SARS-CoV-2: The Prominent Role of Non-structural Proteins (NSPS) in COVID-19

Uma Sankar Gorla*, GSN Koteswara Rao
College of Pharmacy, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, INDIA.

ABSTRACT
COVID-19, an infectious contagious viral (SARS-CoV-2) disease, associated with morbidity and mortality from respiratory pandemic worldwide. Currently, COVID-19 has no targeted treatment strategies and has been declared by WHO as a global health emergency. SARS-CoV-2, an enveloped, positive-sense, betacoronavirus (βCoV) spreading rapidly due to its potential pre- and asymptomatic transmission (silent transmission). The viral replicase single-stranded genome encodes for two open reading frame genes (orf1a and orf1b), that are translated into two polyproteins, pp1a and pp1b. Viral proteases including papain-like protease (PL-PRO) and chymotrypsin-like protease (3CL-PR0) pre-processes and fragments the polyproteins into 16 non-structural proteins (Nsps) that are assembled into replicase-transcriptase complex and exhibit multiple enzymatic activity. To rationalize their evolutionary success and develop improved control strategies, understanding the main functions and interactions of non-structural proteins of SARS-CoV-2 will be essential. Based on the existing published literature, this review summarizes the knowledge on multiple pathologic functions of Nsps in COVID-19. This review is hoped to help the researchers to understand the significant role of non-structural proteins in COVID-19 and provide a reference for future studies.

Key words: COVID-19, SARS-CoV-2, Non-structural proteins, Open reading frames, Polyproteins, Phylogenetic analysis.


INTRODUCTION
Coronavirus disease 2019 (COVID-19), a pandemic infectious respiratory disease caused by an enveloped, positive-sense, single-stranded RNA novel coronavirus named as SARS-CoV-2.¹² As of 23 July 2020, WHO reported 1,50,12,731 cases globally in 216 countries including 6,19,150 deaths for the Corona outbreak 2019 and it has been declared as a global health emergency (https://www.who.int/emergencies/diseases/novel-coronavirus-2019). The virus has locked down the millions and the catastrophes are threatening the global economy.³ COVID-19 is a highly contagious disease associated with morbidity and mortality due to its potential pre- and asymptomatic transmission.⁴ SARS-CoV-2 belongs to the order Nidovirales, the family Coronaviridae, the genus Betacoronavirus and the species Severe acute respiratory syndrome-related coronavirus.⁵ SARS-CoV-2, which is responsible for the COVID-19 pandemic, is spreading rapidly all over the globe due to the transmission from human-to-human.⁶ The SARS-CoV-2 reproductive number is 2.68 with an incubation period of 6.4 days (range 2.1 to 11.1 days) and the time of onset of symptoms is between 1 and 14 days, usually 5 days.⁷ Studies have found that majority of COVID-19 transmission is due to silent spreaders that do not show symptoms.⁸ The silent transmission of SARS-CoV-2 infection is due to virus
shedding patterns in asymptomatic patients.\textsuperscript{9} Recent studies indicates that SARS-CoV-2 shedding may begin 2 to 3 days before the onset of first clinical signs and symptoms.\textsuperscript{10} Regarding the potential routes of SARS-CoV-2 transmission, respiratory route was the most predominant route in humans. However, recent evidences indicates that the SARS-CoV-2 may also be transmitted through the fecal-oral route, since the virus was found in endoscopic and stool specimens from COVID-19 patients.\textsuperscript{11} The ongoing epidemic originally identified as pneumonia of unknown etiology in a group of patients in Wuhan, China.\textsuperscript{12} The Chinese Centre for Disease Control and Prevention (CDC, China) subsequently categorized the clinical manifestations of COVID-19 based on the severity with a wide range of clinical symptoms ranging from asymptomatic patients to septic shock and multiorgan dysfunction.\textsuperscript{13} The most frequently reported symptoms of COVID-19 includes fever, dry cough and fatigue.\textsuperscript{14} In addition to pulmonary symptoms, 2\% to 10\% of COVID-19 patients reported GIT manifestations such as nausea, vomiting, diarrhoea and abdominal pain.\textsuperscript{15} Extra pulmonary and systemic manifestations such as cardiac arrhythmias, cardiac failure, deep vein thrombosis, renal tissue damage, liver damage, seizures, Guillain-Barre syndrome and confusion are caused by a direct attack of SARS-CoV-2 on cardiac muscles, kidneys, blood vessels, liver and central nervous system with ACE2 receptors.\textsuperscript{16} Furthermore, SARS-CoV-2 targets male reproductive system due to high expression of ACE2 and has been detected in semen.\textsuperscript{17} The complications include spermatogenic failure, autoimmune orchitis, hypogonadism, germ cell destruction, testicular dysfunction and infertility.\textsuperscript{18} Despite these recent discoveries, the current review aimed to summarize the existing literature on multiple pathologic functions of non-structural proteins in SARS-CoV-2 replication and virulence.

**SARS-CoV-2 genomic organization**

Phylogenetic studies of the SARS-CoV-2 have shown a nucleotide correlation of 96\% with Bat-SL-RaTG13, 88\% with Bat-SL-Cov-ZC45 and Bat-SL-CoV-ZXC21, 79.6\% with SARS-CoV and 50\% with MERS-CoV.\textsuperscript{19} The analysis revealed that the single-stranded RNA genome (30kb) of SARS-CoV-2 was 29,891 nucleotides in size (9860 amino acids) encoding up to 14 open reading frames (orfs).\textsuperscript{20,21} The 5´-and 3´-untranslated regions of SARS-CoV-2 are identical to those of βCoVs in nucleotide sequence (≥83.6\%). In genetic configuration, the genome is organized as 5´-replicase-transcriptase complex with accessory factors-3´. The 5´ replicase genome encodes two open reading frame genes (orf1a and orf1b), that are translated into two polyproteins, pp1a and pp1b.\textsuperscript{22} These polyproteins were pre-processed and fragmented by viral proteases into 16 non-structural proteins (NSPS), which are assembled into replicase-transcriptase complex and exhibit multiple enzymatic activity (Table 1).\textsuperscript{23} The putative non-structural proteins in the replicase-transcriptase complex includes the papain-like protease (Nsp3, PL-PRO), the main protease (Nsp5, Main protease (M\textsuperscript{pro})), chymotrypsin-like protease (3CL-PRO), the primase complex (Nsp7-Nsp8), the primary RNA-dependent RNA polymerase (Nsp12, RdRp), a helicase (Nsp13, triphosphatase, Hel), a guanine-N7 methyltransferase (Nsp14, exoribonuclease, ExoN), an uridylate-specific endoribonuclease (Nsp15, NendoU) and a hetero-oligomeric complex (Nsp10/Nsp16, N7- and 2´ O-ribose methyl transferase complex).\textsuperscript{24} The genomic 3´ end encodes as many as 13 orfs which includes four structural proteins such as spike (S), envelope (E), membrane (M) and nucleocapsid (N) along with nine potential accessory factors (Figure 1).\textsuperscript{25} Non-structural protein 1

Nsp1, a membrane-associated host translation inhibitor protein which anchors replication complex to the cellular membranes. Nsp1 forms a complex with 40S ribosomal subunit, that degrades host mRNAs by inducing an endonucleolytic cleavage near the 5´-untranslated region and thus inhibits host translation.\textsuperscript{26} SARS-CoV-2 mRNAs are not prone to nsp1-mediated endonucleolytic degradation because of the presence of 5´-end leader sequence.\textsuperscript{27} By suppressing host gene expression, the leader protein promotes efficient expression of viral genes in infected cells and escape from the host immune response.\textsuperscript{28}
Non-structural protein 2
Nsp2 plays a vital role in modulation of host cell survival signalling pathway through interaction with host prohibitins (PHB and PHB2).\textsuperscript{29} Prohibitin (B-cell receptor associated protein-32) and prohibitin 2 (repressor of estrogen receptor activity) are ubiquitously expressed and present in several cell compartments like mitochondria, preserving functional integrity and defending cells from various stresses.\textsuperscript{30} The stabilizing mutation occurring inside the endosome-associated-protein-like domain of the nsp2 protein may account for highly contagious COVID-19.\textsuperscript{31}

Non-structural protein 3
Nsp3, Papain-like protease (PL-PRO) is the largest cystine protease with a proteolytic core domain of 316 amino acids responsible for cleaving first three cleavage sites at the N-terminus of the replicase substrates (poly proteins).\textsuperscript{32} In addition to its role in viral protein maturation, Nsp3 has deubiquitinating/delISGylating activity and processes polyubiquitin chains linked to ‘Lys48’ and ‘Lys63’ from cellular substrates to supress antiviral innate immune response of the host cell.\textsuperscript{33,34} Along with Nsp4 and other host proteins, Nsp3 engages in the rearrangement of cytoplasmic double-membrane vesicle required for the replication of SARS-CoV-2.\textsuperscript{35} Nsp3 also prevents host Nuclear Factor Kappa B (NF-κB) signalling and antagonizes innate immune response of type I interferon by blocking phosphorylation, dimerization and subsequent nuclear translocation of host interferon regulatory factor 3 (IRF3).\textsuperscript{36,37} The destabilizing mutation happening near the phosphatase domain of the Nsp3 may suggest a potential mechanism that differentiates SARS-CoV-2 from SARS. Because of its diverse actions, Nsp3 may provide new avenues for investigating the virus replication cycle in host cells, with the goal of developing therapeutic agents to inhibit replication of SARS-CoV-2.

Non-structural protein 4
Nsp4, a membrane-spanning protein interacts with Nsp3 and other host proteins and plays a significant role in SARS-CoV replication through the rearrangements of viral induced cytoplasmic double-membrane vesicles.\textsuperscript{38} The alpha helical (C-terminal) domain may be involved in protein-protein interactions that anchors the viral replicase-transcriptase complex to modified cytoplasmic membranes necessary for viral replication.\textsuperscript{39}

Non-structural protein 5
Nsp5, chymotrypsin-like protease (3CL-PRO) or Main protease (M\textsuperscript{pro}), a key enzyme responsible for SARS-CoV-2 replication.\textsuperscript{40} The M\textsuperscript{pro} exists in homodimer and has Cys-His dyad on active site which shows protease activity.\textsuperscript{41} Nsp5 digest replicase polyprotein (C-terminus) at 11 conserved sites to generate non-structural proteins (Nsp4-Nsp16), which plays a vital role in mediating viral replication and transcription and serves as a desirable target for discovery and development of antivirals against COVID-19.\textsuperscript{42}

Non-structural protein 6
SARS-CoV-2 non-structural protein 6 (Nsp6) binds with sigma receptor of host endoplasmic reticulum and

<table>
<thead>
<tr>
<th>Table 1: Putative functions of SARS-CoV-2 non-structural proteins.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nsp</strong></td>
</tr>
<tr>
<td>Nsp1</td>
</tr>
<tr>
<td>Nsp2</td>
</tr>
<tr>
<td>Nsp3</td>
</tr>
<tr>
<td>Nsp4</td>
</tr>
<tr>
<td>Nsp5</td>
</tr>
<tr>
<td>Nsp6</td>
</tr>
<tr>
<td>Nsp7/ Nsp8</td>
</tr>
<tr>
<td>Nsp9</td>
</tr>
<tr>
<td>Nsp10</td>
</tr>
<tr>
<td>Nsp11</td>
</tr>
<tr>
<td>Nsp12</td>
</tr>
<tr>
<td>Nsp13</td>
</tr>
<tr>
<td>Nsp14</td>
</tr>
<tr>
<td>Nsp15</td>
</tr>
<tr>
<td>nsp16</td>
</tr>
</tbody>
</table>
initiates the induction of autophagosomes.\textsuperscript{43} Generally, sigma receptor activation regulates endoplasmic reticulum stress response. Nsp6-sigma receptor interaction restricts autophagosomal expansion, which is no longer capable of delivering viral products to host lysosomes.\textsuperscript{43} Nsp3, Nsp4 and Nsp6 all together have the capability to induce double-membrane vesicles in which Nsp3 and Nsp4 are capable of pairing membranes while Nsp6 has membrane proliferation ability.\textsuperscript{46}

**Non-structural protein 7/8 complex**

Nsp7 binds with Nsp8 (8 subunits of each) to form a hexadecamer known as primase complex that interacts with RNA-dependent RNA polymerase (Nsp12) and forms hetero-oligomeric complex (Nsp12-Nsp7/Nsp8) that may participate in SARS-CoV-2 replication.\textsuperscript{46} Moreover, these Nsp's can synthesize significantly longer products than oligonucleotide polymers. Nsp8 was predicted to have adhesins, essential for viral adherence and host invasion.\textsuperscript{47}

**Non-structural protein 9**

Nsp9 acts as a single-stranded RNA-binding protein that mediates both viral replication and virulence. Nsp9, an essential protein binds RNA through oligosaccharide/oligonucleotide fold-like fold that is unique to the class of βCoVs and promotes viral replication.\textsuperscript{48} Nsp9 plays a crucial role in SARS-CoV virulence and may play a similar role in upholding functional integrity of the virus.\textsuperscript{49} Researchers have determined the structure of Nsp9 and identified a peptide binding site which could prompt further research in understanding its role in COVID-19.\textsuperscript{50}

**Non-structural protein 10**

Nsp10, an essential cofactor that activates guanine-N7 methyltransferase (Nsp14) and 2´ O-ribose methyltransferase (Nsp16) and plays an important part in methylation of mRNAs guanosine cap to promote transcription, splicing, polyadenylation and nuclear export of viral mRNA.\textsuperscript{51} The hetero-oligomeric complex (Nsp10-Nsp14-Nsp16) may be a target for the development of antivirals against pathogenic coronaviruses.\textsuperscript{52}

**Non-structural protein 11**

Nsp11, a shortest peptide expressed at the end of orf1a polyprotein of SARS-CoV-2 replicase genome having 85% amino acid similarity with bat-SL-CoVZXC21 and SARS-CoV.\textsuperscript{53}

**Non-structural protein 12**

Nsp12, a RNA-dependent RNA polymerase (RdRp), is primarily responsible for replication and transcription of the SARS-CoV-2 genome.\textsuperscript{54} The Nsp12 has ssRNA and ssDNA-dependent polymerase activity with no priming activity and is involved in post transcriptional gene slicing (PTGS).\textsuperscript{55} SARS-CoV-2 nucleotide polymerase was predicted to have 932 amino acids with N-terminal and a polymerase domain. The amino acid sequence alignment showed that SARS-CoV-2 Nsp12 shared 96.35% similarity to SARS Nsp12. The Nsp12 interacts with Nsp7/Nsp8 complex (Primase complex) and activates to form RNA synthesis machinery to replicate long RNA.\textsuperscript{56} On the other side, the SARS-CoV-2 nsp12 possess seven conserved motifs (motifs A-E) in the polymerase active site, which are involved in a template and nucleotide binding and catalysis.\textsuperscript{57} The hetero-oligomeric complex (Nsp12-Nsp7/Nsp8) could be the potential target (PDB ID:6M71) to inhibit the RdRp activity of Nsp12 and helps researchers to identify a novel antiviral agent against SARS-CoV-2.\textsuperscript{58}

**Non-structural protein 13**

Nsp13, Helicase (Hel) or Nucleoside-triphosphatase (NTPase), a multifunctional protein with a zinc-binding domain in N-terminus unwinds duplex RNA and DNA with a 5´ single-stranded tail in a 5´ to 3´ direction and plays an important role in central dogma of the virus.\textsuperscript{59} Despite its significant role in the replication of SARS-CoV-2 RNA, Nsp13 readily unwinds duplex DNA due to its functional cooperativity rather than structural interactions between helicase monomers.\textsuperscript{60} The helicase activity depends on magnesium and may exhibit various properties during unwinding of RNA.\textsuperscript{61} Due to its NTPase and helicase activities, SARS-CoV-2 helicase plays a crucial role in replication and virulence and is considered as a target for antivirals.\textsuperscript{62}

**Non-structural protein 14**

Nsp14, a bifunctional replicase subunit with exoribonuclease activity (proofreading) that acts in a 3´ to 5´ direction on both ssRNA and dsRNA and a guanine-N7 methyltransferase activity at its C-terminal.\textsuperscript{63,64} The Nsp10 interacts with Nsp14 and Nsp16 in order to form a dodecamer to enhance their enzyme activities and plays a key role in RNA synthesis and replication fidelity.\textsuperscript{65} The N-terminal portion of Nsp14 interacts with ATP-dependent RNA helicase (DDX1) C-terminal region which enhances replication process.\textsuperscript{66}
Non-structural protein 15

Nsp15, a Mn²⁺-dependent nidoviral uridylate-specific endoribonuclease (NendoU) that leaves 2’-3’-cyclic phosphates 5’ to the cleave bond and performs various vital functions associated with RNA processing. Nsp15 acts as interferon (IFN) antagonist and inhibits interferon-β production through an endonuclease activity-independent mechanism. Moreover, Nsp8 and Nsp7/Nsp8 complex enhances endoribonuclease activity of hexameric Nsp15.

Non-structural protein 16

Nsp16, 2’ O-ribose methyltransferase facilitates methylation of mRNA cap 2’ O-ribose to the 5’-cap structure of viral mRNAs to form N7-methylguanosine cap and plays an important role in methylation of viral mRNAs cap which is necessary for evading immune system. The enzymatic activity of Nsp16 is increased by interaction with Nsp10 and is essential for viral replication-transcription in host cells. With the ongoing threat of SARS-CoV-2 virulence, it is important to use the highly conserved Nsp10/Nsp16 heterodimer interface to establish new treatment strategies against SARS-CoV-2.

Due to the lack of FDA approved drugs for the treatment of human coronavirus infection and vaccines to prevent COVID-19, research against SARS-CoV-2 is ongoing globally. Furthermore, there are many SARS-CoV-2 proteins that have been reported as possible targets for drugs. In addition to the structural proteins, non-structural proteins also play a significant role in the replication and virulence of SARS-CoV-2. Bioinformatic tools can be used to predict drug-like properties, ADMET properties, toxicity profiles, bioactivity scores and antiviral properties for any experimental drug molecule. In silico molecular docking studies can be useful to predict the binding affinity between the experimental drug and the target protein (Table 2) and play a vital role in finding an inhibitor through structure-based drug design. The key limitation of the in-silico studies has false positive results and low coefficients of correlation between the predicted binding energies and experimental values provided in previous research. Despite the drawbacks, the molecular interactions, dock scores and binding energies provide potential information that can notify and guide further in vitro, in vivo and clinical studies.

CONCLUSION

The ongoing COVID-19 pandemic clearly reflects a global public health issue. The novel coronavirus, SARS-CoV-2 spreads rapidly due to its pre- and asymptomatic silent transmission. The review summarized multiple pathologic functions of non-structural proteins in SARS-CoV-2 replication and virulence. Several NspS assembled to form replicase-transcriptase and heterooligomeric complexes and exhibit multiple enzymatic activity. Moreover, there is urgent need to identify and characterize potent antiviral compound against specific targets (Nsps) to combat the emerging COVID-19 pandemic. Our review might contribute by providing substantial information regarding the role of SARS-CoV-2 non-structural proteins in COVID-19.

Table 2: Literature representing protein data base crystal structures of SARS-CoV-2 non-structural proteins.

<table>
<thead>
<tr>
<th>Nsp</th>
<th>PDB Id</th>
<th>Protein (PDB)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nsp3</td>
<td>6W6Y</td>
<td>Crystal structure of ADP ribose phosphatase of nsp3 from SARS-CoV-2</td>
<td>77</td>
</tr>
<tr>
<td>Nsp5</td>
<td>6LU7</td>
<td>SARS-CoV-2 main protease in complex with an inhibitor N3</td>
<td>78</td>
</tr>
<tr>
<td>Nsp7/Nsp8</td>
<td>6YHU</td>
<td>Crystal structure of the nsp7-nsp8 complex of SARS-CoV-2</td>
<td>79</td>
</tr>
<tr>
<td>Nsp9</td>
<td>6WXD</td>
<td>SARS-CoV-2 Nsp9 RNA-replicase</td>
<td>80</td>
</tr>
<tr>
<td>Nsp12</td>
<td>6M71</td>
<td>SARS-CoV-2 RNA-dependent RNA polymerase in complex with cofactors</td>
<td>81</td>
</tr>
<tr>
<td>Nsp15</td>
<td>6W01</td>
<td>Crystal structure of nsp15 Endoribonuclease from SARS-CoV-2 in the complex with a citrate</td>
<td>82</td>
</tr>
<tr>
<td>nsp16-nsp10 complex</td>
<td>6W4H</td>
<td>Crystal structure of nsp16-nsp10 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine</td>
<td>83</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

The authors are grateful to the management, College of Pharmacy, Koneru Lakshmaiah Education Foundation, Vaddeswaram, A.P., India.

CONFLICT OF INTEREST

No conflict of interest declared by the authors.

ABBREVIATIONS

CoV: Coronavirus; SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; orfs: Open reading frames; COVID-19: Coronavirus disease 2019; βCoVs: Beta coronavirus; mRNA: Messenger ribonucleic acid; WHO: World health organization; Nsps: Non-structural proteins; PL-PRO: Papain-like protease; Mpro: Main protease; 3CL-PRO: Chymotrypsin-like protease; NendoU: Uriddylate-specific endoribonuclease; RdRp: RNA-dependent RNA polymerase; ExoN: Exoribonuclease; Hel: Helicase; S: Spike; E: Envelop; M: Membrane; N: Nucleocapsid; PHB: prohibitin; NF-xB: Nuclear factor kappa B; PDB: Protein data bank; DNA: Deoxyribonucleic acid; IFN: Interferon; IRF3: Interferon regulatory factor 3.

REFERENCES

Gorla and Rao.: Non-structural Proteins of SARS-CoV-2


**SUMMARY**

- **COVID-19**, a pandemic respiratory disease associated with high morbidity and mortality worldwide.
- Coronavirus disease 2019 is caused by a novel coronavirus named as “SARS-CoV-2”.
- The SARS-CoV-2 genome encodes as many as 14 open reading frames including orf1a, orf1b, 4 structural proteins and 9 accessory proteins.
- The polyproteins pp1a and pp1b translated from open reading frames (orf1a and orf1b) are preprocessed and fragmented in 16 non-structural proteins.
- Several Nsp5s assembled to form replicase-transcriptase and hetero-oligomeric complexes and plays a significant part in viral replication and virulence.
- Npsp could be appropriate therapeutic targets for investigating replication and pathogenesis of SARS-CoV-2.
About Authors

Mr. Uma Sankar Gorla: M. Pharm., (Ph.D.), M.B.A. currently working as Assistant Professor at K L College of Pharmacy, KLEF Deemed to be University, Guntur, A.P. He completed Bachelor of Pharmacy and Master of Pharmacy from Andhra University, Visakhapatnam. He is trained for ICT, online and blended teaching-learning process from IIT, Bombay and NITTR, Chennai. His area of interest includes Docking studies, Molecular modelling, Biochemical estimations, cancer, Alzheimer’s and infectious diseases.

Dr. GSN Koteswara Rao: M.Pharm., Ph.D., MBA working as Associate Professor at K L College of Pharmacy, KLEF Deemed to be University, Guntur is recipient of gold medal and silver medals during his education. He is post graduated from Andhra University and received doctorate degree from JNTUH. He secured 98.6 percentile in GATE-PY 2006 conducted by IIT Kharagpur. He received several awards and also a recipient of UGC fellowship. He presented his research papers at American Association of Pharmaceutical Scientists (AAPS) at San Diego, California, USA. He has 12 years of academic and research experience. To his credit he has published 30 research papers in reputed national and international journals. He participated in more than 70 national and international conferences/symposia/workshops/FDP’s. He organized number of state and national level pharma events. He is trained for ICT, ALM’s and LMS based modern teaching methodologies from IIT Bombay as 3 FDP’s. He did courses in CADD and biomedical.