RESEARCH   **Open Access**

Modular characterization of SARS-CoV-2 nucleocapsid protein domain functions in nucleocapsid-like assembly

Yan Wang[1†], Xiaobin Ling[1,2†], Chong Zhang[1], Jian Zou[1], Bingnan Luo[1], Yongbo Luo[1], Xinyu Jia[1], Guowen Jia[1], Minghua Zhang[3], Junchao Hu[1], Ting Liu[1], Yuanfeiyi Wang[1], Kefeng Lu[1], Dan Li[4], Jinbiao Ma[2*], Cong Liu[5,6*] and Zhaoming Su[1*]

Abstract

SARS-CoV-2 and its variants, with the Omicron subvariant XBB currently prevailing the global infections, continue to pose threats on public health worldwide. This non-segmented positive-stranded RNA virus encodes the multifunctional nucleocapsid protein (N) that plays key roles in viral infection, replication, genome packaging and budding. N protein consists of two structural domains, NTD and CTD, and three intrinsically disordered regions (IDRs) including the N_{IDR}, the serine/arginine rich motif (SR_{IDR}), and the C_{IDR}. Previous studies revealed functions of N protein in RNA binding, oligomerization, and liquid–liquid phase separation (LLPS), however, characterizations of individual domains and their dissected contributions to N protein functions remain incomplete. In particular, little is known about N protein assembly that may play essential roles in viral replication and genome packing. Here, we present a modular approach to dissect functional roles of individual domains in SARS-CoV-2 N protein that reveals inhibitory or augmented modulations of protein assembly and LLPS in the presence of viral RNAs. Intriguingly, full-length N protein (N_{FL}) assembles into ring-like architecture whereas the truncated SR_{IDR}-CTD-C_{IDR} ($N_{182-419}$) promotes filamentous assembly. Moreover, LLPS droplets of N_{FL} and $N_{182-419}$ are significantly enlarged in the presence of viral RNAs, and we observed filamentous structures in the $N_{182-419}$ droplets using correlative light and electron microscopy (CLEM), suggesting that the formation of LLPS droplets may promote higher-order assembly of N protein for transcription, replication and packaging. Together this study expands our understanding of the multiple functions of N protein in SARS-CoV-2.

Keywords Nucleocapsid protein, Filamentous assembly, Liquid–liquid phase separation

---

†Yan Wang and Xiaobin Ling contributed equally.

*Correspondence:
Jinbiao Ma
majb@fudan.edu.cn
Cong Liu
liulab@sioc.ac.cn
Zhaoming Su
zsu@scu.edu.cn
[1] The State Key Laboratory of Biotherapy, Frontiers Medical Center of Tianfu Jincheng Laboratory, National Clinical Research Center for Geriatrics and Department of Geriatrics, West China Hospital, Sichuan University, Chengdu 610044, Sichuan, China

[2] State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry and Biophysics, School of Life Sciences, Fudan University, Shanghai 200438, China
[3] College of Polymer Science and Engineering, Sichuan University, Chengdu 610065, Sichuan, China
[4] Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Ministry of Education, Shanghai Jiao Tong University, Shanghai 200030, China
[5] Interdisciplinary Research Center On Biology and Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 201210, China
[6] State Key Laboratory of Bio-Organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 2 of 14

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) continues to evolve as it spreads around the world, with the Omicron subvariant XBB currently dominating the global infections[1]. SARS-CoV-2 is a non-segmented positive-stranded RNA virus with strong infectivity and high lethality [2]. The genome of SARS-CoV-2 is almost 30,000 nucleotides (nts), which encodes 16 non-structural proteins (nsp1-nsp16) and 4 structural proteins (Spike, S; Envelope, E; Membrane, M; Nucleocapsid, N) [3–6]. The N protein is located near the 3' end of the genome and is one of the most abundant viral proteins [7]. The N protein assembles into nucleocapsids to wrap and protect the viral genomes. The structural proteins S, E and M form the envelopes of SARS-CoV-2 virus particles that wrap the nucleocapsids [8, 9].

N is a multifunctional protein that involves in the virus life cycle and cell response to protect the viral genome and regulate the host immune system during virion packaging, as well as participation in the replication and synthesis of genomic RNAs [10–14]. N consists of two structural domains, the N-terminal RNA binding domain (NTD) and the C-terminal dimerization domain (CTD), which are incorporated into three intrinsic disordered regions (IDRs) at the N-terminus of NTD (N_{IDR}), between the NTD and CTD with the serine/arginine rich motif (SR_{IDR}), and at the C-terminus of CTD (C_{IDR}) (Fig. 1a). The sequences of N and the structures of NTD and CTD are highly conserved among SARS-CoV-2 and other β-coronaviruses (CoV), such as SARS-CoV, MERS-CoV and HCoV-OC43 [3, 15]. The NTD has a unique RNA binding pocket, whereas the CTD dimerizes in solution and occasionally forms tetramers in the presence of C_{IDR} [16–18]. Small molecules designed to target NTD and CTD binding sites could inhibit RNA binding and oligomerization of N, thereby decreasing the virus replication [15]. N is also highly immunogenic due to its high abundancy and interference with host immune systems during viral infections [19–23], suggesting N as a potential target for diagnosis and biotherapy development.

Filamentous nucleocapsid-like structures in β-CoV including SARS-CoV and SARS-CoV-2 characterized by cryo-electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET) have been previously reported [24–30], recent cryo-ET analyses of SARS-CoV-2 virions in situ revealed low-resolution globular-shaped ribonucleoprotein (RNP) complexes [31–35]. In the meantime, recent studies on SARS-CoV-2 N protein revealed liquid–liquid phase separation (LLPS) in the presence of SARS-CoV-2 viral RNAs [29, 36–43]. The SR_{IDR} contains a SR-rich motif that may facilitate multimerization of N and transcription-regulating sequence (TRS)-dependent LLPS [44, 45], indicating that LLPS formation could initiate nucleocapsid-like assembly that is essential for transcription, replication and virion packaging.

Here, we present a modular approach to investigate functions of individual domains of SARS-CoV-2 N in protein assembly and LLPS. Recombinant expression
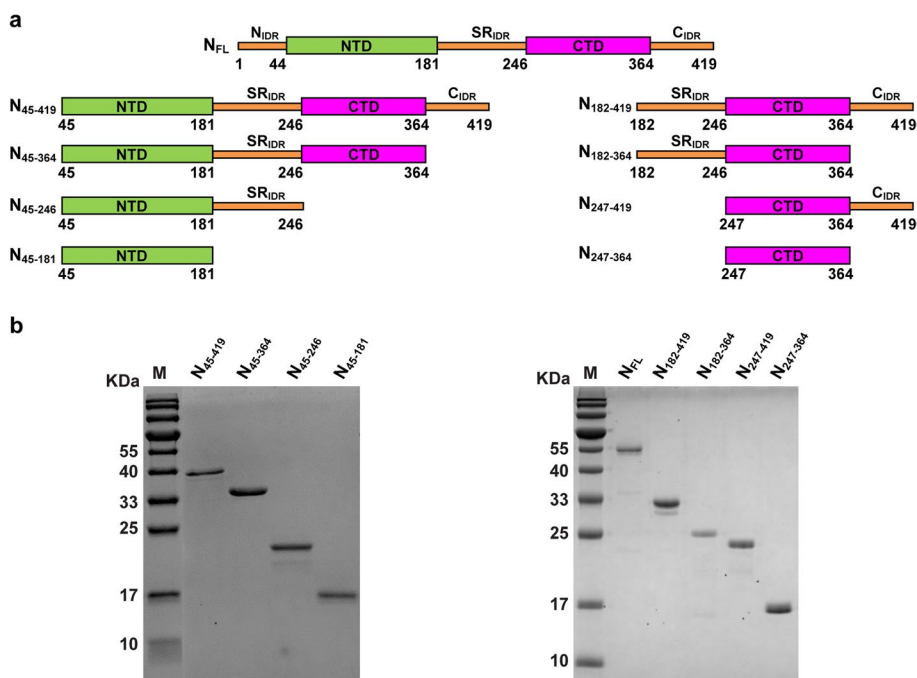

Fig. 1 Modular SARS-CoV-2 N constructs. **a** Cartoon illustration of modular N constructs. **b** SDS-PAGE analysis of purified modular N constructs

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 3 of 14

and purification of the full-length N protein (N_{FL}) from *Escherichia coli* (*E. coli*) revealed ring-like assembly in the presence or absence of cellular RNAs. Intriguingly, the SR_{IDR}-CTD-C_{IDR} ($N_{182-419}$) construct with the truncated N_{IDR}-NTD showed filamentous assembly. Further truncations of either IDR resulted in ring-like assembly, indicating that both SR_{IDR} and C_{IDR} adjacent to CTD are indispensable for filament formation, whereas N_{IDR}-NTD might be inhibitory to filamentous assembly. The higher-order structures, including ring-like and filamentous assembly, were recapitulated in vitro when the N constructs were incubated with previously proposed packaging signals (PS) including 5' UTR (nts 1–478) [46, 47], PS100 (nts 691–789) [48], PS576 (nts 19,786–20,361) [48, 49], PS9 (nts 20,080–21,171 that partially encodes nsp15-16) [50], or PS97 (nts 29,027–29,129) [48, 51] regions of the SARS-CoV-2 genome, whereas 3' UTR could not facilitate protein assembly, validating that these regions likely contain viral assembly and packaging signals. In addition to previous studies that extensively characterized the N_{FL} for LLPS, we further characterized these modular constructs of N for LLPS and observed significantly enlarged droplets containing cylindrical filaments for SR_{IDR}-CTD-C_{IDR} in the presence of SARS-CoV-2 5' UTR, suggesting that formation of LLPS droplets may favor the higher-order assembly of N essential for viral RNA synthesis and virion packaging. Together our results provide modular functions of individual domains of N that further complete our understanding of the multi-functional N protein, which may aid novel diagnosis and antiviral development against SARS-CoV-2.

Results

Modular characterizations of SARS-CoV-2 N constructs in higher-order assembly with cellular and viral RNAs

We designed a modular approach with domain truncations on N-terminus and C-terminus, resulting in eight different modular constructs of N (Fig. 1). Recombinant expression in *E. coli* and purification of the N_{FL} under physiological relevant salt (150 mM NaCl) or high salt (2 M NaCl) conditions resulted in the peak elution volumes (PEVs) of the first peak (p1) as 9.89 mL and 11.06 mL in gel filtration chromatography (Fig. 2a), suggesting possible formations of higher-order assembly. Absorbance measurement showed the existence of cellular nucleic acid under 150 mM NaCl condition, whereas minimal nucleic acid was detected after gel filtration in the presence of 2 M NaCl, which was typically used to remove nonspecifically bound cellular nucleic acids and to induce salt bridge interactions that are frequently observed in protein assembly interfaces [37, 52, 53]. Negative staining revealed heterogeneous ring-like assembly of N under physiological relevant salt condition with the

diameter ranging from 20–35 nm (Fig. 2a-b, black p1), which led to the low-resolution three-dimensional (3D) reconstruction of a representative ring-like structure (Fig. 2c).

We generated eight additional truncated constructs consisted of different modular domains to further characterize their functions in higher-order assembly (Fig. 1). Although all truncated constructs could form higher-order structures to some extent under 2 M NaCl condition (Figs. 2 and 3), indicative of the universal effect of salt on protein assembly [54, 55], only those constructs that contained CTD but not NTD formed higher-order assembly in the presence of RNAs under physiological relevant salt condition (Fig. 4), consistent with previous studies that CTD domain of N was prone to oligomerization [30, 55–57]. The ring-like assembly no longer existed when an additional NTD domain were included, yielding constructs N_{45-419} and N_{45-364}, indicating that NTD played an inhibitory role in higher-order N assembly (Fig. 4d-e). The fact that inclusion of N_{IDR} domain in N_{FL} facilitated ring-like assembly suggested the presence of N_{IDR} domain reduced the inhibitory effect of NTD in higher-order N assembly (Fig. 2a).

$N_{182-419}$ forms filamentous assembly

Previous studies observed filamentous nucleocapsid-like assembly of N in β-CoV including SARS-CoV and SARS-CoV-2 [24–30], whereas recent cryo-ET analyses of SARS-CoV-2 virions in situ revealed low-resolution globular-shaped viral RNPs [31–35], suggesting that the filamentous form of RNP might only exist in the process of RNA synthesis and the beginning of genome packaging to protect the viral genome. Intriguingly, we observed flexible filamentous assembly of $N_{182-419}$ under both high and physiological relevant salt conditions with helix width ranging from 10–30 nm that generally matched the diameter of the previous ring-like assembly (Fig. 2d-e), yielding a low-resolution 3D reconstruction of the representative filament with an averaged diameter of 20 nm (Fig. 2f). The fact that only $N_{182-419}$ construct assembled into filaments suggests that the N_{IDR} and NTD domains might have inhibitory effects, whereas SR_{IDR} and C_{IDR} have displayed augmented effects on filamentous assembly.

SARS-CoV-2 packaging signals induce filamentous assembly in vitro

Previous studies demonstrated that the SARS-CoV-2 5' UTR, PS100, PS576, PS9 and PS97 regions on the viral genome could facilitate N protein assembly that are essential for viral ribonucleoprotein formation, virion packaging and viral replication [46, 50, 58]. In order to evaluate the impacts of viral RNAs on N protein
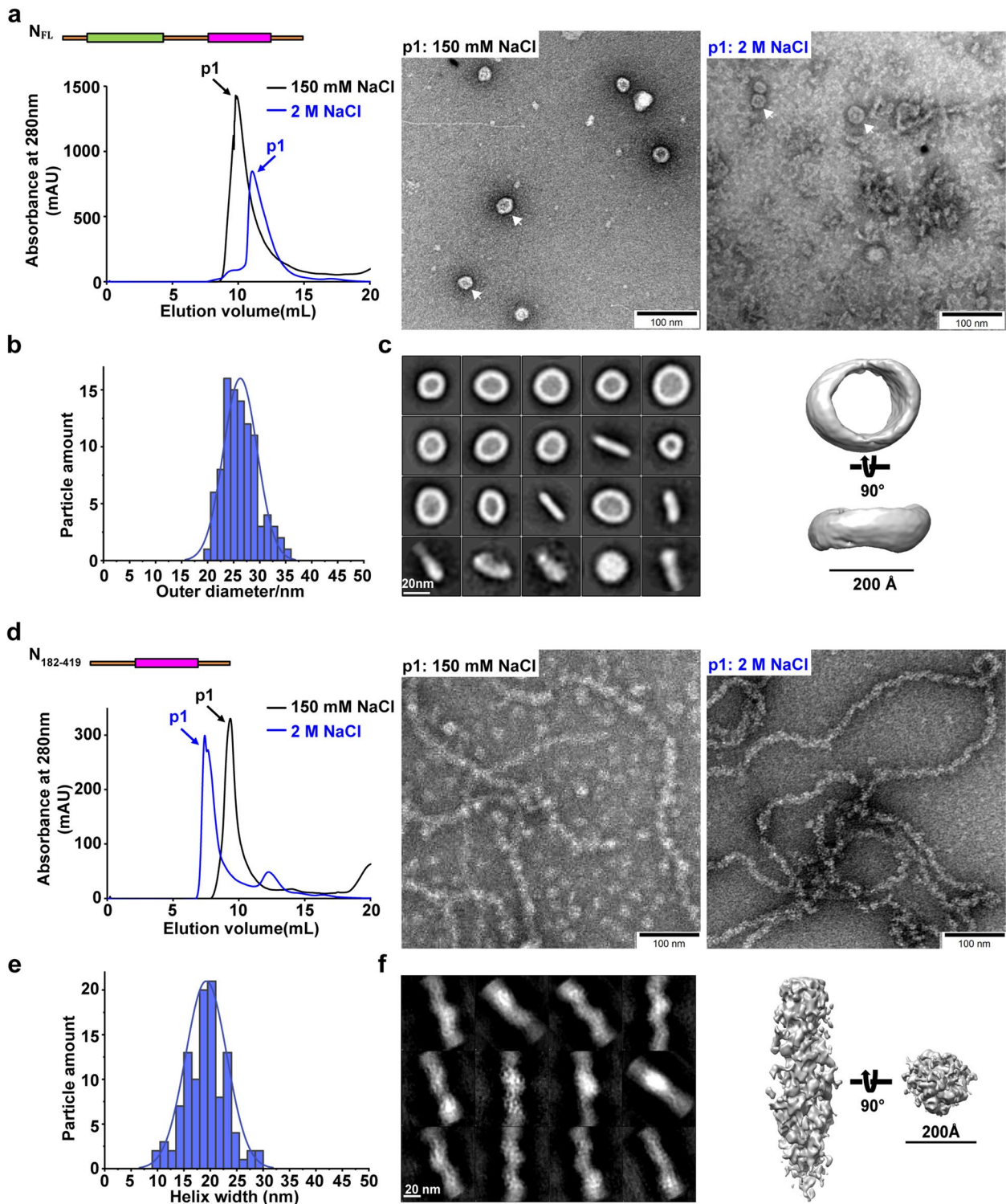
Fig. 2 SARS-CoV-2 N_{FL} and $N_{182-419}$ protein assembly under physiological relevant conditions. **a** The PEVs of SARS-CoV-2 N_{FL} protein under physiological relevant salt condition with retained cellular RNAs (black) and high salt condition with removed cellular RNAs (blue) by gel filtration chromatography and the corresponding representative negative stain micrographs. Scale bars, 100 nm. **b** Diameter distribution of N_{FL} protein ring-like particles. **c** 2D averages and 3D reconstruction of N_{FL} negative stain data. Scale bar, 20 nm. **d** The PEVs of SARS-CoV-2 $N_{182-419}$ protein under physiological relevant salt condition with retained cellular RNAs (black) and high salt condition with removed cellular RNAs (blue) by gel filtration chromatography and the corresponding representative negative stain micrographs. Scale bars, 100 nm. **e** Diameter distribution of $N_{182-419}$ filamentous nucleocapsid-like particles. **f** 2D averages and 3D reconstruction of $N_{182-419}$ negative stain data. Scale bar, 20 nm
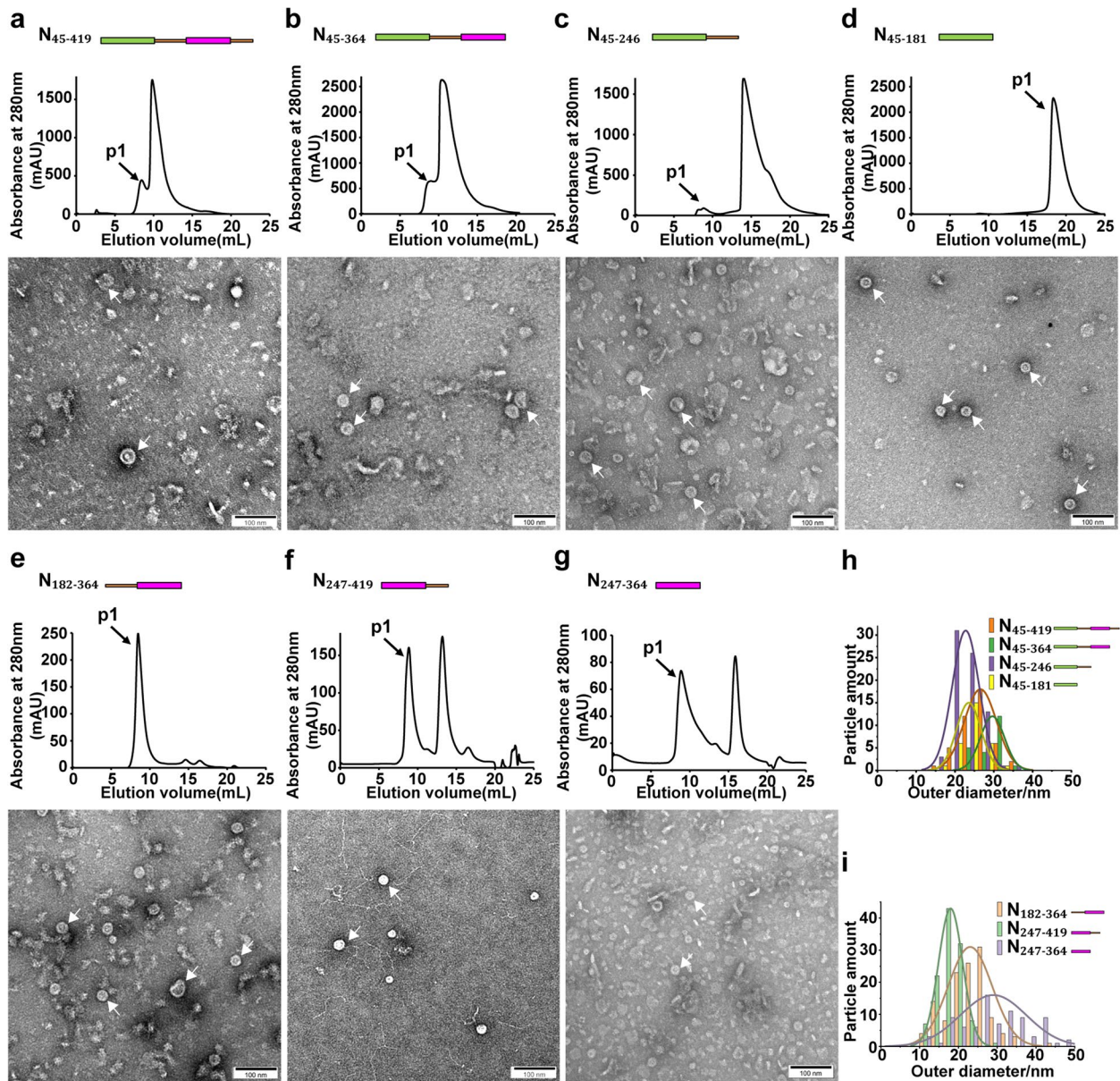
Fig. 3 The assembly of different N constructs under 2 M NaCl high salt condition. **a-g** The gel filtration chromatography profiles and representative negative staining images with ring-like structures of (**a**) N_{45-419}, (**b**) N_{45-364}, (**c**) N_{45-246}, (**d**) N_{45-181}, (**e**) $N_{182-364}$, (**f**) $N_{247-419}$ and (**g**) $N_{247-364}$ under 2 M NaCl condition with cellular RNAs removed. Scale bars, 100 nm. **h-i** Diameter distributions of (**h**) N_{45-419}, N_{45-364}, N_{45-246}, N_{45-181} and (**i**) $N_{182-364}$, $N_{247-419}$, $N_{247-364}$ ring-like assembly showed variable sizes. White arrows indicate ring-like structures

assembly, we first obtained the dimeric N_{FL} and $N_{182-419}$ by buffer exchanging from high salt to low salt (20 mM NaCl) to disrupt existing protein assembly (Fig. 5). The absence of nucleic acids was confirmed by absorbance measurement and SDS-PAGE detection (Supplementary Fig. 1). Either 5′ UTR, PS100, PS576, PS9, PS97 or 3′ UTR of SARS-CoV-2 were folded in the presence of 10 mM Mg^{2+} before incubation with the dimeric N constructs under physiological relevant salt condition (150 mM

NaCl) with a ratio of RNA to protein as 1:100, which is determined according to an estimated average ratio of RNA to protein based on previous structural information of N-RNA complexes of single-stranded RNA (ssRNA) viruses [59], since the full-length SARS-CoV-2 N-RNA complex is not resolved. Gel filtration chromatography revealed large shift of the PEVs of different N constructs from ~20 mL to prior to 10 mL in the presence of different viral RNAs (Fig. 5). The presence of
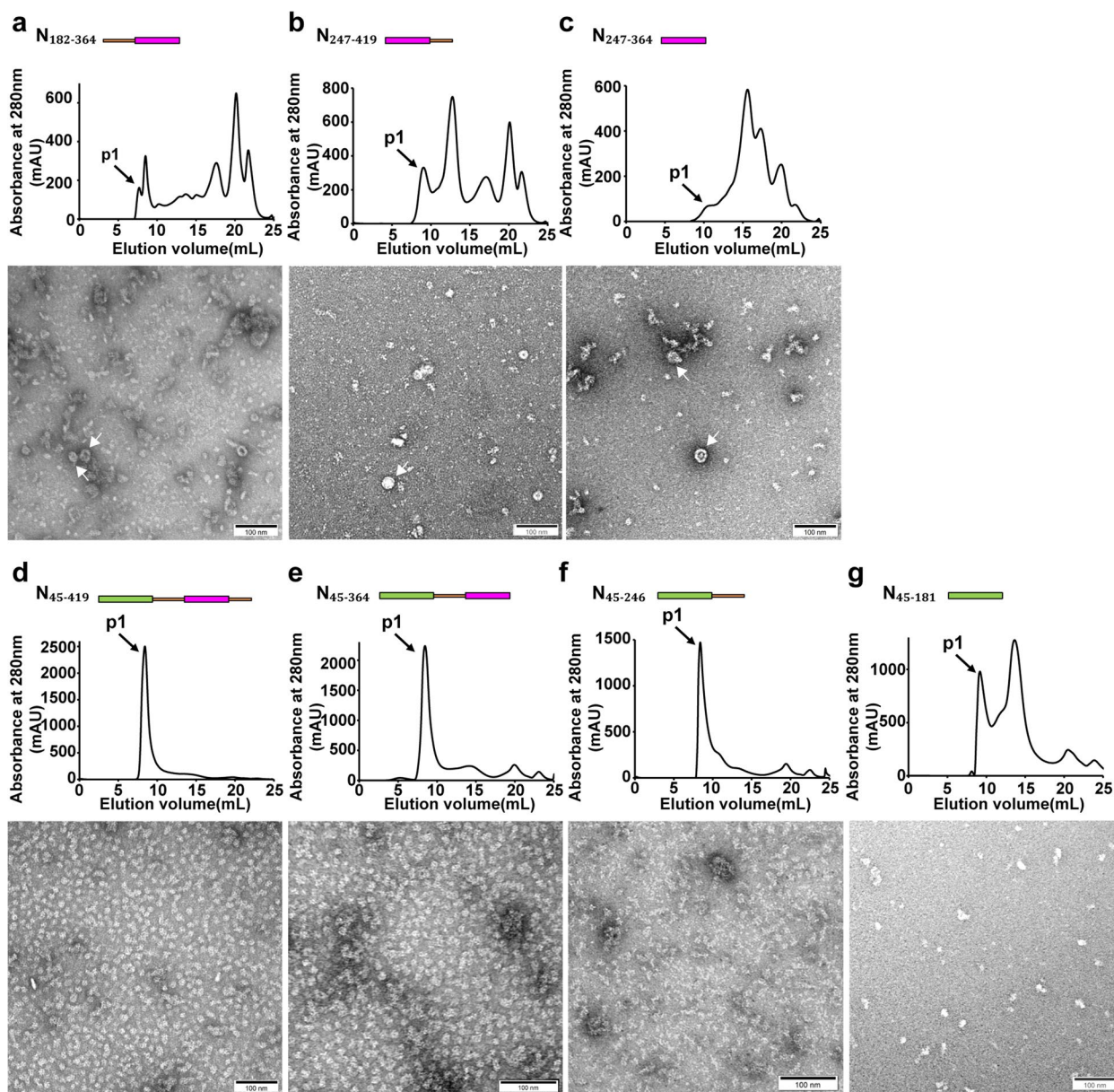
Fig. 4 The assembly of different N constructs under 150 mM NaCl physiological relevant salt condition. **a-g** The gel filtration chromatography profiles and representative negative staining images with ring-like structures of (**a**) $N_{182-364}$, (**b**) $N_{247-419}$, (**c**) $N_{247-364}$, and without ring-like structures of (**d**) N_{45-419}, (**e**) N_{45-364}, (**f**) N_{45-246} and (**g**) N_{45-181} under 150 mM NaCl condition with cellular RNAs. Scale bars, 100 nm. White arrows indicate ring-like structures

(See figure on next page.)

Fig. 5 N_{FL} and $N_{182-419}$ assembly in the presence of SARS-CoV-2 RNAs in vitro. **a-b** The dimeric N_{FL} (**a**) and $N_{182-419}$ (**b**) without cellular RNAs were obtained by buffer exchange from 2 M NaCl high salt condition to 20 mM NaCl extremely low salt condition (red), then incubated under 150 mM NaCl physiological relevant salt condition in the presence of SARS-CoV-2 5' UTR (yellow), 3' UTR (green), PS9 (purple), PS576 (magenta), PS100 (cyan), PS97 (brown), 2 M NaCl high salt condition (blue) or 150 mM NaCl physiological relevant salt condition (black) with representative negative staining images under each condition. Scale bars, 100 nm. White arrows indicate filamentous structures
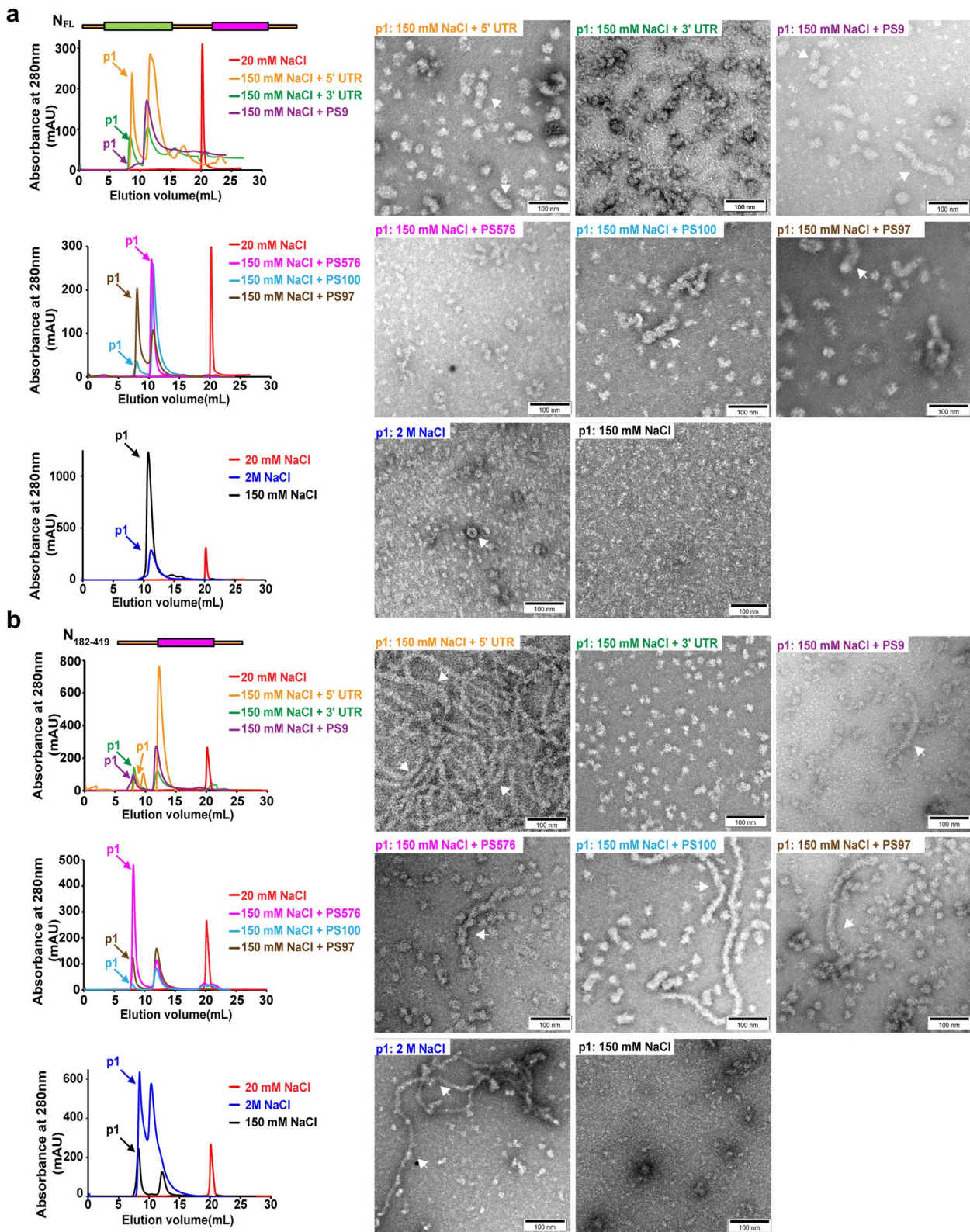
Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 7 of 14

Fig. 5 (See legend on previous page.)

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 8 of 14

viral RNAs were confirmed by absorbance measurement. While the N_{FL} could assemble into short filaments, $N_{182-419}$ could assemble back to long filaments in the presence of PS, but not 3' UTR, validating that PS could facilitate N assembly. Electrophoresis mobility shift assay (EMSA) suggested that all PS bind to N_{FL} and $N_{182-419}$ constructs (Supplementary Fig. 1), indicating that higher-order assembly might be facilitated by direct binding of PS to N proteins. We also tested if we could recapitulate the N protein assembly in vitro by incubations of the N constructs under different salt conditions without viral RNAs. Although we have not observed any N assembly under physiological relevant salt condition (150 mM NaCl), under high salt condition (2 M NaCl) the N_{FL} showed some ring-like assembly whereas $N_{182-419}$ showed filamentous assembly as expected, confirming that the N assemblies we observed could reversibly disassemble into dimeric N constructs.

N and $N_{182-419}$ constructs form variable phase separated droplets that could be enlarged by SARS-CoV-2 5' UTR

SARS-CoV-2 N protein undergoes LLPS to form partially ordered gel-like condensates for RNA compaction that could potentially facilitate viral RNA synthesis and genome packaging [29, 36–39, 41]. However, modular functions of individual domains of N in LLPS and the impacts of LLPS on N assembly remain underexplored.

We found that all N constructs formed droplets of similar diameters in the presence or absence of 5' UTR (Fig. 6, Supplementary Fig. 2), except that the droplets formed by N_{FL} and $N_{182-419}$ were significantly enlarged by adding 5' UTR RNA (Fig. 6a-c). We used correlative light and electron microscopy (CLEM) to correlate fluorescent droplets of $N_{182-419}$ and 5' UTR under cryo-electron microscopy (cryo-EM), in which we observed curved filamentous structures of roughly 20 nm diameter (Fig. 6d). This result suggests that the compact SARS-CoV-2 RNA-N protein condensates formed by LLPS could be enlarged in order to facilitate N assembly for viral RNA synthesis and packaging.

Discussion

Many ssRNA viral nucleocapsids that form filamentous assembly were found to pack as ring-like assembly using X-ray crystallography or cryo-EM [60–62]. SARS-CoV-2 N protein is highly conserved and abundant, and previous studies also suggested that N might assemble into filaments to protect viral RNA during transcription, replication, and packaging [24–30].

In this study, we first observed formation of ring-like structures in N_{FL} that suggest helical assembly of the nucleocapsid in ssRNA viruses [60, 63]. Subsequently, we generated multiple modular constructs of SARS-CoV-2 N protein and uncovered the impact that individual domain
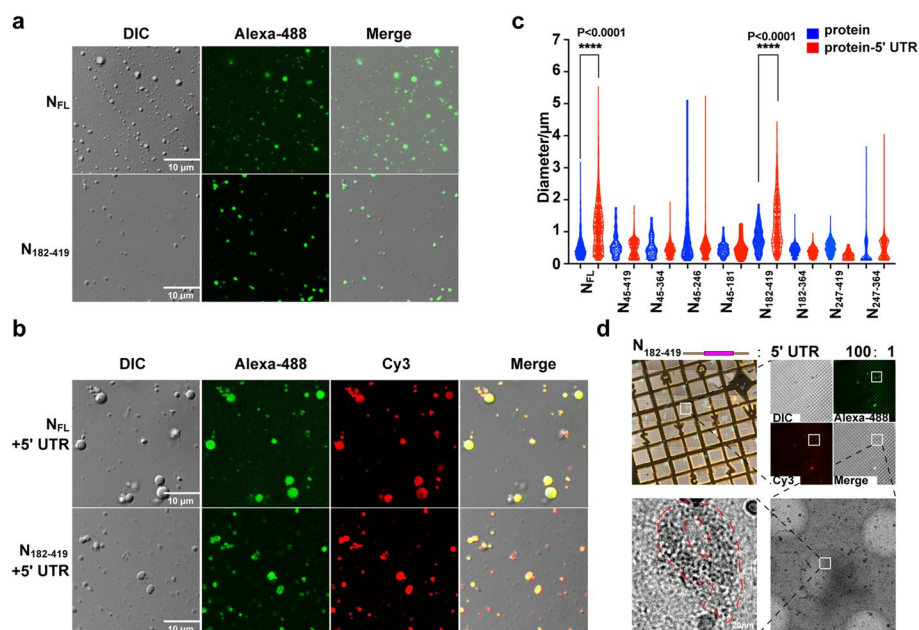

Fig. 6 LLPS droplets formed by N_{FL} and $N_{182-419}$ could be enlarged by viral 5' UTR that facilitates protein assembly. **a** N_{FL} and $N_{182-419}$ induced phase separation and formed droplets of variable sizes. Scale bars, 10 μm. **b** Viral 5' UTR enlarged LLPS droplets formed by N_{FL} and $N_{182-419}$. Scale bars, 10 μm. **c** Diameter analysis of ~1000 droplets of each construct revealed that droplets formed by N_{FL} and $N_{182-419}$ in the presence of 5' UTR were significantly enlarged. **d** CLEM correlated fluorescent LLPS droplets formed by $N_{182-419}$ under cryo-EM, in which curved filamentous structures were observed (outlined by red dashed line). Scale bar, 20 nm. ****$P < 0.0001$ by two-tailed Student's T-test

has on protein assembly. Intriguingly, $N_{182-419}$ that contains SR_{IDR}-CTD-C_{IDR} domains could assemble into filaments that may play essential roles as templates for viral transcriptions and replications: The SR_{IDR} domain could drive LLPS to suppress host immune system that is critical for viral genome packaging [41], and undergo phosphorylation that is critical for transcriptions [64]; The CTD domain could bind to RNA and oligomerize during packaging [30], and bind to nsp3 to play a crucial role in transcriptions [65]; The function of C_{IDR} domain is associated with M that also affects viral packaging [41].

Several segments on the viral genomes of SARS-CoV and SARS-CoV-2 have been previously suggested as the packaging signals, including 5' UTR, PS9, PS69, PS97, PS100 and PS576 [41, 47–49, 48]. Intriguigingly, the filament could be assembled in the presence of PS, whereas addition of 3' UTR viral RNA would not result in filament formation. Multiple studies have shown that mutations and repetitive structural motifs in 5' UTR were critical for viral packaging in various coronaviruses [46, 47, 66–69], whereas PS9 has been recently identified as the minimal packaging signal element for SARS-CoV-2 virus-like particle [50].

All constructs that contained CTD formed either ring-like or filamentous higher-order assembly that could be disrupted by inclusion of the NTD domain in the constructs under the physiological relevant salt condition that preserved cellular RNAs, which suggested that NTD could suppress higher-order N assembly. The fact that N_{FL} could still form ring-like structures indicated that inclusion of the N_{IDR} domain in N_{FL} could reduce the inhibitory effect of NTD against higher-order N assembly. Previous study has reported that the N-terminus, N_{1-209}, could bind to the host factor cyclophilin A, which participates in replication cycle of coronaviruses and

other viral assembly processes [70]. Since numerous proteins from host cells have been suggested to interact with N [71, 72], it is possible that during viral transcriptions and replications in cells, additional viral and host factors may interact with NTD to completely diminish its inhibitory role in higher-order N assembly so that filamentous nucleocapsids could be formed as templates for efficient viral transcription and replication. Future study of N assembly formations in human cells or cell lysates may shed light on the N-terminal functions.

LLPS has been suggested to condense viral proteins and RNAs for efficient packaging and replication [73–76]. SARS-COV-2 N protein has been previously reported to undergo LLPS [38]. Intriguingly, we noticed that addition of viral 5' UTR to both N_{FL} and $N_{182-419}$ enlarged the droplet size with flexible filamentous structures observed in $N_{182-419}$ droplets under CLEM, indicating that larger space might facilitate higher-order N assembly. We propose a working model of N that undergoes LLPS to condensate with SARS-CoV-2 genomic RNA to facilitate RNP assembly (Fig. 7).

In conclusion, we have carried out a modular approach to characterize functions of individual domains of SARS-CoV-2 N in protein assembly and LLPS. These findings further complete our understandings in N protein assembly that may play essential role in RNA replication and packaging.

Materials and methods
Modular construct design and plasmid construction
The pET28a-SARS-CoV-2-N_{FL} plasmid was kindly provided by Guangdong Laboratory Animal Testing Institute. To investigate the function of each domain of SARS-CoV-2 N in protein assembly and LLPS, a modular research approach was designed based on the feature
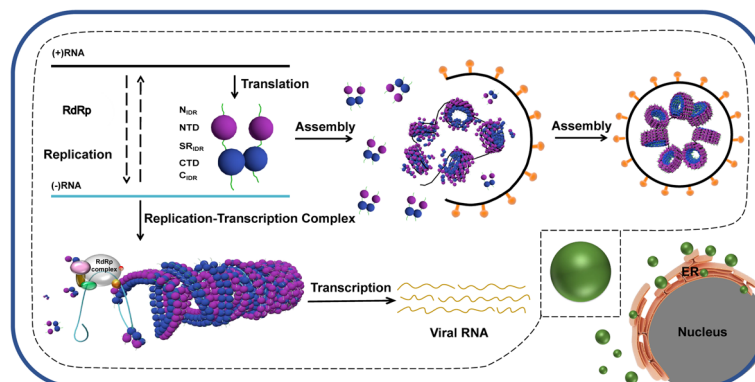

Fig. 7 SARS-CoV-2 N protein assembly facilitated by phase separation. SARS-CoV-2 N protein participates in viral RNA transcription and replication, and forms nucleocapsid-like assembly to protect the viral RNA template during transcription. N protein also initiates viral genomic RNA packaging that eventually forms multiple globular ribonucleoprotein complexes. LLPS induced by N protein might enable the assembly processes

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 10 of 14

that the N protein consists of two domains and three intrinsically disordered regions. Centering on the two domains, we truncated the intrinsic disordered regions at the N-terminus and C-terminus step by step, and constructed eight different N-modular mutants. The SARS-CoV-2-N truncations (pET-28a-N_{45-419}, pET-28a-N_{45-364}, pET-28a-N_{45-246}, pET-28a-N_{45-181}, pET-28a-$N_{182-419}$, pET-28a-$N_{182-364}$, pET-28a-$N_{247-419}$, pET-28a-$N_{247-364}$) were amplified with pET28a-SARS-CoV-2-N_{FL} as the template and cloned into the pET-28a vector between NdeI and XhoI by ClonExpress® II one step cloning kit (Vazyme). The 5' UTR, PS100, PS576, PS9, PS97 or 3' UTR were designed by combining a T7 RNA polymerase promoter sequence and the hepatitis δ virus ribozyme (HDV) sequence to produce uniform 5' ends, in addition to the desired sequence from the full-length SARS-CoV-2 cDNA, which was a kind gift from Beijing Institute of Microbiology and Epidemiology. The combined sequences were inserted into the pUC-19 vector between HindIII and NdeI restriction sites for RNA transcription and purification.

Protein expression and purification

pET28a-SARS-CoV-2-N_{FL} and the truncated protein plasmids (Supplementary Table 1) were individually transformed into *Escherichia coli* (*E. coli*) BL21 (DE3) strain (Novagen). *E. coli* cells were cultured in Luria–Bertani (LB) medium at 37 ℃ with 50 mg/L kanamycin until the OD600 reached 0.6–0.8, then the bacteria were induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18 ℃ for 15–18 h. Bacteria were collected by centrifugation, resuspended in buffer containing 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 10 mM Imidazole, 5% Glycerol, 1 mM phenylmethylsulfonyl fluoride to retain the bacterial nucleic acids and lysed by Ultrasonic Cell Crushe. Cell extracts were centrifuged at 15,000×g for 40 min at 4 ℃. Supernatants were purified with Ni–NTA (GE Healthcare), the target protein was washed with lysis buffer and then eluted with a buffer containing 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 300 mM imidazole. Eluted proteins were concentrated by centrifugal ultrafiltration, loaded onto a pre-equilibrated Superdex™ 200 Increase 10/300 GL column in an ÄKTA-purifier (GE Healthcare), eluted at a flow rate of 0.4 ml/min with the same buffer containing 20 mM Tris–HCl pH 8.0, 150 mM NaCl. Peak fractions were analyzed by SDS-PAGE (15%, w/v) and stained with Coomassie brilliant blue R-250. The peak fractions of proteins retaining the bacterial nucleic acids were concentrated to 0.5 mg/ml for negative stain EM. Bacterial nucleic acids were removed under 2 M NaCl condition with the rest of the procedure unchanged in the above-mentioned protein purification process.

The N protein assembly was disrupted by buffer exchange from 2 M NaCl high salt condition to 20 mM NaCl low salt condition as confirmed by gel filtration chromatography and SDS-PAGE analysis. Fractions were pooled together and concentrated by centrifugal ultrafiltration (Millipore). All sample concentrations were determined by A280 (NanoDrop One^C, Thermo Scientific). The storage protein samples were quick-frozen by liquid nitrogen then kept at-80 ℃ before use.

RNA preparation

The SARS-CoV-2 5' UTR, PS9, PS576, PS100, PS97 and 3' UTR DNA templates in pUC-19 plasmid (Supplementary Table 1) were amplified by Qiagen MegaPrep kits. Linear dsDNA templates were acquired with the regular forward primer and 2'-O-methylated reverse primer to improve 3' end homogeneity by PCR [77] (See Supplementary Table 2 for primers). The resulting DNA templates were isolated by ethanol precipitation. In vitro transcription of SARS-CoV-2 RNAs were carried out with 1.5–2 µg DNA templates, 4 mM NTPs and 1 U/µL RNase inhibitor in 1×transcription buffer containing 40 mM Tris–HCl (pH 7.9), 0.01% TritonX-100, 20 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, and incubated at 37 ℃ for 3 h. The transcription products were mixed with 2×denaturing gel loading buffer containing 95% formamide, 0.025% SDS, 10 mM EDTA, 0.025% xylene cyanol, and 0.025% bromophenol blue, and loaded on an 6–10% 29:1 acrylamide:bis, 7 M urea polyacrylamide gel. The gel was run at 300 V for 6 h, then stained for 5 min in SyBrGold (Invitrogen) and visualized by UV transillumination Molecular Imager (Biorad ChemiDoc™ XRS+). RNA ladder was purchased in Thermo Scientific (SM1833). To acquire SARS-CoV-2 RNAs for assembly with N protein in vitro, RNAs were gel purified and visualized briefly with a 254-nm UV lamp, held far from the gel to minimize RNA damages [78]. Then the RNAs were eluted from the gel overnight in elution buffer containing 30 mM sodium acetate (pH 5.5) and 1 mM EDTA on an active vortexer at 4 ℃ overnight. The resulting gel slurry was then filtered through 0.45 µm filters (Minisart® Syringe Filter, Sartorius). The resulting RNAs were precipitated with isopropyl alcohol to remove urea and salts. The products were pelleted by centrifugation at 15,000×g for 1 h at 4 ℃, washed with 80% cold ethanol for three times and dried in a Speedvac, then resuspend the pellet in RNase-free water. RNA was quantified using a NanoDrop One^C (Thermo Scientific) and kept at -80 ℃ before use.

RNA refolding

The 5' UTR, PS9, PS576, PS100, PS97 and 3' UTR RNAs were added to refolding buffer (10 mM Tris–HCl pH 7.4,

100 mM KCl). The RNA solution was heated at 90 ℃ for 3 min, cooled to 25 ℃ for 10 min, then added $MgCl_2$ to a final concentration of 10 mM, followed by incubation at 50 ℃ for 30 min and 25 ℃ for 15 min. Refolded RNA samples were mixed with 5×gel loading buffer containing 10% glycerol, 0.025% xylene cyanol, and 0.025% bromophenol blue, and loaded on an 6% 29:1 acrylamide: bis polyacrylamide gel. The gel was run at 110 V for 1.5 h, then stained for 5 min in SyBrGold (Invitrogen) and visualized by UV transillumination Molecular Imager (Biorad ChemiDoc™ XRS +). The refolded RNAs were kept on ice before use.

N protein assembly in vitro

All dimeric N constructs were incubated with refolded viral 5' UTR, PS9, PS576, PS100, PS97 or 3' UTR with a protein to RNA molar ratio of 100:1, at 37 ℃ for 30 min in the assembly buffer containing 20 mM Tris–HCl pH 8.0, 10 mM KCl, 140 mM NaCl, 1 mM $MgCl_2$. In the absence of viral RNAs, dimeric N constructs were incubated in buffer containing 20 mM Tris–HCl pH 8.0, 2 M NaCl to obtain assembly under high salt condition. The resulting assembly mixtures were loaded onto a pre-equilibrated Superdex™ 200 Increase 10/300 GL column in an ÄKTA-purifier (GE Healthcare), eluted at a flow rate of 0.4 ml/min with the buffer containing 20 mM Tris–HCl pH 8.0, 10 mM KCl, 140 mM NaCl, 1 mM $MgCl_2$ in the presence of viral RNAs, and 20 mM Tris–HCl pH 8.0, 2 M NaCl in the absence of viral RNAs. Peak fractions were analyzed by SDS-PAGE (15%, w/v) and stained with Coomassie brilliant blue R-250. All fractions were concentrated to 0.5 mg/ml for negative stain EM.

EMSA assay

The 9 pmol refolded SARS-CoV-2 5' UTR, PS9, PS576, PS100 or PS97 binds to N_{FL} protein or $N_{182-419}$ (storage buffer: 20 mM Tris–HCl pH 8.0, 20 mM NaCl) by the ratio of 1:1, 1:2, 1:4, 1:8 in buffer 10 mM Tris–HCl pH 7.4, 10 mM NaCl, 1 mM KCl, 10 mM $MgCl_2$. After incubating at 37 ℃ for 30 min, SARS-CoV-2 5' UTR, PS576, PS100 or PS97 detect with 6% native urea gel and PS9 1% native agarose gel (1000nts RNA for native urea gel is too long, 6% native urea gel can't be detected, but data not show), then stained for 5 min in SyBrGold (Invitrogen) and visualized by UV transillumination Molecular Imager (Biorad ChemiDoc™ XRS +).

Fluorescence labeling and LLPS assay

The dimeric N constructs with removed cellular nucleic acids were labeled by fluorescence dye Oregon-Green488 (Invitrogen, 2,161,802) with a molar ratio 10:1 between protein and fluorescence dye in labeling buffer 50 mM NaPhosphate pH 7.0, 50 mM NaCl. The mixtures were incubated at 25 ℃ for 1 h. To remove the excess fluorescence dye, the resulting mixture was buffer-exchanged with labeling buffer using concentrator columns with 10 kDa cutoff (Ultrafiltration Centrifugal Tube, Millipore). The SARS-CoV-2 5' UTR and 3' UTR were labeled by fluorescence dye Cy3 (Lumiprobe, #41,070) following the Cy3 labelling protocol [79]. The labeled proteins and RNAs were quantified using a NanoDrop One^C (Thermo Scientific).

Oregon-Green488 labeled N constructs were diluted to a final concentration of 25 μM in phase separation buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl) containing 10% PEG 3350, and RNAs were added with a protein to RNA molar ratio of 100:1. Then the mixture was incubated at room temperature for 5 min. A total of 10 μL solution was transferred onto the glass slide, and images were collected using a Zeiss-Axio Observer 7 microscope.

CLEM imaging

A total of 3 μL of the LLPS sample containing $N_{182-419}$ and 5' UTR was applied onto glow-discharged (45 s) Quantifoil Au H2 finder (R 2/1) grids (Quantifoil Micro Tools GmbH, Germany). The grids were blotted with filter paper for 2.5 s in 100% humidity at 4 ℃ with no blotting offset and rapidly frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher). The fluorescent images were collected by LEICA EM Cryo CLEM microscope and LAS X software at 60×magnification.

Frozen grids were then loaded into a Titan Krios cryo-electron microscope (Thermo Fisher) operated at 300 kV with a 50 μm condenser lens aperture and spot size 5. Micrographs were collected using EPU software (Version 2.9.0.1519REL) under various magnifications.

Negative stain EM sample preparation and data collection

Three drops of 20 μL, 20 μL, and 60 μL 0.75% uranium formate (UF) stain solution were applied on the parafilm (Bemis). Then, 3 μL of sample was applied on the glow-discharged (40 s) 300-mesh-Cu grids (Quantifoil) coated with a continuous carbon film and waited 60 s to allow sample adsorption. The grid was blotted from the side with a piece of filter paper and washed with 4 μL of washing buffer. Excessive washing buffer was blotted with filter paper and the grid was stained in the first two drops of UF followed by blotting with filter paper. Subsequently, the grid was stained in the third drop of UF for 40 s. Excessive UF was blotted and the grid was air-dried and stored until imaging. The grid was placed on a side-entry holder and loaded into a JEM-1400 operated at 120 kV, condenser lens aperture 150 μm, spot size 1. Micrographs were collected using RADIUS software on

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 12 of 14

a Morada G3 direct electron camera under magnification of 120,000× (corresponding to a calibrated sampling of 3.23 Å per physical pixel).

Image processing and 3D reconstruction

The 99 negative stain images of N_{FL} assembly were subjected to EMAN2 [80] for neural network particle picking. A total of 6,152 particles were extracted in Relion 3.1 [81] with a box size of 140 pixels. After 2D classifications in cryoSPARC3.1 [82], 2,384 particles were subjected to one round of ab-initio reconstruction and heterogeneous refinement to remove poor-quality particles. The best class was selected for homogeneous refinement, and the map after homogeneous refinement was displayed in UCSF Chimera [83].

The 21 negative stain images of $N_{182-419}$ assembly were processed until 2D classification following the same protocol above with a box size of 320 pixels. After 2D classification, 14,761 particles were subjected to helical reconstruction in cryoSPARC3.1 to obtain a low-resolution 3D density map and displayed in UCSF Chimera.

Statistics and reproducibility

All the experiments were independently repeated for at least three times, and no inconsistent results were observed. Origin software and GraphPad Prism v.9.0.0 was used to perform statistical analyses. To determine the partition coefficient of indicated groups, 3 microscopy images were randomly selected, and the diameter of particles or droplets was acquired with Image J. Data are presented as the mean $\pm$ S.D or $\pm$ S.E.M. The box borders in the boxplots and violin plots represent the upper and lower quartiles (25th and 75th percentiles), and the center line denotes the median. A standard two-tailed unpaired Student's t-test was used for statistical analysis of two groups; $*p < 0.05$, $**p < 0.01$, $****p < 0.0001$, and $p > 0.05$ is considered as non-significant. All the data were reproducible, and details of replicates are described in the figure legends.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43556-023-00129-z.

Additional file 1: Supplementary Fig 1. EMSA analysis of N_{FL} or $N_{182-419}$ binding with SARS-CoV-2 5'UTR, PS9, PS576, PS100 and PS97. SARS-CoV-2 5'UTR, PS9, PS576, PS100 or PS97 (9 pmol) and N_{FL} protein or $N_{182-419}$ were combined at a ratio of 1:1, 1:2, 1:4, 1:8. After incubation at 37°C for 30 min, detect with 6% or 1% Native PAGE gel. **Supplementary Fig 2.** LLPS analysis of different N constructs in the presence and absence of viral 5'UTR. a-b N_{45-419}, N_{45-364}, N_{45-246}, N_{45-181}, $N_{182-364}$, $N_{247-419}$ and $N_{247-364}$ undergo phase separation to form variable droplets (a) without SARS-CoV-2 5'UTR and (b) with 5'UTR colocalized in the droplets. Scale bars, 10 μm. **Supplementary Table 1.** RNA sequences used in this study. **Supplementary Table 2.** PCR primers used in this paper.

Availability of data and materials

The EM reconstructions have been deposited in EMDB under EMD accession codes of EMD-35527 for N_{FL} and EMD-35528 for $N_{182-419}$, respectively. All other data are available from the authors upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. Uraki R, Ito M, Furusawa Y, Yamayoshi S, Iwatsuki-Horimoto K, Adachi E, et al. Humoral immune evasion of the omicron subvariants BQ.1.1 and XBB. Lancet Infect Dis. 2023;23(1):30–2. https://doi.org/10.1016/s1473-3099(22)00816-7.
2. Ni Y, Alu A, Lei H, Wang Y, Wu M, Wei X. Immunological perspectives on the pathogenesis, diagnosis, prevention and treatment of COVID-19. Mol Biomed. 2021;2(1):1. https://doi.org/10.1186/s43556-020-00015-y.
3. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579(7798):270–3. https://doi.org/10.1038/s41586-020-2012-7.
4. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265–9. https://doi.org/10.1038/s41586-020-2008-3.
5. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382(8):727–33. https://doi.org/10.1056/NEJMoa2001017.
6. Hu B, Guo H, Zhou P, Shi Z-L. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol. 2021;19(3):141–54. https://doi.org/10.1038/s41579-020-00459-7.
7. Kim D, Lee J-Y, Yang J-S, Kim JW, Kim VN, Chang H. The architecture of SARS-CoV-2 transcriptome. Cell. 2020;181(4):914-921.e910. https://doi.org/10.1016/j.cell.2020.04.011.

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 13 of 14

8.   Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. The Lancet. 2020;395(10224):565–74. https://doi.org/10.1016/S0140-6736(20)30251-8.

9.   V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication5 implications for SARS-CoV-2. Nature Rev Microbiol. 2021;19(3):155–70. https://doi.org/10.1038/s41579-020-00468-6.

10. McBride R, van Zyl M, Fielding BC. The coronavirus nucleocapsid is a multifunctional protein. Viruses. 2014;6(8):2991–3018. https://doi.org/10.3390/v6082991.

11. Yingying Cong MU, Schepers Hein, Mauthe Mario, V'kovski Philip, Kriegenburg Franziska, Thiel Volker, de HaanCornelis AM, Reggiori F. Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle. J Virol. 2020;94(4):e01925-01919. https://doi.org/10.1128/JVI.01925-19.

12   Chang C-k, Hou M-H, Chang C-F, Hsiao C-D, Huang T-h. The SARS coronavirus nucleocapsid protein – Forms and functions. Anti Res. 2014;103:39–50. https://doi.org/10.1016/j.antiviral.2013.12.009.

13. Scherer KM, Mascheroni L, Carnell GW, Wunderlich LCS, Makarchuk S, Brockhoff M, et al. SARS-CoV-2 nucleocapsid protein adheres to replication organelles before viral assembly at the Golgi/ERGIC and lysosome-mediated egress. Sci Adv. 2022;8(1):eabl4895. https://doi.org/10.1126/sciadv.abl4895.

14. Shapira T, Vimalanathan S, Rens C, Pichler V, Pena-Diaz S, Jordana G, et al. Inhibition of glycogen synthase kinase-3-beta (GSK3beta) blocks nucleocapsid phosphorylation and SARS-CoV-2 replication. Mol Biomed. 2022;3(1):43. https://doi.org/10.1186/s43556-022-00111-1.

15. Peng Y, Du N, Lei Y, Dorje S, Qi J, Luo T, et al. Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design. EMBO J. 2020;39(20):e105938. https://doi.org/10.15252/embj.2020105938.

16. Kang S, Yang M, Hong Z, Zhang L, Huang Z, Chen X, et al. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm Sin B. 2020;10(7):1228–38. https://doi.org/10.1016/j.apsb.2020.04.009.

17. Ye Q, West AMV, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci. 2020;29(9):1890–901. https://doi.org/10.1002/pro.3909.

18   Zhao H, Nguyen A, Wu D, Li Y, Hassan SA, Chen J, et al. Plasticity in structure and assembly of SARS-CoV-2 nucleocapsid protein. PNAS Nexus. 2022;1(2):049. https://doi.org/10.1093/pnasnexus/pgac049.

19   Dutta NK, Mazumdar K, Gordy JT. The nucleocapsid protein of SARS-CoV-2 a target for vaccine development. J Virol. 2020;94(13):e00647. https://doi.org/10.1128/JVI.00647-20.

20. Ni L, Ye F, Cheng ML, Feng Y, Deng YQ, Zhao H, et al. Detection of SARS-CoV-2-Specific Humoral and Cellular Immunity in COVID-19 Convalescent Individuals. Immunity. 2020;52(6):971-977 e973. https://doi.org/10.1016/j.immuni.2020.04.023.

21. Xiang FWX, He X, Peng Z, Yang B, Zhang J, Zhou Q, Ye H, Ma Y, Li H, et al. Antibody Detection and Dynamic Characteristics in Patients With Coronavirus Disease 2019. Clin Infect Dis. 2020;71(8):1930–4. https://doi.org/10.1093/cid/ciaa461.

22. Mu J, Fang Y, Yang Q, Shu T, Wang A, Huang M, et al. SARS-CoV-2 N protein antagonizes type I interferon signaling by suppressing phosphorylation and nuclear translocation of STAT1 and STAT2. Cell Discov. 2020;6:65. https://doi.org/10.1038/s41421-020-00208-3.

23. Xia J, Tang W, Wang J, Lai D, Xu Q, Huang R, et al. SARS-CoV-2 N protein induces acute lung injury in mice via NF-kB activation. Front Immunol. 2021;12:791753. https://doi.org/10.3389/fimmu.2021.791753.

24. Neuman BW, Adair BD, Yoshioka C, Quispe JD, Orca G, Kuhn P, et al. Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. J Virol. 2006;80(16):7918–28. https://doi.org/10.1128/jvi.00645-06.

25. Bárcena M, Oostergetel GT, Bartelink W, Faas FGA, Verkleij A, Rottier PJM, et al. Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. Proc Natl Acad Sci. 2009;106(2):582–7. https://doi.org/10.1073/pnas.0805270106.

26. Davies HA, Dourmashkin RR, Macnaughton MR. Ribonucleoprotein of avian infectious bronchitis virus. J Gen Virol. 1981;53(1):67–74. https://doi.org/10.1099/0022-1317-53-1-67.

27. Macnaughton MR, Davies HA, Nermut MV. Ribonucleoprotein-like structures from coronavirus particles. J Gen Virol. 1978;39(3):545–9. https://doi.org/10.1099/0022-1317-39-3-545.

28. Gui M, Liu X, Guo D, Zhang Z, Yin C-C, Chen Y, et al. Electron microscopy studies of the coronavirus ribonucleoprotein complex. Protein Cell. 2017;8(3):219–24. https://doi.org/10.1007/s13238-016-0352-8.

29. Chen H, Cui Y, Han X, Hu W, Sun M, Zhang Y, et al. Liquid–liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA. Cell Res. 2020;30(12):1143–5. https://doi.org/10.1038/s41422-020-00408-2.

30. Chen C-Y, Chang C-k, Chang Y-W, Sue S-C, Bai H-I, Riang L, et al. Structure of the SARS Coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J Mol Biol. 2007;368(4):1075–86. https://doi.org/10.1016/j.jmb.2007.02.069.

31. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, et al. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nat Commun. 2020;11(1):5885. https://doi.org/10.1038/s41467-020-19619-7.

32. Liu C, Mendonça L, Yang Y, Gao Y, Shen C, Liu J, et al. The architecture of inactivated SARS-CoV-2 with postfusion spikes revealed by Cryo-EM and Cryo-ET. Structure. 2020;28(11):1218-1224.e1214. https://doi.org/10.1016/j.str.2020.10.001.

33. Hardenbrook NJ, Zhang P. A structural view of the SARS-CoV-2 virus and its assembly. Curr Opin Virol. 2022;52:123–34. https://doi.org/10.1016/j.coviro.2021.11.011.

34. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, et al. Molecular architecture of the SARS-CoV-2 virus. Cell. 2020;183(3):730-738.e713. https://doi.org/10.1016/j.cell.2020.09.018.

35. Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature. 2020;588(7838):498–502. https://doi.org/10.1038/s41586-020-2665-2.

36. Cubuk J, Alston JJ, Incicco JJ, Singh S, Stuchell-Brereton MD, Ward MD, et al. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. Nat Commun. 2021;12(1):1936. https://doi.org/10.1038/s41467-021-21953-3.

37. Jack A, Ferro LS, Trnka MJ, Wehri E, Nadgir A, Nguyenla X, et al. SARS-CoV-2 nucleocapsid protein forms condensates with viral genomic RNA. PLoS Biol. 2021;19(10):e3001425. https://doi.org/10.1371/journal.pbio.3001425.

38. Cascarina SM, Ross ED. Phase separation by the SARS-CoV-2 nucleocapsid protein: Consensus and open questions. J Biol Chem. 2022;298(3):101677. https://doi.org/10.1016/j.jbc.2022.101677.

39. Wang S, Dai T, Qin Z, Pan T, Chu F, Lou L, et al. Targeting liquid–liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity. Nat Cell Biol. 2021;23(7):718–32. https://doi.org/10.1038/s41556-021-00710-0.

40. Iserman C, Roden CA, Boerneke MA, Sealfon RSG, McLaughlin GA, Jungreis I, et al. Genomic RNA elements drive phase separation of the SARS-CoV-2 Nucleocapsid. Mol Cell. 2020;80(6):1078-1091.e1076. https://doi.org/10.1016/j.molcel.2020.11.041.

41. Lu S, Ye Q, Singh D, Cao Y, Diedrich JK, Yates JR, et al. The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. Nat Commun. 2021;12(1):502. https://doi.org/10.1038/s41467-020-20768-y.

42   Perdikari TM, Murthy AC, Ryan VH, Watters S, Naik MT, Fawzi NL. SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. EMBO J. 2020;39(24):e106478. https://doi.org/10.15252/embj.2020106478.

43. Luo L, Li Z, Zhao T, Ju X, Ma P, Jin B, et al. SARS-CoV-2 nucleocapsid protein phase separates with G3BPs to disassemble stress granules and facilitate viral production. Sci Bull. 2021;66(12):1194–204. https://doi.org/10.1016/j.scib.2021.01.013.

44. Luo HYF, Chen K, et al. SR-rich motif plays a pivotal role in recombinant SARS coronavirus nucleocapsid protein multimerization. Biochemistry. 2005;44(46):15351–8. https://doi.org/10.1021/bi051122c.

45. Carlson CR, Asfaha JB, Ghent CM, Howard CJ, Hartooni N, Safari M, et al. Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. Mol Cell. 2020;80(6):1092-1103 e1094. https://doi.org/10.1016/j.molcel.2020.11.025.

46. Miao Z, Tidu A, Eriani G, Martin F. Secondary structure of the SARS-CoV-2 5'-UTR. RNA Biol. 2021;18(4):447–56. https://doi.org/10.1080/15476286.2020.1814556.

Wang *et al. Molecular Biomedicine*   (2023) 4:16

Page 14 of 14

47. Escors D, Izeta A, Capiscol C, Enjuanes L. Transmissible gastroenteritis coronavirus packaging signal is located at the 5′ end of the virus genome. J Virol. 2003;77(14):7890–902. https://doi.org/10.1128/jvi.77.14.7890-7902.2003.

48. Woo J, Lee EY, Lee M, Kim T, Cho YE. An in vivo cell-based assay for investigating the specific interaction between the SARS-CoV N-protein and its viral RNA packaging sequence. Biochem Biophys Res Commun. 2019;520(3):499–506. https://doi.org/10.1016/j.bbrc.2019.09.115.

49. Carlson CR, Adly AN, Bi M, Howard CJ, Frost A, Cheng Y, et al. Reconstitution of the SARS-CoV-2 ribonucleosome provides insights into genomic RNA packaging and regulation by phosphorylation. J Biol Chem. 2022;298(11):102560. https://doi.org/10.1016/j.jbc.2022.102560.

50. Syed AM, Taha TY, Tabata T, Chen IP, Ciling A, Khalid MM, et al. Rapid assessment of SARS-CoV-2-evolved variants using virus-like particles. Science. 2021;374(6575):1626–32. https://doi.org/10.1126/science.abl6184.

51. Chen SC, van den Born E, van den Worm SH, Pleij CW, Snijder EJ, Olsthoorn RC. New structure model for the packaging signal in the genome of group IIa coronaviruses. J Virol. 2007;81(12):6771–4. https://doi.org/10.1128/jvi.02231-06.

52. Selzer L, Su Z, Pintilie GD, Chiu W, Kirkegaard K. Full-length three-dimensional structure of the influenza A virus M1 protein and its organization into a matrix layer. PLoS Biol. 2020;18(9):e3000827. https://doi.org/10.1371/journal.pbio.3000827.

53. Heaphy S, Finch JT, Gait MJ, Karn J, Singh M. Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. Proc Natl Acad Sci U S A. 1991;88(16):7366–70. https://doi.org/10.1073/pnas.88.16.7366.

54. Peng R, Zhu T, Oladejo BO, Musyoki AM, Cui Y, Shi Y, et al. In vitro assembly of Ebola virus nucleocapsid-like complex expressed in E. coli. Protein Cell. 2016;7(12):888–98. https://doi.org/10.1007/s13238-016-0314-1.

55. Chang CK, Chen CM, Chiang MH, Hsu YL, Huang TH. Transient oligomerization of the SARS-CoV N protein–implication for virus ribonucleoprotein packaging. PLoS One. 2013;8(5):e65045. https://doi.org/10.1371/journal.pone.0065045.

56. Zinzula L, Basquin J, Bohn S, Beck F, Klumpe S, Pfeifer G, et al. High-resolution structure and biophysical characterization of the nucleocapsid phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. Biochem Biophys Res Commun. 2021;538:54–62. https://doi.org/10.1016/j.bbrc.2020.09.131.

57. Luo H, Chen J, Chen K, Shen X, Jiang H. Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization. Biochemistry. 2006;45(39):11827–35. https://doi.org/10.1021/bi0609319.

58. Masters PS. Coronavirus genomic RNA packaging. Virology. 2019;537:198–207. https://doi.org/10.1016/j.virol.2019.08.031.

59. Luo M, Terrell JR, Mcmanus SA. Nucleocapsid structure of negative strand RNA Virus. Viruses. 2020;12(8):835.

60. Tawar RG, Duquerroy S, Vonrhein C, Varela PF, Damier-Piolle L, Castagné N, et al. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science. 2009;326(5957):1279–83. https://doi.org/10.1126/science.1177634.

61. Leung DW, Borek D, Luthra P, Binning JM, Anantpadma M, Liu G, et al. An intrinsically disordered peptide from ebola virus VP35 controls viral RNA synthesis by modulating nucleoprotein-RNA interactions. Cell Rep. 2015;11(3):376–89. https://doi.org/10.1016/j.celrep.2015.03.034.

62. Culver JN. Tobacco mosaic virus assembly and disassembly: determinants in pathogenicity and resistance. Annu Rev Phytopathol. 2002;40(1):287–308. https://doi.org/10.1146/annurev.phyto.40.120301.102400.

63. Green TJ, Zhang X, Wertz GW, Luo M. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. Science. 2006;313:357–60. https://doi.org/10.1126/science.1126953.

64. Wu C-H, Chen P-J, Yeh S-H. Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription. Cell Host Microbe. 2014;16(4):462–72. https://doi.org/10.1016/j.chom.2014.09.009.

65. Khan MT, Zeb MT, Ahsan H, Ahmed A, Ali A, Akhtar K, et al. SARS-CoV-2 nucleocapsid and Nsp3 binding: an in silico study. Arch Microbiol. 2021;203(1):59–66. https://doi.org/10.1007/s00203-020-01998-6.

66. Cai T, Yu Z, Wang Z, Liang C, Richard S. Arginine methylation of SARS-Cov-2 nucleocapsid protein regulates RNA binding, its ability to suppress stress granule formation, and viral replication. J Biol Chem. 2021;297(1):100821. https://doi.org/10.1016/j.jbc.2021.100821.

67. Alam A, Islam OK, Hasan MS, Islam MR, Mahmud S, Al-Emran HM, et al. Dominant clade-featured SARS-CoV-2 co-occurring mutations reveal plausible epistasis: An in silico based hypothetical model. J Med Virol. 2022;94(3):1035–49. https://doi.org/10.1002/jmv.27416.

68. Chen SC, Olsthoorn RCL, Yu CH. Structural phylogenetic analysis reveals lineage-specific RNA repetitive structural motifs in all coronaviruses and associated variations in SARS-CoV-2. Virus Evol. 2021;7(1):veab021. https://doi.org/10.1093/ve/veab021.

69. Bartas M, Brázda V, Bohálová N, Cantara A, Volná A, Stachurová T, et al. In-depth bioinformatic analyses of nidovirales including human SARS-CoV-2, SARS-CoV, MERS-CoV viruses suggest important roles of non-canonical nucleic acid structures in their lifecycles. Front Microbiol. 2020;11:1583. https://doi.org/10.3389/fmicb.2020.01583.

70. Lutomski CA, El-Baba TJ, Bolla JR, Robinson CV. Multiple roles of SARS-CoV-2 N protein facilitated by proteoform-specific interactions with RNA, host proteins, and convalescent antibodies. JACS Au. 2021;1(8):1147–57. https://doi.org/10.1021/jacsau.1c00139.

71. Chen Z, Wang C, Feng X, Nie L, Tang M, Zhang H, et al. Interactomes of SARS-CoV-2 and human coronaviruses reveal host factors potentially affecting pathogenesis. EMBO J. 2021;40(17):e107776. https://doi.org/10.15252/embj.2021107776.

72. Wu W, Cheng Y, Zhou H, Sun C, Zhang S. The SARS-CoV-2 nucleocapsid protein: its role in the viral life cycle, structure and functions, and use as a potential target in the development of vaccines and diagnostics. Virol J. 2023;20(1):6. https://doi.org/10.1186/s12985-023-01968-6.

73. Narayanan K, Maeda A, Maeda J, Makino S. Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. J Virol. 2000;74(17):8127–34. https://doi.org/10.1128/jvi.74.17.8127-8134.2000.

74. Alenquer M, Vale-Costa S, Etibor TA, Ferreira F, Sousa AL, Amorim MJ. Influenza A virus ribonucleoproteins form liquid organelles at endoplasmic reticulum exit sites. Nat Commun. 2019;10(1):1629. https://doi.org/10.1038/s41467-019-09549-4.

75. Heinrich BS, Maliga Z, Stein DA, Hyman AA, Whelan SPJ. Phase transitions drive the formation of vesicular stomatitis virus replication compartments. mBio. 2018;9(5):e02290. https://doi.org/10.1128/mBio.02290-17.

76. Nikolic J, Le Bars R, Lama Z, Scrima N, Lagaudrière-Gesbert C, Gaudin Y, et al. Negri bodies are viral factories with properties of liquid organelles. Nat Commun. 2017;8(1):58. https://doi.org/10.1038/s41467-017-00102-9.

77. Kao C, Zheng M, Rüdisser S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. RNA. 1999;5(9):1268–72. https://doi.org/10.1017/s1355838299991033.

78. Kladwang W, Hum J, Das R. Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. Sci Rep. 2012;2:517. https://doi.org/10.1038/srep00517.

79. Qin PZPAM. Site-specific labeling of RNA with fluorophores and other structural probes. Methods. 1999;18(1):60–70. https://doi.org/10.1006/meth.1999.0757.

80. Chen M, Dai W, Sun SY, Jonasch D, He CY, Schmid MF, et al. Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. Nat Methods. 2017;14(10):983–5. https://doi.org/10.1038/nmeth.4405.

81. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife. 2018;7:e42166. https://doi.org/10.7554/eLife.42166.

82. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290–6. https://doi.org/10.1038/nmeth.4169.

83. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12. https://doi.org/10.1002/jcc.20084.

Publisher's Note