Journal of Structural Biology 217 (2025) 108162

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Research Article

Characterization of SARS-CoV-2 nucleocapsid protein oligomers

Domenica Farci [a,b,c,d,1,*], André T. Graça [b,1], Michael Hall [b], Patrycja Haniewicz [a], Sami Kereïche [e], Peter Faull [f,g], Joanna Kirkpatrick [f], Enzo Tramontano [h], Wolfgang P. Schröder [b], Dario Piano [a,c,d,1,*]

[a] Department of Plant Physiology, Institute of Biology, Warsaw University of Life Sciences - SGGW, Warsaw, Poland
[b] Department of Chemistry, Umeå University, Umeå, Sweden
[c] Laboratory of Plant Physiology and Photobiology, Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy
[d] ReGenFix Laboratories, R&D Department, Sardara, Italy
[e] Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, Czech Republic
[f] The Francis Crick Institute, London, United Kingdom
[g] Proteomics Facility, University of Texas at Austin, Austin, USA
[h] Laboratory of Molecular Virology, Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy

ARTICLE INFO

Keywords:
SARS-CoV-2
Covid-19
Self-cleavage
Nucleocapsid protein
Cryo-electron microscopy
Proprotein convertases
Furin

ABSTRACT

Oligomers of the SARS-CoV-2 nucleocapsid (N) protein are characterized by pronounced instability resulting in fast degradation. This property likely relates to two contrasting behaviors of the N protein: genome stabilization through a compact nucleocapsid during cell evasion and genome release by nucleocapsid disassembling during infection. *In vivo*, the N protein forms rounded complexes of high molecular mass from its interaction with the viral genome. To study the N protein and understand its instability, we analyzed degradation profiles under different conditions by size-exclusion chromatography and characterized samples by mass spectrometry and cryo-electron microscopy. We identified self-cleavage properties of the N protein based on specific Proprotein convertases activities, with Cl^- playing a key role in modulating stability and degradation. These findings allowed isolation of a stable oligomeric complex of N, for which we report the 3D structure at ~6.8 Å resolution. Findings are discussed considering available knowledge about the coronaviruses' infection cycle.

1. Introduction

Nucleocapsid proteins are essential components of the viral particle, as they are deputed to genome protection and virion structural stabilization (Muriaux and Darlix, 2010; McBride et al., 2014). In the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), the nucleocapsid (N) protein typically occurs in globular units of high molecular weight suggesting that it assembles with a high oligomerization index *in vivo* (Klein et al., 2020; Yao et al., 2020; Turoňová et al., 2020). These oligomers might represent a genome-packing building block expected to behave as a self-assembling structural unit, a typical feature of nucleocapsid proteins (Ye et al., 2020). This self-assembling process is thought to be promoted by the protein-genome interactions exploiting the high affinity of nucleocapsid proteins for nucleic acids (Chen et al., 2007; Ye et al., 2020).

Structurally, the N protein has two main domains: the N-terminal domain (NTD), which binds RNA and recognizes specific sequences or structures within the viral genome, and the C-terminal domain (CTD), primarily responsible for dimerization (Morse et al., 2023). The NTD has a globular structure with a positively charged surface that facilitates interaction with the negatively charged RNA, while the CTD is highly flexible, allowing N proteins to form dimers and subsequently higher-order oligomers essential for genome packaging. Moreover, the N protein has large intrinsically disordered regions (IDRs) flanking the NTD and CTD (Cubuk et al., 2021). These flexible regions allow for dynamic interactions with other viral or host proteins, providing adaptability and resilience in different cellular environments (Wu et al., 2023).

This combination of well-ordered domains (NTD and CTD) and disordered regions (IDRs) gives the N protein a high degree of flexibility, which is functionally advantageous for RNA interaction and host factor binding. However, this flexibility presents challenges for obtaining high-resolution structural data for the full-length protein and its functional

* Corresponding authors at: Department of Plant Physiology, Institute of Biology, Warsaw University of Life Sciences - SGGW, Warsaw, Poland.
E-mail addresses: domenica.farci@unica.it (D. Farci), dario.piano@unica.it (D. Piano).
[1] Shared co-first authorship.

https://doi.org/10.1016/j.jsb.2024.108162
Received 25 August 2024; Received in revised form 5 November 2024; Accepted 12 December 2024
Available online 13 December 2024
1047-8477/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

oligomers, as the disordered regions tend to form variable structures in solution, complicating crystallization and 2D/3D classification in crystallography or cryo-Electron Microscopy (cryo-EM) studies, respectively.

While engaged with the genome by forming the nucleocapsid, the N protein further contributes to the virion stabilization by specific protein–protein interactions with the capsid and the matrix (He et al., 2004). This enhances the nucleocapsid cohesion and has a further role in the static stabilization of the whole virion (Muriaux and Darlix, 2010; McBride et al., 2014). During infection, after the receptor-mediated adhesion to the cell surface of the host, the virion is embedded into a membrane vesicle gaining entrance into the cell (Wang et al., 2004; Shang et al., 2020). This process of fusion is followed by the disaggregation of the vesicle and the release of the virion into the cytosol (Wang et al., 2004; Shang et al., 2020). The intracellular release is a critical step, during which the vesicle dismantlement takes place and the whole capsid falls apart into its components. While our understanding of this process is sufficiently elucidated with respect to the capsid, especially considering the common fate with the membrane's vesicle, the disaggregation of the nucleocapsid remains a puzzling process, particularly when considering its tight packing. The compactness of the nucleocapsid structure suggests the presence of specific chemical-physical signals that might trigger the disassembling to allow the subsequent steps of replication.

As common in SARS-CoV (Wang et al., 2004; Huang et al., 2009), also the N protein from SARS-CoV-2 is not only characterized by pronounced instability leading to consequent protein degradation but also exhibits an intrinsic tendency to aggregate, often forming higher-order oligomers (Perdikari et al., 2020). These properties have significantly limited structural studies of this protein in its full-length form (Kang et al., 2020; Zinzula et al., 2021; Jia et al., 2022) and precluded further investigations into its possible oligomeric states.

To understand the bases for the N protein instability, we characterized the degradation profiles under different conditions, identifying a specific correlation of this instability with the presence of Chloride (Cl^-). Further studies based on mass spectrometry confirmed the protein identity and excluded the presence of potential contaminants that might have been responsible for its degradation. The results shown here suggest the N protein having Cl^--related self-cleavage properties based on specific Proprotein convertases activities. Proprotein convertases are a group of enzymes identified for cleaving precursor proteins at specific sites activating or modifying protein function (Seidah and Prat, 2012). Specifically, in chemical-physical conditions in which Cl^- is absent, the degradation activity is drastically reduced, and the N protein occurs in a dominant state with a high oligomerization index.

In vivo, the N protein forms rounded complexes resulting from its interaction with the viral genome (Klein et al., 2020; Yao et al., 2020). In this work, micrographs revealed the oligomeric state of the assembly occurring in absence of Cl^-, resembling the *in vivo* structure, with a characteristic rounded shape and an approximate diameter of ~160 Å. We have structurally characterized this complex by cryo-electron microscopy single-particle analysis (cryo-EM SPA) to obtain its intermediate-resolution 3D structure, which is further supported by cryo-electron tomography (cryo-ET) on intact virions. With respect to cryo-ET, our findings align with previous cryo-ET studies (Klein et al., 2020; Yao et al., 2020). Results show new insights on the properties of the N protein and have implications with respect to the viral infection cycle, also providing essential hints for biomedical applications.

2. Materials and methods

2.1. Heterologous expression and cell culturing

Escherichia coli (BL21) competent cells were transformed with a pET21a plasmid containing the gene for a full-length Nucleocapsid protein with 6x Hi-tagged in the C-terminal (GMV-V-2019nCoV-029, Genemedi). Pre-cultures were grown in 50 mL Luria Bertani (LB) media at 25 °C for 12 h, and then transferred to 1.5 L flask until reaching the optical density ($OD_{600\ nm}$) of 0.6. Protein expression was then induced by the addition of 1 mM Isopropyl-β-D-thiogalactopyranoside (IPTG) and the transformed cells were grown in LB media at 37 °C for 24 h under shaking.

2.2. Nucleocapsid protein isolation

Bacterial cells were harvested by centrifugation (5000*x g*, 10 min at 4 °C), resuspended in a 50 mM Phosphate Buffer pH 7.5 (Buffer A − BA) added with 100 U DNase I (DNase I recombinant, RNase-free Roche), and disrupted using a French Pressure Cell (3x 1100 psi). Intact cells were removed by low-speed centrifugation (5000*x g*, 10 min at 4 °C) and the final cytosolic and membrane fractions were separated by a high-speed centrifugation step (48000*x g*, 10 min at 4 °C). Considering that the larger amount of protein was found associated to the membrane fraction, most likely due to its natural interaction with biological membranes required during viral replication and assembly, we performed the N protein isolation from the membrane fraction. The pellet obtained was then resuspended in BA and subject to mild solubilization at room temperature for 60 min under slow stirring with 10 % (w/v) n-dodecyl-β-D-maltoside (β-DDM) at a final detergent concentration of ~1.4 % and a protein concentration of ~20 mg/ml. The sample was then centrifuged (48000*x g*, 10 min at 4 °C) and the fraction in solution was injected into a NiNTA column (5 mL HisTrap HP, GE, Healthcare), washed for 25 column volumes at 1 mL/min with 50 mM Tris-Acetate buffer pH 7.5 (Buffer B − BB) and eluted at the same flow with 50 mM Tris-Acetate buffer pH 7.5, 300 mM Imidazole (Buffer C − BC). The eluted protein fraction was concentrated in protein concentrators of 100 kDa cutoff (Vivaspin 20, Sartorius) then subjected to a step of Size Exclusion Chromatography (SEC) with a Superose 6 column (10/300 GL, GE, Healthcare) in BB at a flow rate of 0.5 mL/min. This step allowed to remove the imidazole and the presence of any contaminant. The resulting SEC pool, corresponding to the N oligomer, was concentrated (as described above) to either 0.2 mg/mL or 0.8 mg/mL protein to perform the experiments on diluted or concentrated N protein, respectively, and used for further SEC characterization experiments or stored at −80 °C after flash freezing in liquid nitrogen. In these studies, all chromatography columns were subjected to the ReGenFix procedure (https://www.regenfix.eu/) for regeneration and calibration before every run.

2.3. Denaturing polyacrylamide gel electrophoresis (PAGE)

Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) with a 10 % (w/v) separating and a 4 % (w/v) stacking gel was used (Farci et al., 2022). Isolated cell fraction, NiNTA pool, and SEC pool were denatured with RotiLoad (Roth), boiled for 5 min, and centrifuged for 5′ at room temperature before loading. After the electrophoretic separation, the gels were stained with Coomassie Brilliant Blue G-250 to allow visualization.

2.4. Mass spectrometry

The gel bands for the SEC pool were excised from SDS-PAGEs and digested with trypsin prior to further processing according to (Farci et al., 2022). LC-MS/MS using an Ultimate 3000 nano UHPLC system (Thermo Fisher Scientific) coupled to an Orbitrap Fusion Lumos (Thermo Fisher Scientific) mass spectrometer was used to analyse the extracted peptides. Data analysis was performed with MaxQuant software (version 1.5.3.28). The obtained data were searched using the Andromeda search engine (Cox et al., 2011) against FASTA databases for Uniprot entries for *E. coli,* SARS-CoV-2, and common contaminants. A false discovery rate of 1 % was used at both peptide and protein level. To determine the most abundant peptides, intensity-based absolute

quantification (iBAQ) values were used.

2.5. Bioinformatic analyses

The N sequence was screened for post-translational modifications using the tools available on the DTU Health Tech web platform, specifically the NetNGlyc 1.0 (Gupta and Brunak, 2002) (https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/) and the ProP 1.0 (Duckert et al., 2004) (https://services.healthtech.dtu.dk/services/ProP-1.0/).

2.6. Sample preparation for cryo-electron microscopy

Grids preparation and data acquisition have been done at Umeå Core Facility for Electron Microscopy (UCEM) and Swedish National Cryo-EM facility, Umeå, Sweden.

Considering the protein's pronounced instability and intrinsic tendency to aggregate, during grid preparation we tested several conditions (e.g., grid's type, protein concentration) to reduce these tendencies. Quantifoil 2/2 + 2 nm C holey carbon grids were glow-discharged prior to use. Purified N protein samples were blotted and vitrified using a Vitrobot plunge-freezing machine (Mark IV, Thermo Fisher Scientific) at 4 °C and 100 % humidity (blot force 0, blotting time 5 s, wait time 1 s, drain time 0 s), and placed in autogrid prior to image acquisition.

Datasets were acquired with a Titan-Krios TEM (Thermo Fisher Scientific) at an operating voltage of 300 kV and equipped with a Cs-corrector (cs 2.7 mm), a Quantum GIF energy filter (slit width set to 20 eV), and a post-GIF K2 camera (Gatan) in counting mode. Images were automatically recorded by the EPU software v2.8.0 at a nominal magnification of 165000×, yielding a final image pixel size of 0.82 Å. Image defocus was set in EPU to vary between −1.5 μm and −3.0 μm, the total electron dose used to acquire a single image was ~59.5 electrons/Å^2 (dose per fraction was ~1.5 electrons/Å^2).

2.7. Single-particle data processing

The cryoSPARC software v.3.2.2 (Punjani et al., 2017) was used during the whole data processing workflow as described below. First, 8636 movies (.tif files; 0.82 Å, 300 kV, 2.7 mm and total exposure dose 59.5 e/Å^2) were imported, patch motion corrected, and the contrast transfer function (CTF) was estimated for each micrograph. Micrographs with extensive motion, large ice thickness, and poor CTF fit were excluded remaining 8459 micrographs. Particles were picked using an automatic blob picker (circular blobs with 120–180 Å diameter; minimum distance separation of 0.5 diameters). Inspect particle picks job to perform to exclude over-picked particles such as ice contaminants or low signal-to-noise-ration (SNR) picks. The 2847315 picked particles were extracted using a box size of 512 pixels, binned 4x (to 128 pix Fourier crop size). One round of 2D classification with 30 classes (0.6 window inner radius and 0.7 window outer radius) was used to exclude obvious junk particles, cutting down to 1637450 particles used on further processing steps. Next, 12 rounds of 2D classification were performed separating particles into 200 different classes with a window inner size of 0.6 radius and 0.7 window outer radius. Next, 399714 particles were re-extracted at full box size (512px). 3D classification was performed generating 200 different volumes out of 278626 particles. All 200 volumes were superimposed for comparison. 52 out of the 200 3D classes (= 95686 particles) contained the same structural features at low resolution and were therefore segregated into a single group for further steps. *Ab-initio* 3D reconstruction job was performed to obtain an initial common 3D volume to all 95686 particles. The *ab-initio* volume shared similar features to the 3D classes previously analyzed. Non-uniform refinement on the same particle stack yielded a first 3D volume at ~6.8 Å resolution (EMD-17430). The resolution has been estimated according to the 0.143 Gold-standard Fourier shell correlation (FSC) criteria. Further data processing was performed by local refinement dividing into four main subregions of the obtained 3D volume. No symmetry was imposed. Unfortunately, the final 3D model did not show any significant improvement (EMD-16868), confirming the many conformational states and high dynamic of the protein, therefore this map was not used for further analyses and not mentioned in any part of the results. Isosurface visualization was accomplished using the Chimera software (Pettersen et al., 2004). A scheme for the cryo-EM analysis has been provided in Sup. Fig. 1a.

2.8. Cryo-electron tomography data processing

Raw tilt series of SARS-CoV-2 virions were obtained from the EMPIAR database (*5*; https://doi.org/10.6019/EMPIAR-10453) and processed by the *etomo* software (imod and PEET) (Cope et al., 2011). Gold particles were used as fiducial markers and only tomograms with minimal alignment errors (between 0.1 and 0.2 nm) were selected for the final reconstruction, which was calculated using weighted back-projection algorithms (Cope et al., 2011).

For subtomogram averaging on 3D reconstructions, sub-tomograms with a size of 256^3 voxels were manually selected and extracted from regions corresponding to the N protein inside the virion particle. Subtomogram alignment and averaging with missing-wedge compensation was performed with the PEET software without imposing any symmetry. Absolute values of cross-correlation were used for alignment potentially helping to prevent noise by reinforcing to match features in the reference. Isosurface visualization was accomplished using the Chimera software (Pettersen et al., 2004). A scheme for the cryo-EM analysis has been provided in Sup. Fig. 1b.

3. Results

3.1. In absence of Cl^- the stability of N increases leading to a dominant oligomer

The full-length N protein of SARS-CoV viruses is intrinsically unstable (Wang et al., 2004; Huang et al., 2009; Zinzula et al., 2021), a property that makes it undergoing degradation and to aggregate into higher-order oligomers (Perdikari et al., 2020) in relatively short time. This fact has severely limited the *in vitro* studies only to its truncated forms (Kang et al., 2020; Zinzula et al., 2021; Jia et al., 2022) so that, so far, a structure of the full-length N is still missing. These properties also precluded almost completely the characterization of the possible N oligomeric states. We performed the isolation of the full-length N protein under the conditions reported in literature (Kang et al., 2020; Zinzula et al., 2021; Jia et al., 2022). Accordingly, most of those conditions led to heterogeneous samples with extended degradations occurring within a few hours after protein elution. To these features, persistent oligomerization/aggregation properties, even under denaturing conditions, were also observed (Fig. 1a).

Aiming at overcoming these problems, we developed a procedure for stabilizing the N protein possibly preserving its supramolecular organization. A batch purification screening let to assess the effect of independent multiple conditions on the protein, identifying a critical role played by Cl^- (Sup. Table 1) and allowing the selection of Cl^--free buffer conditions based either on Phosphate Buffer or Tris-Acetate (Sup. Fig. 2). On the selected conditions, the protein concentration promoted degradation in Phosphate Buffer, while in Tris-Acetate this property was significantly limited hence it was chosen for further steps of characterization (Sup. Fig. 2). Samples resulting from Tris-Acetate isolation, when characterized by denaturing electrophoresis, resolved in a reduced amount of oligomers and a main monomeric band instead of the two observed for the conditions above (Fig. 1b vs Fig. 1a). In Tris-Acetate, the N protein appeared significantly more stable, and its degradation was less pronounced showing a chromatographic profile with a dominant form carrying a high oligomerization index likely influenced by the length of the possible bound nucleic acid (the elution peak, centered at ~8.8 mL, is equivalent to an apparent mass of ~800 kDa). Interestingly,

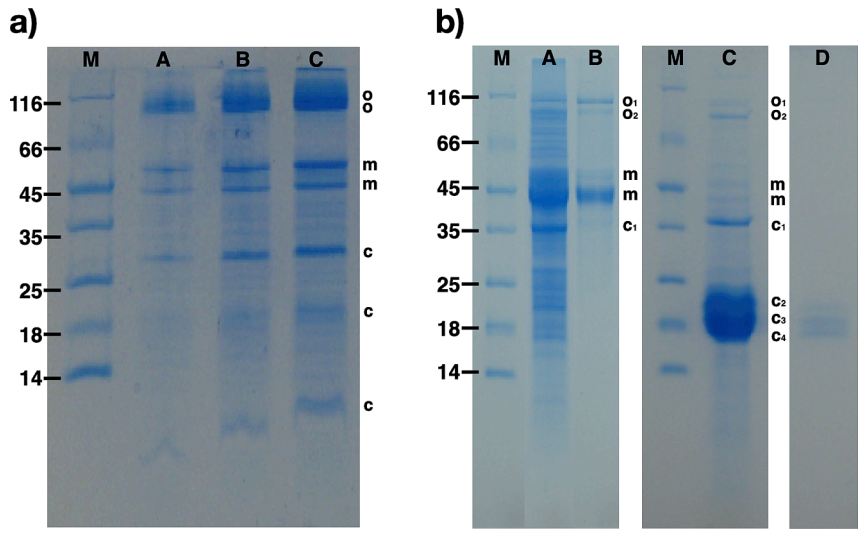


Fig. 1. Stability profiles of the N protein under different isolation conditions: In the image a) under the main conditions reported in literature, and, in particular, in the presence of Cl^-, the protein pool is composed by monomers (m), several degradation bands (c), and higher oligomers (o). In the lanes A, B, and C are 1x, 2x, and 3x loaded samples, respectively. The image b) shows the optimized isolation of N pool in Cl^--free Tris-Acetate buffer at the first (Ni-NTA – lane A) and the second (SEC – lane B) chromatography step. In lanes C and D is indicated the exhaustive degradation of the N protein after 3 days of incubation at 4 °C (10x and 1x loaded samples, respectively). The lanes M are the molecular marker. With m is indicated the monomer (full length) of N; with c_1, c_2, c_3 and c_4 are indicated the degradation fragments in decreasing size order; with o_1 and o_2 are indicated the oligomerization bands likely resulting from a strong interaction of the full length (m) or the heavier degradation band (c_1), respectively.

this condition also showed a significant void volume peak indicating possible high molecular assemblies occurring between oligomeric particles. This isolation procedure led to a minimal presence of monomers and degradations (Fig. 2a).

However, even in the absence of Cl^-, the isolated protein pools reached partial (diluted protein in Tris-Acetate) or exhaustive proteolysis (concentrated protein in Tris-Acetate) after 72 h at 4 °C (Fig. 2b and c). The exhaustive proteolysis was further evidenced by denaturing electrophoresis (Fig. 1b, lanes C and D). Moreover, the same degradation profiles occurred regardless of the presence of protease inhibitors

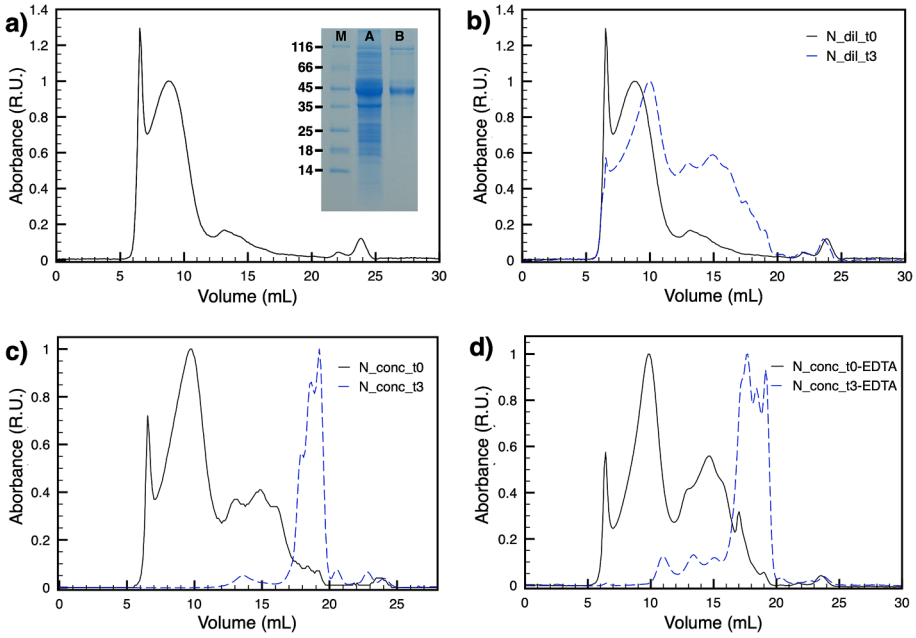


Fig. 2. The N protein is destabilized by time and concentration: The image shows the SEC profiles of the N pool after a first Ni-NTA step in Tris-Acetate and Cl^--free conditions. In a) the main part of the protein pool separates in a peak centered at ~8.8 mL (apparent mass ~800 kDa) preceded by a significant void volume component indicating possible assemblies/aggregates of particles. The inset shows a gel of the injected pool (A), the resulting SEC pool (B, elution peak centered at ~8.8 mL) that was also subjected to cryo-EM analysis, and the molecular marker (M). In b) the protein pool as obtained in (a) has been reinjected immediately (black-continuous line) or after 72 h at 4 °C (dashed-blue line). The latter profile shows an increased degradation and less interactions between oligomers. In c) the protein pool as obtained in (a) has been analysed under the same timing conditions as in (b) but concentrated about 10 times (see material and methods for details). The profile shows a much more pronounced degradation with respect to (b) that is visible at t_0 (black-continuous line) and goes to exhaustion after 72 h at 4 °C (t_3, dashed-blue line). Degradation as in (b) and (c) occurred also in the presence of protease inhibitors cocktails (Roche, data not shown). In d) the protein pool has been analysed in the same conditions as in (c) but in presence of EDTA (t_0 black-continuous line; t_3 dashed-blue line) showing a minimal effect as a chelator. All profiles and isolations performed for these experiments are Cl^- free in Tris-Acetate buffer.

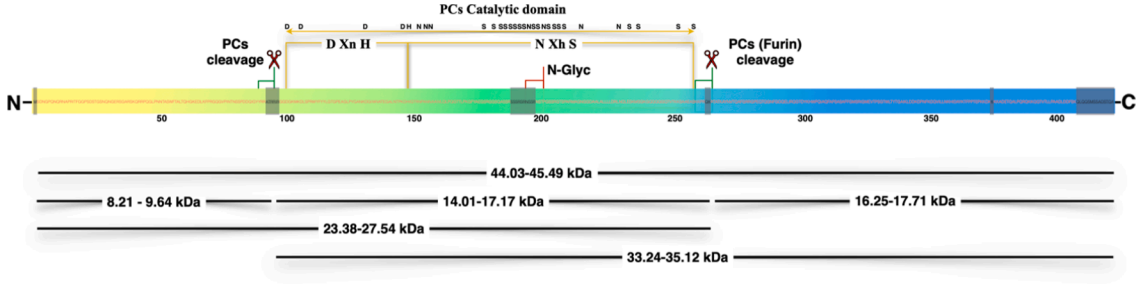
cocktails (Roche) in either diluted of concentrated samples (data not shown). Finally, EDTA had no influence on the oligomeric profiles, allowing to exclude the need of metal ions for the assembly formation (Fig. 2d).

3.2. The N protein shows a self-cleavage degradation profile

Based on the above observations, the N protein undergoes several steps of proteolysis. The full-length form (~45 kDa) is first cleaved to a main ~36 kDa fragment which is further split into three fragments of ~24, ~17, ~15 kDa (Fig. 1b bands c_1, c_2, c_3, and c_4, respectively). Moreover, the ~45 kDa full-length form and the ~36 kDa fragment are both subjected to a stable dimerization that partially persists to denaturing conditions and is likely more significant in native conditions, as shown by two less representative bands at apparent masses of ~90 and ~70 kDa (Fig. 1b bands o_1 and o_2, respectively). Degradation profiles were characterized by Mass Spectrometry (MS) to confirm the identity of the protein and of its fragments, as well as the possible presence of proteases. This analysis assigned to the N protein all the bands resolved by SDS-PAGE, with no potential contaminants responsible for the observed degradation (Table 1). Therefore, these results suggested possible self-cleavage properties associated with the N protein that, according to the data shown in the previous paragraph, are further enhanced by Cl^-. We next investigated the characteristic parts of the N sequence that were not covered by the MS analysis (see scheme in Table 1 and Sup. Table 2). This approach led to the identification of two regions, one at the C-terminal and the other at the N-terminal side, characterized by boxes with multiple Lysine (K) or/and Arginine (R) residues, and a third central region localized between the first two and rich in Serines (S) (see scheme in Table 1). Considering that the molecular weights of the fragments resulting from a possible cleavage in the two K/R-rich boxes are compatible with those of the experimental fragments observed by the SDS-PAGE (Fig. 1b), we investigated whether these two regions could be specific protease cleavage sites. The two sites rich in K and R were found to be highly specific for the cleavage of Proprotein convertases (PCs) (Seidah and Prat, 2012; Garten, 2018), a possibility also supported by the positive influence of Cl^- on the cleavage (Seidah and Prat, 2012; Garten, 2018). The signal sequences for the two regions are YRRATRR (with a score of 0.507 on a threshold of 0.5) and KKPRQKR (with a score of 0.815 on a threshold of 0.5), in which the terminal R, residues 93 and 262 respectively, correspond to the actual cleavage sites. Furthermore, the cleavage site in position 262 was identified as specific for Furin cleavage, a main representative and the most characterized of the PCs family. Next, to identify the possible meaning of the third-central S-rich region, localized between the two cleavage sites (see scheme in Table 1), we investigated a masking effect due to N-glycosylation, which would make this region not easily accessible to the MS analysis. Prediction analyses for N-glycosylation hint at the sequence NSSRNST (residues 192–198 localized at the end of the S-rich region) as a possible site of N-glycosylation in the residues 192 and 196 (Sup. Table 2). The S-rich region sits into a more extended and flexible one characterized by the presence of possible PCs key reactive residues found in the non-consecutive order DXnHXmNXhS (with n, m or h a number ≥ 0 of residues X), a characteristic that, together with N-glycosylation, is highly conserved in PCs (Garten, 2018). Among different possible sites in which the reactive residues can be located, there is one having consecutive residues D and H (residues 144 and 145). Similarly, also the residues N and S are found to be consecutive in two sites of the N-glycosylation region NSSRNST (residues 192–193 and residues 196–197, respectively). These properties make the region

Table 1

Scheme of the degradation pattern and MS results. The string from light to dark grey to is a schematic representation of the N protein sequence. In a darker background are indicated the regions not covered by the MS analysis and that are common between all the analysed bands. The two putative PCs cleavage sites are outlined by scissors. In the catalytic domain, the N-glycosylation site, the DH and NS regions crucial for the activity, and the position of the putative key residues are also indicated. The numbers from 50 to 400 give an indicative reference of the sequence regions. The line scheme below indicates the size range of the possible degradation products with respect to the N monomer. The associated table provides the MS details for every analysed band and shows the agreement between the calculated masses inferred from the theoretical degradation pattern and the apparent masses measured by electrophoresis. In the table, the letters m, c and o represent the monomer, the specific cleaved form and the specific bands of oligomerization, respectively, as also done in Fig. 1.



ID	Description	Acronym	Organism	Gel Band	Apparent Mass (kDa)	Calculated Mass averaged (kDa)	Coverage (%)	Unique peptides
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**m**	44,5	44,8	92	78
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**c_1**	35	34,2	77	42
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**c_2**	24	25,5	81	55
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**c_3**	17	17	74	44
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**c_4**	15	15,6	74	51
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**o_1**	90	–	–	–
P0DTC9	Nucleocapsid protein	N	Severe acute respiratory syndrome coronavirus 2	**o_2**	70	–	67	78

between the two K/R-rich boxes a potential PCs catalytic domain supporting the hypotheses about the self-cleavage activity of N (see scheme in Table 1).

3.3. N oligomers occur as a non-symmetric round particle

The oligomeric pool separated by size-exclusion chromatography (main peak centered at ~8.8 mL in Fig. 2a) occurred as relatively homogenous fields of pseudo-spheric particles when analysed by cryo-EM (Fig. 3), which is notable given the known protein's pronounced instability, leading to degradation and an inherent tendency to aggregate into higher-order oligomers (Perdikari et al., 2020). *In vivo*, the N protein was observed to form rounded complexes resulting from its interaction with the viral genome (Klein et al., 2020; Yao et al., 2020; Turoňová et al., 2020). Here, we investigated the N oligomeric structure by cryo-EM single-particle analysis. The data set, initially composed of ~3 million particles from a total of ~8500 movies, yielded ~96000 particles that allowed solving the N oligomer structure at ~6.8 Å resolution (Fig. 4). The N oligomer appears with a pseudo-spherical shape and a diameter of ~160 Å. The surface is rough and strongly anisotropic with regions differing in their degree of irregularity and porosity (Fig. 4). Several well-defined holes can be observed (Fig. 4, green highlights), in agreement with what already observed *in situ* at lower resolution (Klein et al., 2020; Yao et al., 2020), and among them two are found to be contiguous (Fig. 4) eventually opening the way into two internal cavities. Another characteristic feature is also represented by a pseudo-tubular protrusion representing a possible nucleic acid strand emerging from the particle (Fig. 4, red highlight). More of these formations emerge when observing the structure at lower sigma. Considering the knowledge from previous cryo-ET studies and our single-particle map, we attempted subtomogram averaging of N oligomers in intact virions. The final subtomogram averaging map, even if its lower resolution does not allow clear correspondence with the single-particle analysis, shows features such as a rough surface, holes, cavities as well as dimensions that are coherent with the ones observed by single-particle analysis (Sup. Fig. 3).

4. Discussion

In SARS-CoV-2 the Nucleocapsid protein is shown to be the most abundant during the viral infection cycle. Being the most crucial structural component providing cohesion, stability, and protection to the viral genome (Cong et al., 2020), the N protein is a characteristic antigen with extensively conserved regions (Caruso et al., 2022). Therefore, together with the Spike protein and the Envelope protein (Noori and Sardar, 2022; Cubisino et al., 2024), it has been a main target for clinical detection and pharmacological studies (Bai et al., 2021).

However, the N protein is also known for its physiological role in the viral infection cycle being essential as RNA-binding protein during nucleocapsid assembly and as a regulation factor for RNA metabolism (replication and transcription) (Cong et al., 2020). Because of these multiple functions and, in particular, of its involvement in the genome super-packing, the N protein of Coronaviruses is shown to be highly dynamic (Cubuk et al., 2021; Bai et al., 2021; Caruso et al., 2022) and prone to degradation (Wang et al., 2004; Huang et al., 2009). These properties limited the studies of full-length N proteins including their structural characterization. Consequently, high-resolution structural studies were performed only on its N-terminal or C-terminal truncate versions (Kang et al., 2020; Zinzula et al., 2021; Jia et al., 2022). Here, by analysing by MS the degradation pattern and the sequence coverage of each natural fragment, we identified two regions of the N protein that are putatively important elements of the instability and one region resembling a PCs active site. The first two regions, both rich in K and R residues, are ideal recognition sites for PCs, and one of these seems to be highly specific for Furin. The degradation pattern associated with the N protein and the MS analysis did not indicate the presence of protease contaminant in the sample, suggesting a possible self-cleavage activity. This observation was further confirmed by the presence of a third region, rich in S residues, carrying specific features for the PCs active site.

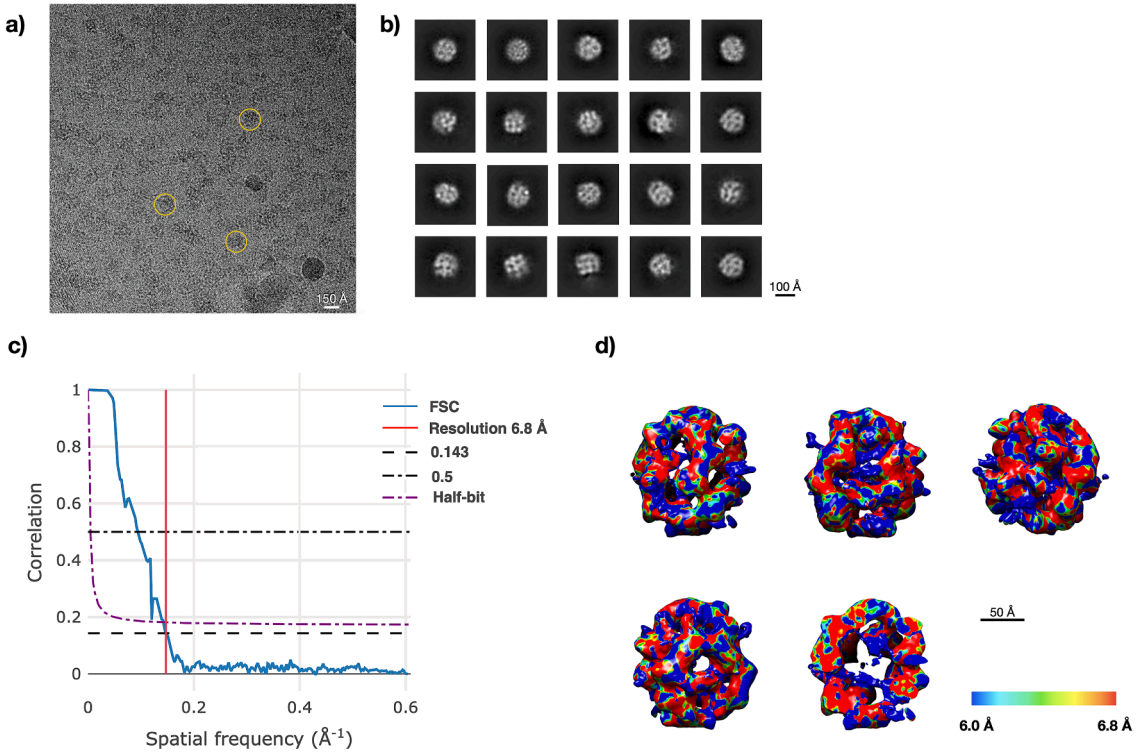


Fig. 3. Characterization of the N oligomers by cryo-EM. In a) a micrograph from a grid with a relatively homogeneous field of N particles is shown. A few typical particles are highlighted (yellow circles); in (b) the main representative 2D classes are shown; in c) and d) are shown the FSC plot and the local resolution map respectively. Scale bars are indicated.

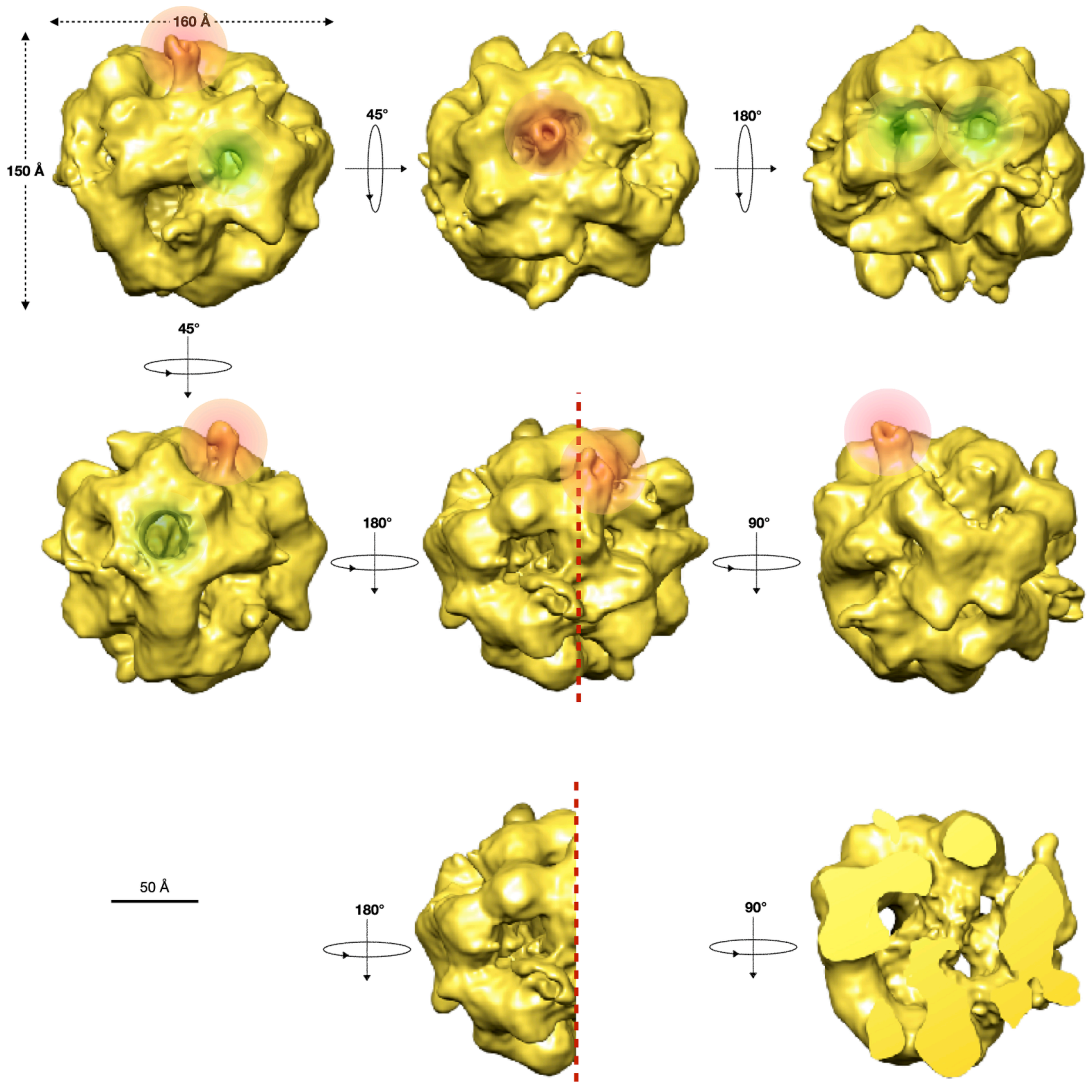


Fig. 4. Structure of N oligomers at intermediate resolution. The N assembly appears as a rounded particle with a rough and heterogeneous surface as show by different preferential views. The green highlights show the main details of the surface holes that provide access to the internal cavities and the red highlight shows a possible nucleic acid strand emerging from the particle. The scale bar is indicated.

Not coincidentally, this region is the highly dynamic and still not characterized; in fact, it is the one localized between the N- and the C-terminal fragments, the only extensively studied regions so far. The occurrence of degradation activities on viral proteins during the infection cycle is not unusual (Bessa et al., 2022; V'kovski et al., 2021), and the role of Furin in the degradation of the SARS-CoV-2 S protein has also been reported (Peacock et al., 2021). Furthermore, proteolytic events involving the N protein have been observed in human-derived cell lines (Lutomski et al., 2021), suggesting potential functional significance and eventually highlighting the need for a deeper understanding of this process. Accordingly, in the present study, the susceptibility of the N protein for Cl^- has been identified and described, showing the Cl^- to act on the protein strongly enhancing its otherwise slow degradation (Fig. 2). This property of Cl^- suggests both a triggering and a modulation mechanism aimed at regulating the degradation rate on the base of the Cl^- levels, which is a precise signal of localization and moment of the infection cycle. In the absence of Cl^- signal, the N protein has been found in a relatively stable and homogenous oligomeric "default" state. This state is characterized by a low rate of degradation and, in agreement with previous reports (Klein et al., 2020; Yao et al., 2020; Turoňová et al., 2020), by a characteristic rounded shape as independently shown here by cryo-EM single-particle analysis and tomography. However, while these particles on one side are here found to be highly reproducible giving homogeneous preparations, on the other they are characterized by intrinsic dynamics that did not allow obtaining high-resolution cryo-EM maps and the consequent structural modelling. The observed results suggest that the N dynamics might be at the base of precise mechanisms of stabilization/destabilization expected to be relevant *in vivo*. Through these complex transitions, the N protein is able to drive the passage between the high oligomeric state with structural function and the dimeric/monomeric/fragmented state with biochemical function, as typically occurs in different moments of the infection cycle. The first state aims at interacting with the genomic RNA protecting it through a tight packing and limiting its accessibility; this state contributes to building a viral particle able to contrast the environmental stressors during the extracellular part of the viral infection cycle. Contrarily, the second state aims at promoting the unpacking of the genome and the interactions with replication and transcription factors, while other N copies interact with different machineries of the host which are related to several aspects of the infection (Cong et al., 2020; Cubuk et al., 2021). In this context, the transition between the two states must be triggered by specific signals among which the Cl^- seems to play a pivotal role. According to these observations, previous reports showed the retention of Cl^- in the structure of the N protein C-terminal dimers

(Jia et al., 2022), supporting this form as the one occurring after Cl^- interaction on the oligomer and likely to be equivalent to one of the forms that we observed here under exhaustive degradation (Fig. 2).

5. Conclusions

The present findings provide evidence for the complex oligomeric transitions to which the N protein is subjected *in vivo*, passing from packed viral particle to virion disassembling and *vice versa*. The peculiar properties of the N protein emerged by the present work and the related elusive structural dynamics connected to its oligomeric form suggest how further structural insights of this complex might be solved only by acquiring large cryo-EM single-particle analysis datasets of samples in stabilizing buffer conditions. Most likely, only the application of deep neuronal networks can be meaningful to further segregate the cryo-EM data and interpreting the conformational landscape of this large complex. This is a challenge that could potentially bring to important advances into different aspect of this pandemic pathogen.

CRediT authorship contribution statement

Domenica Farci: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **André T. Graça:** Conceptualization, Data curation, Validation, Methodology, Resources, Writing – original draft. **Michael Hall:** Writing – review & editing, Writing – original draft, Validation, Resources, Methodology, Investigation, Data curation. **Patrycja Haniewicz:** Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Sami Kereïche:** Writing – review & editing, Writing – original draft, Investigation. **Peter Faull:** Writing – review & editing, Writing – original draft, Validation, Resources, Methodology, Investigation, Formal analysis. **Joanna Kirkpatrick:** Writing – review & editing, Writing – original draft, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation. **Enzo Tramontano:** Writing – review & editing, Writing – original draft, Validation, Resources. **Wolfgang P. Schröder:** Writing – review & editing, Writing – original draft, Validation, Resources. **Dario Piano:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

This work was supported by the Polish National Agency for Academic Exchange (NAWA, Poland) with the Urgency Grant program (2020) code PPN/GIN/2020/1/00028/DEC/1 (to D.P., D.F., P.H., W.P. S., A.T.G., and M.H.).

2.9. Data availability

Materials, data, and associated protocols are available within the manuscript and the public repository Electron Microscopy Data Bank (EMDB) as well as the final 3D volumes and other relevant information about data acquisition and processing (EMD-17430 and EMD-16868). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD042943. Data will be released upon publication. This paper does not report any original code.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful to the Polish National Agency for Academic Exchange (NAWA, Poland) for funding this study during the Sars-CoV-2 pandemic. The data were collected at the Umeå Core Facility for Electron Microscopy, a node of the Cryo-EM Swedish National Facility, funded by the Knut and Alice Wallenberg, Family Erling Persson and Kempe Foundations, SciLifeLab, Stockholm University and Umeå University. Authors thank the Electron Microscopy Public Image Archive (EMPIAR), a public resource for raw electron microscopy images. This is a remarkable tool to accelerate research and make use of any possible resource. We gratefully acknowledge Prof. Joseph Davis's team at the Massachusetts Institute of Technology (MIT) for their collaboration in using CryoDRGN (https://dspace.mit.edu/handle/1721.1/133247) to explore the protein's dynamics, though the complexity of these dynamics prevented clear distinctions.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2024.108162.

Data availability

Data will be made available on request.

References

Bai, Z., Cao, Y., Liu, W., Li, J., 2021. The SARS-CoV-2 Nucleocapsid protein and its role in viral structure, biological functions, and a potential target for drug or vaccine mitigation. Viruses. 13 (6), 1115. https://doi.org/10.3390/v13061115.

Bessa, L.M., Guseva, S., Camacho-Zarco, A.R., Salvi, N., Maurin, D., Perez, L.M., Botova, M., Malki, A., Nanao, M., Jensen, M.R., Ruigrok, R.W.H., Blackledge, M., 2022. The intrinsically disordered SARS-CoV-2 nucleoprotein in dynamic complex with its viral partner nsp3a. Sci Adv. 8 (3), eabm4034. https://doi.org/10.1126/sciadv.abm4034.

Caruso, I.P., Dos Santos, A.V., do Amaral, M.J., de Andrade, G.C., de Araújo, G.R., de Araújo, T.S., de Azevedo, J.M., Barbosa, G.M., Bartkevihi, L., Bezerra, P.R., Dos Santos Cabral, K.M., de Lourenço, I.O., Malizia-Motta, C.L.F., de Luna Marques, A., Mebus-Antunes, N.C., Neves-Martins, T.C., de Sá, J.M., Sanches, K., Santana-Silva, M.C., Vasconcelos, A.A., da Silva Almeida, M., de Amorim, G.C., Anobom, C. D., Da Poian, A.T., Gomes-Neto, F., Pinheiro, A.S., Almeida, F.C.L., 2022. Insights into the specificity for the interaction of the promiscuous SARS-CoV-2 nucleocapsid protein N-terminal domain with deoxyribonucleic acids. Int. J. Biol. Macromol. 203, 466–480. https://doi.org/10.1016/j.ijbiomac.2022.01.121.

Chen, C.Y., Chang, C.K., Chang, Y.W., Sue, S.C., Bai, H.I., Riang, L., Hsiao, C.D., Huang, T.H., 2007. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J. Mol. Biol. 368 (4), 1075–1086. https://doi.org/10.1016/j.jmb.2007.02.069.

Cong, Y., Ulasli, M., Schepers, H., Mauthe, M., V'kovski, P., Kriegenburg, F., Thiel, V., de Haan, C.A.M., Reggiori, F., 2020. Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in Coronaviral life cycle. J. Virol. 94 (4), e01925–e02019. https://doi.org/10.1128/JVI.01925-19.

Cope, J., Heumann, J., Hoenger, A., 2011. Cryo-electron tomography for structural characterization of macromolecular complexes. Curr. Protoc. Protein Sci. Chapter 17. Unit 17.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., Mann, M., 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 10 (4), 1794–1805. https://doi.org/10.1021/pr101065j.

Cubisino, S.A.M., Milenkovic, S., Conti-Nibali, S., Musso, N., Bonacci, P., De Pinto, V., Ceccarelli, M., Reina, S., 2024. Electrophysiological properties and structural prediction of the SARS-CoV-2 viroprotein E. Front. Mol. Biosci. 28 (11), 1334819. https://doi.org/10.3389/fmolb.2024.1334819.

Cubuk, J., Alston, J.J., Incicco, J.J., Singh, S., Stuchell-Brereton, M.D., Ward, M.D., Zimmerman, M.I., Vithani, N., Griffith, D., Wagoner, J.A., Bowman, G.R., Hall, K.B., Soranno, A., Holehouse, A.S., 2021. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. Nat Commun. 12 (1), 1936. https://doi.org/10.1038/s41467-021-21953-3.

Duckert, P., Brunak, S., Blom, N., 2004. Prediction of proprotein convertase cleavage sites. Protein Eng Des Sel. 17 (1), 107–112. https://doi.org/10.1093/protein/gzh013.

Farci, D., Cocco, E., Tanas, M., Kirkpatrick, J., Maxia, A., Tamburini, E., Schröder, W.P., Piano, D., 2022. Isolation and characterization of a main porin from the outer membrane of *Salinibacter ruber*. J. Bioenerg. Biomembr. 54 (5–6), 273–281. https://doi.org/10.1007/s10863-022-09950-7.

Garten, W., 2018. Characterization of proprotein convertases and their involvement in virus propagation. Activation of Viruses by Host Proteases. 16, 205–248. https://doi.org/10.1007/978-3-319-75474-1_9.

Gupta, R., Brunak, S., 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac. Symp. Biocomput. 2002, 310–322.

He, R., Leeson, A., Ballantine, M., Andonov, A., Baker, L., Dobie, F., Li, Y., Bastien, N., Feldmann, H., Strocher, U., Theriault, S., Cutts, T., Cao, J., Booth, T.F., Plummer, F. A., Tyler, S., Li, X., 2004. Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus. Virus Res. 105 (2), 121–125. https://doi.org/10.1016/j.virusres.2004.05.002.

Huang, C.Y., Hsu, Y.L., Chiang, W.L., Hou, M.H., 2009. Elucidation of the stability and functional regions of the human coronavirus OC43 nucleocapsid protein. Protein Sci. 18 (11), 2209–2218. https://doi.org/10.1002/pro.225.

Jia, Z., Liu, C., Chen, Y., Jiang, H., Wang, Z., Yao, J., Yang, J., Zhu, J., Zhang, B., Yuchi, Z., 2022. Crystal structures of the SARS-CoV-2 nucleocapsid protein C-terminal domain and development of nucleocapsid-targeting nanobodies. FEBS J. 289 (13), 3813–3825. https://doi.org/10.1111/febs.16239.

Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Z., Zhou, Z., Chen, Q., Yan, Y., Zhang, C., Shan, H., Chen, S., 2020. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm. Sin. b. 10 (7), 1228–1238. https://doi.org/10.1016/j.apsb.2020.04.009.

Klein, S., Cortese, M., Winter, S.L., Wachsmuth-Melm, M., Neufeldt, C.J., Cerikan, B., Stanifer, M.L., Boulant, S., Bartenschlager, R., Chlanda, P., 2020. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nat. Commun. 11 (1), 5885. https://doi.org/10.1038/s41467-020-19619-7.

Lutomski, C.A., El-Baba, T.J., Bolla, J.R., Robinson, C.V., 2021. Multiple Roles of SARS-CoV-2 N protein facilitated by proteoform-specific interactions with RNA, host proteins, and convalescent antibodies. JACS Au. 1 (8), 1147–1157. https://doi.org/10.1021/jacsau.1c00139.

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

Morse, M., Sefcikova, J., Rouzina, I., Beuning, P.J., Williams, M.C., 2023. Structural domains of SARS-CoV-2 nucleocapsid protein coordinate to compact long nucleic acid substrates. Nucleic Acids Res. 51 (1), 290–303. https://doi.org/10.1093/nar/gkac1179.

Muriaux, D., Darlix, J.L., 2010. Properties and functions of the nucleocapsid protein in virus assembly. RNA Biol. 7 (6), 744–753. https://doi.org/10.4161/rna.7.6.14065.

Noori, R., Sardar, M., 2022. An outlook on potential protein targets of COVID-19 as a druggable site. Mol. Biol. Rep. 49 (11), 10729–10748. https://doi.org/10.1007/s11033-022-07724-3.

Peacock, T.P., Goldhill, D.H., Zhou, J., Baillon, L., Frise, R., Swann, O.C., Kugathasan, R., Penn, R., Brown, J.C., Sanchez-David, R.Y., Braga, L., Williamson, M.K., Hassard, J. A., Staller, E., Hanley, B., Osborn, M., Giacca, M., Davidson, A.D., Matthews, D.A., Barclay, W.S., 2021. The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. Nat. Microbiol. 6 (7), 899–909. https://doi.org/10.1038/s41564-021-00908-w.

Perdikari, T.M., Murthy, A.C., Ryan, V.H., Watters, S., Naik, M.T., Fawzi, N.L., 2020. SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. EMBO J. 39 (24), e106478. https://doi.org/10.15252/embj.2020106478.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296.

Seidah, N.G., Prat, A., 2012. The biology and therapeutic targeting of the proprotein convertases. Nat. Rev. Drug. Discov. 11 (5), 367–383. https://doi.org/10.1038/nrd3699.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., Li, F., 2020. Structural basis of receptor recognition by SARS-CoV-2. Nature. 581 (7807), 221–224. https://doi.org/10.1038/s41586-020-2179-y.

Turoňová, B., Sikora, M., Schürmann, C., Hagen, W.J.H., Welsch, S., Blanc, F.E.C., von Bülow, S., Gecht, M., Bagola, K., Hörner, C., van Zandbergen, G., Landry, J., de Azevedo, N.T.D., Mosalaganti, S., Schwarz, A., Covino, R., Mühlebach, M.D., Hummer, G., Krijnse Locker, J., Beck, M., 2020. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. Science. 370 (6513), 203–208. https://doi.org/10.1126/science.abd5223.

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., Thiel, V., 2021. Coronavirus biology and replication: implications for SARS-CoV-2. Nat. Rev. Microbiol. 19 (3), 155–170. https://doi.org/10.1038/s41579-020-00468-6.

Wang, Y., Wu, X., Wang, Y., Li, B., Zhou, H., Yuan, G., Fu, Y., Luo, Y., 2004. Low stability of nucleocapsid protein in SARS virus. Biochemistry. 43 (34), 11103–11108. https://doi.org/10.1021/bi049194b.

Wu, W., Cheng, Y., Zhou, H., Sun, C., Zhang, S., 2023. 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. 20 (1), 6. https://doi.org/10.1186/s12985-023-01968-6.

Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., Zhang, J., Weng, T., Zhang, Z., Wu, Z., Cheng, L., Shi, D., Lu, X., Lei, J., Crispin, M., Shi, Y., Li, L., Li, S., 2020. Molecular Architecture of the SARS-CoV-2 Virus. Cell. 183 (3), 730–738.e13. https://doi.org/10.1016/j.cell.2020.09.018.

Ye, Q., West, A.M.V., Silletti, S., Corbett, K.D., 2020. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci. 29 (9), 1890–1901. https://doi.org/10.1002/pro.3909.

Zinzula, L., Basquin, J., Bohn, S., Beck, F., Klumpe, S., Pfeifer, G., Nagy, I., Bracher, A., Hartl, F.U., Baumeister, W., 2021. 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. 538, 54–62. https://doi.org/10.1016/j.bbrc.2020.09.131.