Understanding the phase separation characteristics of nucleocapsid protein provides a new therapeutic opportunity against SARS-CoV-2

Dear Editor

To date, tens of millions of people have been infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), causing the outbreak of the respiratory disease named the coronavirus disease 2019 (COVID-19). As a newly emerged member of the coronavirus family, SARS-CoV-2 is an enveloped positive-strand RNA virus, which has probably the largest genome (approximately 30 kb) among all RNA viruses. The nucleocapsid (N) protein of SARS-CoV-2 is mainly responsible for recognizing and wrapping viral RNA into helically symmetric structures (Malik, 2020). It was also reported that N protein can boost the efficiency of transcription and replication of viral RNA, implying its vital and multifunctional roles in the life cycle of coronavirus (Surjit and Lal, 2008; Chang et al., 2014). Recently, several independent research teams have reported that N protein of SARS-CoV-2 is capable of undergoing liquid-liquid phase separation (LLPS) (Iserman et al., 2020; Perdikari et al., 2020; Savastano et al., 2020).

Here, we comprehensively determined the characteristics of the phase separation driven by the N protein of SARS-CoV-2 (termed SARS-CoV-2 N), and found that LLPS is involved in the interplay between the N protein-viral RNA complex of SARS-CoV-2 (termed SARS-CoV-2 N-RNA) and other viral proteins, such as nsp12. Importantly, we identified two small molecules targeting the SARS-CoV-2 N protein, which can intervene the phase separation properties of the N protein-viral RNA complex, thus probably improving the accessibility of other antiviral drugs (e.g., remdesivir) to their viral targets (e.g., nsp12/RdRp).

First, IUPred2 (Erdos and Dosztanyi, 2020) and PLAAC (Lancaster et al., 2014) programs showed that SARS-CoV-2 N is highly disordered, and contains three intrinsically disordered regions (IDRs), with one also displaying prion-like activity (Figs. 1A and S1). Then, we expressed and purified the recombinant SARS-CoV-2 N protein with an mEGFP-tag (a monomeric variant of EGFP, A206K) or a His-tag using a prokaryotic expression system to understand the properties of N-driven LLPS in vitro (Fig. S2A and S2B). Confocal fluorescence microscopy showed that SARS-CoV-2 N was readily self-associated to form numerous micron-sized spherical condensates (Fig. 1B and 1C). Further time-lapse observations revealed that the SARS-CoV-2 N condensates fused and coalesced into larger ones upon their intersections (Fig. 1D and Video S1), verifying the liquid-like properties of SARS-CoV-2 N condensates. We also used fluorescence recovery after photobleaching (FRAP) to deeply study the dynamics of internal molecules within the N protein condensates. Recovery of fluorescence within the bleached regions (Fig. 1E) showed that SARS-CoV-2 N can partially freely diffuse within the condensed phase, consistent with their liquid-like behavior. In addition, phase condensation of SARS-CoV-2 N was sensitive to the increase of ionic strength, suggesting that electrostatic interactions are important for its condensation (Fig. S3). Consistently, we also demonstrated that SARS-CoV-2 N can also undergo LLPS in cellulo and display liquid-like behavior, by ectopically expressing an mCherry-tagged version of SARS-CoV-2 N in Vero E6 cells (Fig. 1F, 1G, 1H and Video S2).

Given that N protein of coronavirus prefers to bind to the intergenic regions and exhibits high binding affinity with the UCUAA pentanucleotide repeats (Stohlman et al., 1988; Nelson et al., 2000; Lin et al., 2014), we found that the addition of this viral RNA to the purified SARS-CoV-2 N solution at a physiological salt concentration (150 mmol/L NaCl) resulted in robust co-phase separation of these two components, which was more predominant than that of a single one alone (Fig. 1I). Compared to the condensates of SARS-CoV-2 N alone, the droplets of N protein-RNA (termed SARS-CoV-2 N-RNA) complex were more spherical (Fig. 1J and Video S3) and exhibited slightly higher molecular exchange rates than those of N protein alone (Figs. 1K and S4). Moreover, although the synthesized viral RNA, the host-derived RNA (β-actin RNA), the dsRNA and the ssDNA (both derived from viral RNA sequences) all exhibited similar
Understanding the phase separation characteristics of nucleocapsid protein

**LETTER**

**A**

| SARS-CoV-2-N | IDR1 | NTD | IDR2 | CTD | IDR3 |
|--------------|------|-----|------|-----|------|
| 1 | 41 | 174 | 249 | 364 | 419 |

**B**

- mEGFP
- DIC Merge
- Ax546
- DIC Merge

**C**

- Bleaching
- Prebleach 0 min 2 min 4 min 8 min 10 min

**D**

- mEGFP-N 0 s 14 s 37 s 68 s 126 s

**E**

- Percentage of recovery
- t_{1/2} = 199 s

**F**

- Prebleach Bleach Recovery

**G**

- Percentage of recovery
- SARS-CoV-2-N in vivo
  - t_{1/2} = 29 s

**H**

- Prebleach Bleach Recovery

**I**

- SARS-CoV-2-N mEGFP
- RNA-30nt Cy5

**J**

- SARS-CoV-2-N + RNA mEGFP
- Cy5

**K**

- Percentage of recovery
- SARS-CoV-2-N + RNA
  - t_{1/2} = 136 s

**L**

- Area occupied (%)
  - Viral RNA, β-actin RNA
  - Nucleotides (μmol/L)

**M**

- mEGFP-N (μmol/L)
  - N-FL, ΔIDR1, ΔNTD, ΔIDR2, ΔCTD, ΔIDR3

**N**

- Protein:RNA = 1:1 (μmol/L)
  - N-WT, R203K, S118L, S196L, S197L

© The Author(s) 2021
Figure 1. SARS-CoV-2 N undergoes phase separation and viral RNA facilitates phase condensation of SARS-CoV-2 N.

(A) Bioinformatic analysis of the amino acid sequence of SARS-CoV-2 N. Schematic representation of the domain structure is shown on the top. IDR, intrinsically disorder region; NTD, N-terminal domain; CTD, C-terminal domain; |IUPred, prediction of intrinsic disorder; PLD, prediction of prion-like region (PLAAC).

(B) In vitro phase separation assays of 25 µmol/L mEGFP tagged SARS-CoV-2 N protein (mEGFP-N). Scale bar, 20 µm. 

(C) In vitro phase separation assays of 25 µmol/L full-length SARS-CoV-2 N protein labeled with Alx546. Scale bar, 20 µm. 

(D) Fusion of mEGFP-N (50 µmol/L) condensates. Data are representative of three independent experiments. Scale bar, 2.5 µm. 

(E) In vitro FRAP analysis of the condensates (n = 3) of mEGFP-N (3 µmol/L). Top, representative snapshots of condensates before and after bleaching. Bottom, average fluorescence recovery traces of mEGFP-N condensates. Data are representative of three independent experiments and presented as mean ± SD. Scale bar, 1 µm. 

(F) In cellulo FRAP analysis of the puncta of the expressed mCherry-N (SARS-CoV-2 N tagged with mCherry) in Vero E6 cells. Insets, representative snapshots of puncta before and after bleaching. The bleached punctum is marked with an arrow. Scale bars, 10 µm (bottom), 2 µm (insets). 

(G) The average fluorescence recovery traces of the puncta (n = 3) of the expressed mCherry-N protein in Vero E6 cells presented in (F). Data are representative of three independent experiments and presented as mean ± SD. 

(H) Fluorescence time-lapse microscopy of Vero E6 cells expressing mCherry-N. In (F) and (H), the “Red Fire” lookup table in the NIS-Elements Viewer was used to highlight the intensity difference. Two mCherry-N puncta are zoomed-in. Scale bars, 5 µm (bottom), 1 µm (insets). 

(I) Left, in vitro phase separation assays of mEGFP-tagged SARS-CoV-2 N (mEGFP-N) alone (375 nmol/L) or Cy5-labeled 30-nt viral RNA alone (1.5 µmol/L). Right, puncta formed by mEGFP-N (375 nmol/L) mixed with 30-nt viral RNA (375 nmol/L) in vitro. Scale bar, 5 µm. 

(J) Fusion upon contact of the condensates of mEGFP-N (50 µmol/L) with 30-nt viral RNA (25 µmol/L). Scale bar, 2.5 µm. 

(K) In vitro FRAP analysis of the condensates (n = 3) of mEGFP-N (3 µmol/L) with 30-nt viral RNA (1.5 µmol/L). Top, representative snapshots of condensates before and after bleaching. Bottom, average fluorescence recovery traces of mEGFP-N with viral RNA in condensates. Data are representative of three independent experiments and presented as mean ± SD. Scale bar, 1 µm. 

(L) Quantitative comparison of phase condensation of mEGFP-N with different nucleic acids. 

(M) Quantitative comparison of the phase condensation of full-length and truncated mEGFP-N proteins with 30-nt viral RNA (1.5 µmol/L). 

In (L) and (M), %Area Occupied = [Sum of area occupied by N protein condensates] × 100/[The whole area]. 

(N) In vitro phase separation assays of the Alx546-labeled wild-type (WT) N protein and four mutants with mutations on the serine-arginine (SR) rich region with viral RNA of different concentrations. The ratio of N protein to viral RNA was 1:1. Scale bar, 20 µm. 

The reaction buffer of in vitro phase separation assays in (B), (D and E) and (I–N) consists of 20 mmol/L HEPES, pH 7.4, 150 mmol/L NaCl, 5% glycerol, 1 mmol/L EGTA and 1 mmol/L MgCl₂. The assays in (C) were performed in 20 mmol/L Tris-HCl, pH 7.5 and 150 mmol/L NaCl.
Understanding the phase separation characteristics of nucleocapsid protein

**A**

![Chemical structures of CVL218 and PJ34](Image)

- **K_D = 4.7 μmol/L** for CVL218
- **K_D = 696 μmol/L** for PJ34

**B**

![Graphs of RU vs. Time](Image)

- DMSO: Response
- CVL218: Response
- PJ34: Response

**C**

![Graph of Percentage of Recovery vs. Time](Image)

- DMSO: t_{1/2} = 186 s
- CVL218: t_{1/2} = 158 s
- PJ34: t_{1/2} = 122 s

**D**

![Images of mCherry-N under DMSO, CVL218, and PJ34](Image)

**E**

![Bar graph of Number of puncta](Image)

- 1% DMSO
- CVL218
- PJ34

**F**

![Merge images of nsp12 + RNA](Image)

**G**

![Images of mEGFP-N Cy5 under DMSO, CVL218, and PJ34](Image)

**H**

![Images of mEGFP-N Cy5 under CVL218 and PJ34](Image)

**I**

![Images of nsp12 + RNA under CVL218 and PJ34](Image)

**J**

![Bar graph of Average diameter of condensates (μm)](Image)

- DMSO
- CVL218
- PJ34
- nsp12

**K**

![Graph of Percentage of recovery vs. Time](Image)

**L**

![Images of Bleaching prebleach and 0 min to 12 min](Image)

**M**

![Diagram showing nucleocapsid protein](Image)

© The Author(s) 2021
Surface plasmon resonance (SPR) assays of CVL218 (A) and PJ34 (B) to the immobilized full-length SARS-CoV-2-N. Top, the chemical structures of CVL218 and PJ34, respectively. Bottom, SPR binding curves of SARS-CoV-2 N to CVL218 and PJ34, respectively. (C) In vitro FRAP analysis of droplets (n = 3) formed by mEGFP-N protein with viral RNA (mEGFP-N, 3 μmol/L; RNA, 3 μmol/L) under the treatment of 20 μmol/L DMSO, CVL218 and PJ34, respectively. Data are representative of three independent experiments and presented as mean ± SD. (D) Droplet formation of the expressed mCherry-N (SARS-CoV-2 N tagged with mCherry) in Vero E6 cells after 48h transfection under three different treatments (20 μmol/L DMSO, CVL218 or PJ34, respectively). Scale bar, 10 μm. (E) Quantitative comparison of the numbers of droplets presented in (D). In total 11 transfected cells were considered in each treatment. Data are shown as mean ± SD and P values were determined by one-way analysis of variance (ANOVA) with Dunnett’s multiple comparison test, **P < 0.01, ***P < 0.001. (F) In vitro phase separation assays for 3 μmol/L Alexa546 labeled nsp12 alone, with 3 μmol/L mEGFP-N (SARS-CoV-2 N tagged with mEGFP, green) and with the complex of 3 μmol/L mEGFP-N and 3 μmol/L Cy5-labeled viral RNA (red), respectively. The molar ratio between mEGFP-N and RNA was 1:1. Scale bar, 2 μm. (G–I) In vitro phase separation assays of 3 μmol/L mEGFP-N protein with 3 μmol/L viral RNA in the absence and presence of 3 μmol/L nsp12 under the treatment of 20 μmol/L DMSO (G), CVL218 (H) and PJ34 (I), respectively. Scale bar, 2 μm. (J) Quantification of the effect of CVL218 or PJ34 treatment on the average diameters of the condensates presented in (G–I). The diameters were measured from the fluorescence microscopy images and shown as mean ± SD over three independent experiments. P values were determined by one-way analysis of variance (ANOVA) with Tukey’s multiple comparison test, **P < 0.01, ***P < 0.001, ****P < 0.0001. (K) In vitro FRAP analysis of the condensates (n = 3) of SARS-CoV-2 N-RNA-nsp12 complex (mEGFP-N, 3 μmol/L; RNA, 3 μmol/L; nsp12, 3 μmol/L) under the treatment of 20 μmol/L DMSO, CVL218 and PJ34, respectively. Data are representative of three independent experiments and presented as mean ± SD. (L) Representative snapshots of the condensates before and after bleaching presented in (K). Scale bar, 2 μm. (M) A model mechanism of the inhibition of viral replication and transcription of SARS-CoV-2 by small molecules in a phase separation dependent manner. Nsp12 alone cannot undergo phase separation in vitro, but it can be recruited into the droplets of SARS-CoV-2 N-RNA complex, despite the fact that nsp12 and viral RNA can form solid-state condensates. Comparing to those with DMSO treatment, the diameters and mobility of SARS-CoV-2 N-RNA-nsp12 droplets obviously increased after the treatment of CVL218/PJ34, which can attenuate the local density of the condensates and thus promote the entrance of other antiviral drugs (e.g., remdesivir) into their targets (e.g., nsp12/RdRp). The reaction buffer of in vitro phase separation assays in (F–L) consists of 20 mmol/L HEPES, pH 7.4, 150 mmol/L NaCl, 5% glycerol, 1 mmol/L EGTA and 1 mmol/L MgCl₂.
alone to 2.93 μmol/L in mixture. This result indicated that the combination of CVL218 and remdesivir can enhance the therapeutic efficacy of individual drugs alone against SARS-CoV-2, which thus can partially support the model mechanism of antiviral activities against SARS-CoV-2 for the reported drugs in a phase separation manner (Fig. 2M).

In summary, our results indicated that the N protein-driven LLPS is a promising target for the design of antiviral drugs, and the deep understanding of the functional roles of N protein in regulating the accessibility of RTCs will thus advance the development of anti-SARS-CoV-2 therapies.

**FOOTNOTES**

This work was supported in part by the National Natural Science Foundation of China (61872216 and 81630103 to JZ, 31900862 to DZ, 31871443 to PL), the National Key R&D Program (2019YFA0508403 to PL, 2020YFA0803300 to HL), the Turing AI Institute of Nanjing and the Zhongguancun Huihua Institute for Frontier Information Technology.

The data supporting the results of this article are included within the article and its supplementary files.

Dan Zhao, Weifan Xu, Pilong Li and Jianyang Zeng conceived the research project. Pilong Li and Jianyang Zeng supervised the study. Dan Zhao, Weifan Xu and Xiaofan Zhang designed and performed experiments, and analyzed the data. Xiaoting Wang, Shenyang Wu and Xiaolong Feng performed protein expression and purification in vitro. Yiyue Ge, Jingxin Li and Fengcai Zhu performed inhibitor combination assay. Enming Yuan, Yuanpeng Xiong and Peng Lang analyzed the SARS-CoV-2 genomic data. Nian Wu and Haitao Shu analyzed the structural data of SARS-CoV-2-N. Xiaokun Shen supplied PJ34 and CVL218 and provided discussions on the SPR results. Haitao Li participated in the project discussion and provided suggestions. Dan Zhan, Weifan Xu, Xiaofan Zhang, Pilong Li and Jianyang Zeng wrote the manuscript with support from all authors.

We thank Dr. Tingting Li for helpful discussions on phase separation. We acknowledge the assistance of Protein Preparation and Identification Facility (Technology Center for Protein Science, Tsinghua University) for protein expression and purification, High Throughput Screening (HTS) Core Facility for SPR experimental guidance (Center of Pharmaceutical Technology, Tsinghua University) and SLSTU-Nikon Biological Imaging Center (Center of Pharmaceutical Technology, Tsinghua University) for imaging support.

Dan Zhao1, Weifan Xu2,3, Xiaofan Zhang1, Xiaoting Wang4, Yiyue Ge5, Enming Yuan1, Yuanpeng Xiong6, Shenyang Wu7, Shuya Li1, Nian Wu1, Tingzhong Tian1, Xiaolong Feng8, Haitao Shu1, Peng Lang1, Jingxin Li5, Fengcai Zhu5,9, Xiaokun Shen10, Haitao Li3,11, Pilong Li2,3,10, Jianyang Zeng10

1 Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China
2 Beijing Advanced Innovation Center for Structural Biology Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing 100084, China
3 Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China
4 Silexon AI Technology Co., Ltd, Nanjing 210033, China
5 NHC Key Laboratory of Enteric Pathogenic Microbiology, Jiangsu Provincial Center for Diseases Control and Prevention, Nanjing 210009, China
6 Bioinformatics Division, BNRIST/Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
7 Protein Preparation and Identification Facility, Technology Center for Protein Science, Tsinghua University, Beijing 100084, China
8 Institute of Pathology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430074, China
9 Center for Global Health, Nanjing Medical University, Nanjing 210009, China
10 Convalife (Shanghai) Co., Ltd, Shanghai 210203, China
11 Ministry of Education Key Laboratory of Protein Sciences, Beijing Advanced Innovation Center for Structural Biology, Beijing Frontier Research Center for Biological Structure, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, China

Correspondence: pilongli@mail.tsinghua.edu.cn (P. Li), zengjy321@tsinghua.edu.cn (J. Zeng)

Accepted January 27, 2021

**OPEN ACCESS**

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**REFERENCE**

Carlson, C.R., Asfaha, J.B., Ghent, C.M., Howard, C.J., Hartooni, N., and Morgan, D.O. (2020). Phosphorylation modulates liquid-liquid phase separation of the SARS-CoV-2 N protein. bioRxiv.

Chang CK, Hou MH, Chang CF, Hsiao CD, Huang TH (2014) The SARS coronavirus nucleocapsid protein—forms and functions. Antiviral Res 103:39–50

Cong, Y.Y., Ulasl, M., Schepers, H., Mauhle, M., Viko, P., Kriegenburg, F., Thiel, V., de Haan, C.A.M., and Reggiori, F. (2020). Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle. J Virol 94.
Ge YY, Tian TZ, Huang SL, Wan FP, Li JX, Li SY, Yang H, Hong LX, Wu N, Yuan EM et al (2020) A data-driven drug repositioning framework discovered a potential therapeutic agent targeting COVID-19. bioRxiv.
Hurst KR, Koetzner CA, Masters PS (2013) Characterization of a critical interaction between the coronavirus nucleocapsid protein and nonstructural protein 3 of the viral replicate-transcriptase complex. J Virol 87:9159–9172
Iserman C, Roden C, Boerneke M, Sealfon R, McLaughlin G, Jungreis I, Park C, Boppana A, Fritch E, Hou YJ et al (2020) Specific viral RNA drives the SARS CoV-2 nucleocapsid to phase separate. bioRxiv.
Klein IA, Boija A, Afeyan LK, Hawken SW, Fan M, Dall’Agnese A, Oksuz O, Henninger JE, Shrinivas K, Sabari BR et al (2020) Partitioning of cancer therapeutics in nuclear condensates. Science 368:1386–1392
Lin SY, Liu CL, Chang YM, Zhao J, Perlman S, Hou MH (2014) Structural basis for the identification of the N-terminal domain of coronavirus nucleocapsid protein as an antiviral target. J Med Chem 57:2247–2257
Malik YA (2020) Properties of Coronavirus and SARS-CoV-2. Malays J Pathol 42:3–11
Nelson GW, Stohlman SA, Tahara SM (2000) High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. J General Virol 81:181–188
Perdikari, T.M., Murthy, A.C., Ryan, V.H., Watters, S., Naik, M.T., and Fawzi, N.L. (2020). SARS-CoV-2 nucleocapsid protein undergoes liquid-liquid phase separation stimulated by RNA and partitions into phases of human ribonucleoproteins. bioRxiv.
Savastano, A., Opakua, A.L., Rankovic, M., Zweckstetterc, M. (2020). Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. bioRxiv.
Stohlman SA, Baric RS, Nelson GN, Soe LH, Welter LM, Deans RJ (1988) Specific interaction between coronavirus leader RNA and nucleocapsid protein. J Virol 62:4288–4295
Surjit M, Lal SK (2008) The SARS-CoV nucleocapsid protein: a protein with multifarious activities. Infect Genet Evol 8:397–405

Supplementary Information The online version of this article (https://doi.org/10.1007/s13238-021-00832-z) contains supplementary material, which is available to authorized users.