral:** Investigation, Writing - Original Draft; **Guilherme Caldas de Andrade:** Investigation; **Gabriela Rocha de Araújo:** Investigation; **Talita Stelling de Araújo:** Investigation; **Jéssica Moreira de Azevedo:** Investigation; **Glauce Moreno Barbosa:** Investigation; **Leonardo Bartkevihi:** Investigation; **Peter Reis Bezerra:** Investigation; **Katia Maria dos Santos Cabral:** Investigation; **Isabella Otênio de Lourenço:** Investigation; **Clara L. F. Malizia-Motta:** Investigation; **Aline de Luna Marques:** Investigation; **Nathane Cunha Mebus-Antunes:** Investigation; **Thais Cristtina Neves-Martins:** Investigation; **Jéssica Maróstica de Sá:** Investigation; **Karoline Sanches:** Investigation; **Marcos Caique Santana-Silva:** Investigation; **Ariana Azevedo Vasconcelos:** Investigation; **Marcius da Silva Almeida:** Conceptualization, Formal analysis, Supervision; **Gisele Cardoso de Amorim:** Conceptualization, Formal analysis, Supervision; **Cristiane Dinis Anobom:** Conceptualization, Formal analysis, Supervision; **Andrea T. Da Poian:** Conceptualization, Formal analysis, Supervision; **Francisco Gomes-Neto:** Conceptualization, Formal analysis, Supervision; **Anderson S. Pinheiro:** Conceptualization, Formal analysis, Supervision, Writing - Original Draft; **Fabio C. L. Almeida:** Conceptualization, Formal analysis, Supervision, Project administration, Funding acquisition, Writing - Original Draft, Writing - Review & Editing.

Declaration of competing interest

The authors declare that no conflict of interest exists.

Acknowledgments

The author IPC gratefully acknowledges the financial support by postdoctoral fellowship from FAPERJ and the PROPe UNESP. The authors are grateful for the access to the Santos Dumont supercomputer at the National Laboratory of Scientific Computing (LNCC), Brazil. We also acknowledge the Covid-19 NMR Consortium (https://covid19-nmr.de/) for providing an excellent environment for scientific discussions. This work is a collaborative effort of the Rio BioNMR Network (https://www.rmnrio.org/). This study made use of NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource (BTRR), which is supported by NIH grant P41GM111135 (NIGMS).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2022.01.121.

References

[1] T. Singhal, A review of coronavirus Disease-2019 (COVID-19), indian, J. Pediatr. 87 (2020) 281–286, https://doi.org/10.1007/s12098-020-03263-6.
[2] C.C. Lai, T.P. Shih, W.C. Ko, H.J. Tang, P.R. Hsueh, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges, Int. J. Antimicrob. Agents 55 (2020), 105924, https://doi.org/10.1016/j.ijantimicag.2020.105924.
[3] L.Deng F. Jiang L. Zhang, COVID-19, J. Gen. Intern. Med. 2020 (2019) 1–5, https://doi.org/10.1007/s11606-020-05762-w.
[4] I. Sola, F. Almazán, S. Zúñiga, L. Enjuanes, Continuous and discontinuous RNA synthesis in coronaviruses, Annu. Rev. Virol. 2 (2015) 265–288, https://doi.org/10.1146/annurev-virology-100114-055218.
[5] S. Zúñiga, I. Sola, S. Alonso, L. Enjuanes, Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis, J. Virol. 78 (2004) 980–994, https://doi.org/10.1128/jvi.78.2.980-994.2004.
[6] D. Kim, J.Y. Lee, J.S. Yang, J.W. Kim, V.N. Kim, H. Chang, The architecture of SARS-CoV-2 transcriptome, Cell 181 (2020) 914–921.e10, https://doi.org/10.1016/j.cell.2020.04.011.
[7] N.E. Grossoehme, L. Li, S.C. Keane, P. Liu, C.E. Dann, J.L. Leibowitz, D.P. Giedroc, Coronavirus N protein N-terminal domain (NTD) specifically binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS RNA duplexes, J. Mol. Biol. 394 (2009) 544–557, https://doi.org/10.1016/j.jmb.2009.09.040.
[8] P. V'kovski, M. Gerber, J. Kelly, S. Pfaender, N. Ebert, S.Braga Lagache, C. Simillion, J. Portmann, H. Stalder, V. Gaschen, R. Bruggmann, M.H. Stoffel, M. Heller, R. Dijkman, V. Thiel, Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling, Elife 8 (2019), https://doi.org/10.7554/eLife.42037.
[9] M. Yang, S. He, X. Chen, Z. Huang, Z. Zhou, Z. Zhou, Q. Chen, S. Chen, S. Kang, Structural insight into the SARS-CoV-2 nucleocapsid protein C-terminal domain reveals a novel recognition mechanism for viral transcriptional regulatory sequences, Front. Chem. 8 (2021), https://doi.org/10.3389/fchem.2020.624765.
[10] C.K. Chang, M.H. Hou, C.F. Chang, C.D. Hsiao, T.H. Huang, The SARS coronavirus nucleocapsid protein - forms and functions, Antivir. Res. 103 (2014) 39–50, https://doi.org/10.1016/j.antiviral.2013.12.009.
[11] R. McBride, M. van Zyl, B.C. Fielding, The coronavirus nucleocapsid is a multifunctional protein, Viruses 6 (2014) 2991–3018, https://doi.org/10.3390/v6082991.
[12] S. Zúñiga, J.L.G. Cruz, I. Sola, P.A. Mateos-Gómez, L. Palacio, L. Enjuanes, Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription, J. Virol. 84 (2010) 2169–2175, https://doi.org/10.1128/jvi.02011-09.
[13] S.G. Sawicki, D.L. Sawicki, S.G. Siddell, A contemporary view of coronavirus transcription, J. Virol. 81 (2007) 20–29, https://doi.org/10.1128/jvi.01358-06.
[14] I.P. Caruso, K. Sanches, A.T. Da Poian, A.S. Pinheiro, F.C.L. Almeida, Dynamics of the SARS-CoV-2 nucleoprotein N-terminal domain triggers RNA duplex destabilization, Biophys. J. 120 (2021) P2814–P2827, https://doi.org/10.1016/J.BPJ.2021.06.003.
[15] 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.
[16] J. Wang, J.M. Choi, A.S. Holehouse, H.O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R.V. Pappu, S. Alberti, A. A. Hyman, A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins, Cell 174 (2018) 688–699.e16, https://doi.org/10.1016/j.cell.2018.06.006.
[17] C.P. Brangwynne, C.R. Eckmann, D.S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, A.A. Hyman, Germline P granules are liquid droplets that localize by controlled dissolution/condensation, Science 324 (2009) 1729–1732, https://doi.org/10.1126/science.1172046, 80-.
[18] S. Alberti, A.A. Hyman, Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing, Nat. Rev. Mol. Cell Biol. 22 (2021) 196–213, https://doi.org/10.1038/s41580-020-00326-6.
[19] S. Guseva, S. Milles, M.R. Jensen, N. Salvi, J.P. Kleman, D. Maurin, R.W. H. Ruigrok, M. Blackledge, Measles virus nucleo- and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly, Sci. Adv. 6 (2020), https://doi.org/10.1126/sciadv.aaz7095.
[20] T.M. Perdikari, A.C. Murthy, V.H. Ryan, S. Watters, M.T. Naik, N.L. Fawzi, SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs, EMBO J. (2020) 1–15, https://doi.org/10.15252/embj.2020106478.
[21] A. Savastano, A. Ibáñez de Opakua, M. Rankovic, M. Zweckstetter, Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates, Nat. Commun. 11 (2020), https://doi.org/10.1038/s41467-020-19843-1.
[22] C. Iserman, C.A. Roden, M.A. Boerneke, R.S.G. Sealfon, G.A. McLaughlin, I. Jungreis, E.J. Fritch, Y.J. Hou, J. Ekena, C.A. Weidmann, C.L. Theesfeld, M. Kellis, O.G. Troyanskaya, R.S. Baric, T.P. Sheahan, K.M. Weeks, A.S. Gladfelter, Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid, Mol. Cell 80 (2020) 1078–1091.e6, https://doi.org/10.1016/j.molcel.2020.11.041.
[23] J. Cubuk, J.J. Alston, J.J. Incicco, S. Singh, M.D. Stuchell-Brereton, M.D. Ward, M. I. Zimmerman, N. Vithani, D. Griffith, J.A. Wagoner, G.R. Bowman, K.B. Hall, A. Soranno, A.S. Holehouse, The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA, Nat. Commun. 12 (2021) 1–17, https://doi.org/10.1038/s41467-021-21953-3.
[24] K.A. Timani, Q. Liao, L. Ye, Y. Zeng, J. Liu, Y. Zheng, L. Ye, X. Yang, K. Lingbao, J. Gao, Y. Zhu, Nuclear/nucleolar localization properties of C-terminal nucleocapsid protein of SARS coronavirus, Virus Res. 114 (2005) 23–34, https://doi.org/10.1016/j.virusres.2005.05.007.
[25] R.R.R. Rowland, V. Chauhan, Y. Fang, A. Pekosz, M. Kerrigan, M.D. Burton, Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells, J. Virol. 79 (2005) 11507–11512, https://doi.org/10.1128/jvi.79.17.11507-11512.2005.
[26] W.N. Wulan, D. Heydet, E.J. Walker, M.E. Gahan, R. Ghildyal, Nucleocytoplasmic transport of nucleocapsid proteins of enveloped RNA viruses, Front. Microbiol. 6 (2015), https://doi.org/10.3389/fmicb.2015.00553.
[27] L.A. Caldas, F.A. Carneiro, F.L. Monteiro, I. Augusto, L.M. Higa, K. Miranda, A. Tanuri, W. de Souza, Intracellular host cell membrane remodelling induced by SARS-CoV-2 infection in vitro, Biol. Cell. 113 (2021) 281–293, https://doi.org/10.1111/BOC.202000146.
[28] A. Panagiotopoulos, M. Tseliou, I. Karakasiliotis, D.M. Kotzampasi, V. Daskalakis, N. Kesesidis, G. Notas, C. Lionis, M. Kampa, S. Pirintsos, G. Sourvinos, E. Castanas, P-cymene impairs SARS-CoV-2 and influenza a (H1N1) viral replication: in silico predicted interaction with SARS-CoV-2 nucleocapsid protein and H1N1 nucleoprotein, Pharmacol. Res. Perspect. 9 (2021), e00798, https://doi.org/10.1002/PRP2.798.
[29] T.K. Tang, M.P.J. Wu, S.T. Chen, M.H. Hou, M.H. Hong, F.M. Pan, H.M. Yu, J. H. Chen, C.W. Yao, A.H.J. Wang, Biochemical and immunological studies of nucleocapsid proteins of severe acute respiratory syndrome and 229E human coronaviruses, in: Proteomics, Proteomics, 2005, pp. 925–937, https://doi.org/10.1002/pmic.200401204.
[30] W. Zeng, G. Liu, H. Ma, D. Zhao, Y. Yang, M. Liu, A. Mohammed, C. Zhao, Y. Yang, J. Xie, C. Ding, X. Ma, J. Weng, Y. Gao, H. He, T. Jin, Biochemical characterization of SARS-CoV-2 nucleocapsid protein, Biochem. Biophys. Res. Commun. 527 (2020) 618–623, https://doi.org/10.1016/j.bbrc.2020.04.136.
[31] M. Takeda, C.Ke Chang, T. Ikeya, P. Güntert, Y.Hsiang Chang, Y.Lan Hsu, T. Huang Huang, M. Kainosho, Solution structure of the C-terminal dimerization domain of SARS coronavirus nucleocapsid protein solved by the SAIL-NMR method, J. Mol. Biol. 380 (2008) 608–622, https://doi.org/10.1016/j.jmb.2007.11.093.
[32] R. Zhou, R. Zeng, A. von Brunn, J. Lei, Structural characterization of the C-terminal domain of SARS-CoV-2 nucleocapsid protein, Mol. Biomed. 1 (2020) 1–11, https://doi.org/10.1186/s43556-020-00001-4.
[33] N. Altincekic, S.M. Korn, N.S. Qureshi, M. Dujardin, M. Ninot-Pedrosa, R. Abele, M. J.Abi Saad, C. Alfano, F.C.L. Almeida, I. Alshamleh, G.C. de Amorim, T. K. Anderson, C.D. Anobom, C. Anorma, J.K. Bains, A. Bax, M. Blackledge, J. Blechar, A. Böckmann, L. Brigandat, A. Bula, M. Bütikofer, A.R. Camacho-Zarco, T. Carlomagno, I.P. Caruso, B. Ceylan, A. Chaikuad, F. Chu, L. Cole, M.G. Crosby, V. de Jesus, K. Dhamotharan, I.C. Felli, J. Ferner, Y. Fleischmann, M.-L. Fogeron, N.K. Fourkiotis, C. Fuks, B. Fürtig, A. Gallo, S.L. Gande, J.A. Gerez, D. Ghosh, F. Gomes-Neto, O. Gorbatyuk, S. Guseva, C. Hacker, S. Häfner, B. Hao, B. Hargittay, K. Henzler-Wildman, J.C. Hoch, K.F. Hohmann, M.T. Hutchison, K. Jaudzems, K. Jović, J. Kaderli, G. Kalniņš, I. Kaņepe, R.N. Kirchdoerfer, J. Kirkpatrick, S. Knapp, R. Krishnathas, F. Kutz, S.Zur Lage, R. Lambertz, A. Lang, D. Laurents, L. Lecoq, V. Linhard, F. Löhr, A. Malki, L.M. Bessa, R.W. Martin, T. Matzel, D. Maurin, S.W. McNutt, N.C. Mebus-Antunes, B.H. Meier, N. Meiser, M. Mompeán, E. Monaca, R. Montserret, L.Mariño Perez, C. Moser, C. Muhle-Goll, T.C. Neves-Martins, X. Ni, B. Norton-Baker, R. Pierattelli, L. Pontoriero, Y. Pustovalova, O. Ohlenschläger, J. Orts, A.T.Da Poian, D.J. Pyper, C. Richter, R. Riek, C.M. Rienstra, A. Robertson, A.S. Pinheiro, R. Sabbatella, N. Salvi, K. Saxena, L. Schulte, M. Schiavina, H. Schwalbe, M. Silber, M. da S. Almeida, M. A. Sprague-Piercy, G.A. Spyroulias, S. Sreeramulu, J.-N. Tants, K. Tärs, F. Torres, S. Töws, M.Á. Treviño, S. Trucks, A.C. Tsika, K. Varga, Y. Wang, M.E. Weber, J. E. Weigand, C. Wiedemann, J. Wirmer-Bartoschek, M.A.Wirtz Martin, J. Zehnder, M. Hengesbach, A. Schlundt, Large-scale recombinant production of the SARS-CoV-

2 proteome for high-throughput and structural biology applications, Front. Mol. Biosci. 8 (2021) 89, https://doi.org/10.3389/fmolb.2021.653148.
[34] I.P. Caruso, G.C. Guimarães, V.B. Machado, M.A. Fossey, D. Willbold, F.C. L. Almeida, F.P. Souza, Biophysical and dynamic characterization of fine-tuned binding of the human respiratory syncytial virus M2–1 Core domain to long RNAs, J. Virol. 94 (2020), https://doi.org/10.1128/jvi.01505-20.
[35] J.R. Lakowicz, Instrumentation for fluorescence spectroscopy, in: Princ. Fluoresc. Spectrosc, Springer US, 2006, pp. 27–61, https://doi.org/10.1007/978-0-387-46312-4_2.
[36] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293, https://doi.org/10.1007/BF00197809.
[37] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, E.D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, Proteins Struct. Funct. Genet. 59 (2005) 687–696, https://doi.org/10.1002/prot.20449.
[38] M.W. Maciejewski, A.D. Schuyler, M.R. Gryk, I.I. Moraru, P.R. Romero, E.L. Ulrich, H.R. Eghbalnia, M. Livny, F. Delaglio, J.C. Hoch, NMRbox: a resource for biomolecular NMR computation, Biophys. J. 112 (2017) 1529–1534, https://doi.org/10.1016/j.bpj.2017.03.011.
[39] G.C.P. Van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. Van Dijk, S.J. De Vries, A.M.J.J. Bonvin, The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes, J. Mol. Biol. 428 (2016) 720–725, https://doi.org/10.1016/j.jmb.2015.09.014.
[40] J.P. Linge, M.A. Williams, C.A.E.M. Spronk, A.M.J.J. Bonvin, M. Nilges, Refinement of protein structures in explicit solvent, Proteins Struct. Funct. Genet. 50 (2003) 496–506, https://doi.org/10.1002/prot.10299.
[41] S.J. De Vries, M. Van Dijk, A.M.J.J. Bonvin, The HADDOCK web server for data-driven biomolecular docking, Nat. Protoc. 5 (2010) 883–897, https://doi.org/10.1038/nprot.2010.32.
[42] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.
[43] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696–3713, https://doi.org/10.1021/acs.jctc.5b00255.
[44] M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T.E. Cheatham, P. Jurečka, Refinement of the Cornell et al. Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles, J. Chem. Theory Comput. 7 (2011) 2886–2902, https://doi.org/10.1021/ct200162x.
[45] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935, https://doi.org/10.1063/1.445869.
[46] M.H.M. Olsson, C.R. Søndergaard, M. Rostkowski, J.H. Jensen, PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions, J. Chem. Theory Comput. 7 (2011) 525–537, https://doi.org/10.1021/ct100578z.
[47] A.J. Lemkul, Scripts and Programs, OSF, 2018. https://osf.io/bafn4/. (Accessed 13 June 2020).
[48] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10037–10041, https://doi.org/10.1073/pnas.181342398.
[49] R. Kumari, R. Kumar, A. Lynn, G-mmpbsa -a GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. 54 (2014) 1951–1962, https://doi.org/10.1021/ci500020m.
[50] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J. C. Smith, P.M. Kasson, D.Van Der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (2013) 845–854, https://doi.org/10.1093/bioinformatics/btt055.
[51] X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier, D. Moras, J. Lehn, M. Garrett, U. Koert, D. Meyer, J. Fischer, Peptide Folding: When Simulation Meets Experiment, 1998.
[52] W.L. Delano, The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA, 2002.
[53] 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.
[54] 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.
[55] P.D. Ross, S. Subramanian, Thermodynamics of protein association reactions: forces contributing to stability, Biochemistry 20 (1981) 3096–3102, https://doi.org/10.1021/bi00514a017.
[56] I. Alshareedah, T. Kaur, J. Ngo, H. Seppala, L.A.D. Kounatse, W. Wang, M. M. Moosa, P.R. Banerjee, Interplay between short-range attraction and long-range repulsion controls reentrant liquid condensation of ribonucleoprotein-RNA complexes, J. Am. Chem. Soc. 141 (2019) 14593–14602, https://doi.org/10.1021/jacs.9b03689.
[57] S. Maharana, J. Wang, D.K. Papadopoulos, D. Richter, A. Pozniakovsky, I. Poser, M. Bickle, S. Rizk, J. Guillén-Boixet, T.M. Franzmann, M. Jahnel, L. Marrone, Y. T. Chang, J. Sterneckert, P. Tomancak, A.A. Hyman, S. Alberti, RNA buffers the phase separation behavior of prion-like RNA binding proteins, Science 360 (2018) 918–921, https://doi.org/10.1126/science.aar7366, 80-.
[58] B. Chazotte, Labeling nuclear DNA using DAPI, Cold Spring Harb. Protoc. 6 (2011), https://doi.org/10.1101/pdb.prot5556.
[59] S. Kroschwald, S. Maharana, A. Simon, Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments, Matters 3 (2017), e201702000010, https://doi.org/10.19185/matters.201702000010.
[60] C. Chang, C.E. Cornell, S.L. Keller, 1,6-hexanediol, which is used to disrupt protein-rich liquid droplets in the cell cytoplasm, does not disrupt model cell membranes, Biophys. J. 114 (2018) 259a, https://doi.org/10.1016/j.bpj.2017.11.1505.
[61] A.M. Ladouceur, B.S. Parmar, S. Biedzinski, J. Wall, S.G. Tope, D. Cohn, A. Kim, N. Soubry, R. Reyes-Lamothe, S.C. Weber, Clusters of bacterial RNA polymerase are biomolecular condensates that assemble through liquid-liquid phase separation, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 18540–18549, https://doi.org/10.1073/pnas.2005019117.
[62] K.K. Turoverov, I.M. Kuznetsova, A.V. Fonin, A.L. Darling, B.Y. Zaslavsky, V. N. Uversky, Stochasticity of biological soft matter: emerging concepts in intrinsically disordered proteins and biological phase separation, Trends Biochem. Sci. 44 (2019) 716–728, https://doi.org/10.1016/j.tibs.2019.03.005.
[63] C.-K. Chang, Y.-L. Hsu, Y.-H. Chang, F.-A. Chao, M.-C. Wu, Y.-S. Huang, C.-K. Hu, T.-H. Huang, Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging, J. Virol. 83 (2009) 2255–2264, https://doi.org/10.1128/jvi.02001-08.
[64] C.R. Carlson, J.B. Asfaha, C.M. Ghent, C.J. Howard, N. Hartooni, M. Safari, A. D. Frankel, D.O. Morgan, Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions, Mol. Cell 80 (2020) 1092–1103.e4, https://doi.org/10.1016/j.molcel.2020.11.025.
[65] W.J. Gehring, Y.Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A.F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, K. Wüthrich, Homeodomain-DNA recognition, Cell 78 (1994) 211–223, https://doi.org/10.1016/0092-8674(94)90292-5.
[66] K.A. Spencer, J.A. Hiscox, Characterisation of the RNA binding properties of the coronavirus infectious bronchitis virus nucleocapsid protein amino-terminal region, FEBS Lett. 580 (2006) 5993–5998, https://doi.org/10.1016/j.febslet.2006.09.052.
[67] J. Iwahara, G.M. Clore, Detecting transient intermediates in macromolecular binding by paramagnetic NMR, Nature 440 (2006) 1227–1230, https://doi.org/10.1038/nature04673.