Notification no: 600/2026

Notification date: 26.05.26

Date of Award: 08.05.2026

Name of the Scholar: Abdus Samad

Name of the Supervisor: Prof. Shama Parveen

Name of the Co-supervisor: Prof Asimul Islam

Faculty: Faculty of Life Sciences

Title: Mechanistic Insights into the SARS CoV-2 Nucleocapsid Protein: Structure, Dynamics, and Inhibitors Screening through Computational and *in vitro* Strategies

Findings

SARS CoV-2 is a positive sense single stranded RNA virus of 29.8 kB, which encodes two large open reading frames (ORF1a & ORF1ab) into translate into form sixteen (16) non-structural proteins and four (4) structural proteins [spike (S), envelope (E), membrane (M), nucleocapsid (N)] of SARS CoV-2. N protein binds with the viral RNA and forms ribonucleoacid complexes (RNPs) and interacts with the M protein, resulting in viral assembly and packing during the replication cycle.

In this study, the N protein gene, which encodes the RNA-binding domain, C-terminal domain, and linker region, was transformed into *E. coli* DH5α cells to increase plasmid copy number. The confirmed plasmid was retransformed into *E.coli* BL21 codon+ cells, and the expression of the N protein was optimized using different inducer concentrations (IPTG) at 16 °C and 37 °C. The optimization data were further used to purify the N protein on a large-scale using affinity and gel filtration chromatography. The desired 37 kDa N protein observed in the eluted fraction by SDS-PAGE was further validated by western blot using anti-His antibodies. The eluted fraction was further dialyzed to remove impurities and refold the N protein into its native conformation. The native conformation of the N protein was evaluated using UV-Visible spectroscopy. UV-visible spectroscopy confirmed the folded and native structures of the N protein.

The conformational stability and folding behaviour of the N protein were investigated at varying pH using circular dichroism spectroscopy, intrinsic and extrinsic fluorescence spectroscopy, dynamic light scattering, differential scanning calorimetry, and further complemented by MD simulation studies. The results from circular dichroism and fluorescence spectroscopy are in close agreement. At pH 2.0 and 3.0, the N protein loses its tertiary structure but maintains its native conformation at mild acidic and alkaline pH values. An intermediary complex was observed at pH 2.0 and 3.0 and was confirmed using the extrinsic ANS dye. The DLS studies revealed that the hydrodynamic diameter of the protein increased at acidic pH values compared to alkaline pH values due to the loss of tertiary structure of the protein and validated the formation of the intermediary-like complex at this highly acidic pH. Structural changes of the N protein as a function of various pH values were also determined using 100 ns molecular dynamics simulations. The N protein showed maximum structural deviation (RMSD) at pH 2.0 and 12.0 and minimum structural deviation at pH 8.0 and 10.0, further complementing the biophysical characterization results. The Rg value and SASA plot inferred that the N protein has a more compact structure at native pH 8 than at acidic pH 2.0, 3.0 and 4.0. Moreover, the secondary structure analysis showed that the total number of average

residues participating in the secondary structure formation is more at the pH 8.0, while the coil percentage is more at acidic pH 2.0 and 3.0, further validating the biophysical findings. Further, the formation of an intermediate-like state at acidic pH may be a defence mechanism for the protein to survive under hostile conditions. Overall, the findings from both the *in silico* and *in vitro* aspects complement each other.

The stability of N protein in the presence of urea, a strong denaturant, was measured using biophysical techniques such as fluorescence spectroscopy, circular dichroism, and differential scanning calorimetry (DSC), and the results were complemented with 300 ns MD simulations using Gromacs 2022.2. The biophysical studies revealed that the N protein loses its tertiary structure cooperatively, starting at 2.00 M of urea-induced denaturation, with a redshift in λ_{max}. The CD spectroscopy revealed that there is a loss in random coil content, hence secondary structure content and a shift of θ_{max} towards higher wavelength. The DSC experiments revealed increase in T_m along with a decrease in ΔC_p value, inferring formation of intermediate like state. All the *in vitro* studies were further complemented using 300 ns MD simulation using Gromacs 2022.2. These data provide information of structural characteristics and dynamics of the N protein in absence of a high-resolution crystal structure and redefine the current understanding of the structure of the N protein.

Screening 64,594 antiviral compounds against the N-terminal (6VYO) and C-terminal (6YUN) domains identified ten candidates with favourable pharmacokinetic and binding profiles. Docking and molecular dynamics highlighted Compounds 4 and 5 for the NTD and 8 and 9 for the CTD as the most promising inhibitors.

Furthermore, *in vitro* validation confirmed direct binding of compounds 4, 5, 8, and 9 to the SARS-CoV-2 N protein. Fluorescence spectroscopy indicated stable complexes, with compounds 8 and 9 showing higher affinities and near 1:1 binding stoichiometry. ITC confirmed spontaneous, exothermic, primarily enthalpy-driven interactions, with compounds 4, 8, and 9 displaying favourable thermodynamic profiles. SPR further showed distinct binding kinetics, with compounds 4 and 9 exhibiting slower dissociation and more sustained target engagement.

Beyond binding studies, the functional relevance of these interactions was evaluated using LLPS. Fluorescence microscopy revealed that the N protein undergoes phase separation in the presence of RNA analog, forming dynamic biomolecular condensates. Importantly, all tested compounds induced pronounced morphological alterations in these condensates, converting liquid-like droplets into irregular, solid-like aggregates. This disruption of LLPS suggests interference with the multivalent interactions essential for viral ribonucleoprotein assembly and highlights a potential mechanism by which these compounds may impair viral replication.

Cytotoxicity assessment using the MTT assay in Vero E6 cells demonstrated that all compounds exhibited relatively high CC50 values, indicating low cytotoxicity at concentrations exceeding those required for effective protein binding and LLPS disruption. Compounds 4 and 9 showed, supporting their suitability for further biological evaluation.

Overall, this study provides an integrated structural, biophysical, and computational analysis of the SARS-CoV-2 nucleocapsid protein, establishing it as a viable antiviral target. The identification and experimental validation of small-molecule inhibitors that bind the N protein, disrupt LLPS, and exhibit low cytotoxicity offer a promising foundation for the development of novel therapeutics against SARS-CoV-2 and related coronaviruses.