Structure of the SARS-CoV-2 Nsp1/5′-Untranslated Region Complex and Implications for Potential Therapeutic Targets, a Vaccine, and Virulence

Naveen Vankadari,* Nandhini Nisha Jeyasankar,§ and Wilma Jerom Lopes§

ABSTRACT: SARS-CoV-2 is the cause of the ongoing Coronavirus disease 19 (COVID-19) pandemic around the world causing pneumonia and lower respiratory tract infections. In understanding the SARS-CoV-2 pathogenicity and mechanism of action, it is essential to depict the full repertoire of expressed viral proteins. The recent biological studies have highlighted the leader protein Nsp1 of SARS-CoV-2 importance in shutting down the host protein production. Besides, it still enigmatic how Nsp1 regulates for translation. Here we report the novel structure of Nsp1 from SARS-CoV-2 in complex with the SL1 region of 5′UTR of SARS-CoV-2, and its factual interaction is corroborated with enzyme kinetics and experimental binding affinity studies. The studies also address how leader protein Nsp1 of SARS-CoV-2 recognizes its self RNA toward translational regulation by further recruitment of the 40S ribosome. With the aid of molecular dynamics and simulations, we also demonstrated the real-time stability and functional dynamics of the Nsp1/SL1 complex. The studies also report the potential inhibitors and their mode of action to block viral protein/RNA complex formation. This enhance our understanding of the mechanism of the first viral protein Nsp1 synthesized in the human cell to regulate the translation of self and host. Understanding the structure and mechanism of SARS-CoV-2 Nsp1 and its interplay with the viral RNA and ribosome will open the arena for exploring the development of live attenuated vaccines and effective therapeutic targets for this disease.

The Coronavirus Disease 2019 (COVID19) pandemic caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has been declared to be a public health emergency and is known to have originated from the Wuhan province of China.1,2 The potentially virulent and quickly mutating COVID19 strains have infected more than 36 million people and caused more than 1 million deaths worldwide (https://coronavirus.jhu.edu/map.html). Similar to other coronaviruses (α-, β-, γ-, and δ-coronaviruses), SARS-CoV-2 is an enveloped, single-stranded, positive-sense RNA virus, and its close phylogenetic species are known to infect a large number of vertebrate animal species.3,4 The SARS-CoV-2 genome consists of an ~30 kb linear, nonsegmented, 5′-capped, polystrictonic and polyadenylated RNA molecule consisting of 14 open reading frames (ORFs) in total.5,6 Among these, ORF1ab codes for a maximum number of viral proteins, including structural proteins that are involved in the formation of the viral shape or structure for the direct and physical interaction with the host cell for entry and following infection, whereas the nonstructural proteins are involved in other intracellular viral trafficking via its physical and direct interactions with host intercellular proteins. The largest coding region (pp1ab) codes for 16 nonstructural proteins (Nsp1–16) that are crucial players in the replication and propagation of the virus.7–9 SARS-CoV-2 shares a high sequence and infection mode similar to that of earlier β-coronaviruses such as the Middle East Respiratory Syndrome (MERS) outbreak in 2012 and the SARS-CoV epidemic in 2002,10 infecting the lower respiratory tract of the host. However, the COVID19 outbreak was considered to be more serious with higher morbidity and mortality than for other coronaviruses in history (www.cdc.gov). In addition to the difference in the sequence between the species of coronavirus, overwhelming mutations and deletions were also noticed in the viral genome of SARS-CoV-2.11 Collective changes in the viral genome could have resulted in an altered physical interaction pattern or changes in the intermolecular affinity, resulting in change in the chemical aspects of the interaction mode and binding energy. Similarly, the novel insertion of the Furin cleavage site in the spike
glycoprotein,12−14 and other key modifications in the viral proteins make the virus more virulent, resulting in rapid spread. Furthermore, finding potential drug targets is difficult as a result of changes in the biophysical interaction mode or characteristics of SARS-CoV-2 due to the above variations.13 Because of its serious and exponential rate of infection across the globe and with the emergence of a global second wave (www.cdc.gov), it warrants our understanding of the more fundamental aspects of disease spread, the mode of infection, biophysical—chemical interactions, and its regulation at the molecular and atomic levels for better therapeutic development. Currently, COVID19 patients are prescribed broad-spectrum antiviral drugs such as Remdesivir (that blocks RdRp), Arbidol (that impedes spike proteins), Ritonavir, Hydroxychloroquine, and a combination of several other drugs (https://www.nps.org.au/antivirals_COVID-19) as the first line of defense.15,16 Even though specific antiviral drugs targeting different host and viral proteins are being clinically evaluated,17 a potential vaccine is the ultimate and permanent solution for the COVID19 pandemic. Toward this end, many strategies for the development of COVID19 vaccines are being implemented,18 and the global medical fraternity is advancing toward it. Meanwhile, it is worth clarifying that biophysical and computationally derived models not only exhibit biophysical characteristics such as electrostatics and hydrogen bonding but also are implicitly encoded in the fingerprints, and their impacts on the binding affinity are reflected in their functional aspects. Hence, it calls for the detailed understanding of the molecular mechanism, biophysical interactions, and chemistry behind every viral protein for its interplay with the host proteins toward the infection, which could aid the development of better therapeutic targets and vaccines to curb the disease. Although key viral proteins of SARS-CoV-2 have been functionally and structurally characterized,19−22 the structural and functional role/significance of viral proteins involved in its self-viral RNA translation and their physical interaction in the regulation of translation are the least-addressed topics and require an in-depth investigation. Understanding the physical and chemical aspects of viral RNA and protein interaction via various interaction forces (hydrogen bonding, van der Waals, electrostatic, etc.) and their binding affinities have great implications for potential drug targets.

One among the important proteins is nonstructural protein 1 (Nsp1), also called a leader protein since it is the first protein produced by the N-terminal region of the viral genome (ORF1ab). In most coronaviruses infecting humans, including SARS and MERS, the Nsp1 protein has been proven to be necessary for viral replication and translation regulation.23 Interestingly, Nsp1 is not highly conserved among the family of β-coronavirus but has been shown to have the same physical properties and biological function.24,25 It is also well known that Nsp1 of SARS-CoVs physically and directly interact with the host ribosome (40S subunit), which is crucial for host translation and attacks the host mRNA by two strategies: either blocking the 40S subunit of the ribosome to end the host protein translation or endonucleolytic cleavage near the 5′ UTR of the host mRNA, thus making them incompetent for translation.26,27 However, viral RNA is resistant to this cleavage due to the unique structural interaction between Nsp1 and viral RNA in the 5′ UTR region. Although the biophysical functional significance of Nsp1 in blocking the host ribosome has been reported very recently,28 studies concerning either the dynamics of Nsp1 in relation to ribosome docking or the mechanism of Nsp1 physically recognizing the key region of viral RNA (5′ UTR) to regulate viral protein expression are elusive with respect to SARS-CoV-2. It has been shown in SARS-CoV-1 that Nsp1 directly binds to the 5′-Untranslated region (5′ UTR) of the viral RNA, and this specific interaction enhances viral gene replication in the infected cells by recruiting the host ribosome at that site.29 The 5′ UTR of most RNA viruses has unique stem–loop structures (SL) and are known to recruit key translation regulation proteins; however, the studies concerning the physical association of SARS-CoV-2 RNA in this area are yet to be unraveled. On the other hand, Nsp1 is also known to play a key role in inhibiting the production of host defense components such as type-I interferons25 interferon-gamma, and interleukins, which are key signaling molecules secreted by the infected cells that alert the other cells about the viral entry.26,27 Hence, Nsp1 plays a key role in the host cellular antiviral defense mechanism (regulating the interferon response) and alters the innate immune system to benefit viral replication and immune evasion,25,26,30 making SARS-CoV Nsp1 a potential therapeutic target. Here, we set out to characterize the molecular interplay between the Nsp1 and 5′ UTR of SARS-CoV-2, thus dissecting the physical and chemical constrains for direct interaction in viral replication.

Foremost, to distinguish the sequence similarity and differences in the 5′ UTR region of SARS-CoV-2 from those of SARS-CoV-1 and MERS, we performed multiple sequence alignment (MSA) using an online alignment tool, Clustal Omega (www.clustalomega.ebi.org) (Figure S1A). The genomic RNA sequences of SARS-CoV-2, SARS-CoV, and MERS were retrieved from the curated NCBI virus database (https://www.ncbi.nlm.nih.gov/genome/viruses/) with accession numbers NC_045512, AY278741, and KU740200, respectively. As expected, the 5′ UTRs of SARS-CoV-1 and SARS-CoV-2 showed a high similarity of 93% because they belong to the same clade (Figure S1A). However, the sequence comparison with MERS was found to be of low similarity of only 46% and also differs in terms of sequence length (Figure S1A). Despite these differences between the clades, several conserved regions or segments (7−14, 48−58, 140−147, 198−212, and 250−260) were noticeable in both SARS and MERS. On the other hand, the highest sequence similarity between the 5′ UTR region of SARS-CoV-1 and SARS-CoV-2 also implies the greater possibility of similar physical mode binding for the translational regulation and interaction with viral proteins.

We next looked for the structural architecture of 5′ UTR of SARS-CoV-2 RNA. It is known that like most other RNA viruses, the 5′ UTR region of SARS-CoV-2 poses the secondary (2D) and 3D structures, which are defined as the stems and loops of the 5′ UTR. The stem–loop regions of the 5′ UTR are acting as the platform and recruit the various viral and host proteins toward the regulation or incitation of translation and transcription processes. To understand these secondary structural features of complete 5′ UTR, we used two independent programs: SPOT-RNA (https://sparks-lab.org/server/spot-rna/) and RNAfold (http://rna.tbi.univie.ac.at/). (Similar results were obtained with other programs.) As shown in Figure S1B, the secondary structure of 5′ UTR of SARS-CoV-2 RNA comprises eight stem–loops (SL), among which SL3 and SL6 are the longest. The first SL1 region comprises 7−33 nucleotides which are aligned well with SARS-CoV-1 and similar to most other stem–loop regions, SL3 consists of 85−126 nucleotides, and SL6 has 187−219 nucleotides. It is
also noticeable that 99–108 nucleotides of SL3 pair with 212–220 nucleotides of SL6 to form the tertiary helical folds. The overall secondary fold and topology arrangement of the stem–loop regions of 5′ UTR has further validated the loop formation with the aid of base-pair probability and overall energy (kcal/mol) minimization values (Figure S1C).

It has been experimentally shown that the 10–28 nucleotides at the 5′UTR of SARS-CoV-1 and MERS physically interact with Nsp1. Interestingly, our MSA results are evident in demonstrating the highest sequence similarity, in particular to SARS-CoV-2 aligned with SARS-CoV-1. This raises the highest possibility in suggesting that the 10–28 nucleotides of the 5′UTR region of SARS-CoV-2 RNA could potentially interact with Nsp1. Hence, on the basis of our rationale and prediction we next modeled the 3D structure of the 5′UTR-SL1 region (7–33 nucleotides, with extended regions) of SARS-CoV-2 using two independent programs: RNAComposer (www.rnacomposer.cs.put.poznan.pl/) and simRNA (www.genesilico.pl/SimRNAweb/) for more validity (Figure 1D). The overall energy-minimized structure of SL1 resembles the Watson and Crick base pairing of dsDNA with a right-hand twisted or helical fold, which is due to the high complementary base pairing, suggesting high stability. This base pairing or double-stranded RNA formation is exception to 18–21 region due to low complementarity and results in apical loop formation, a region of greater physical interaction with the target proteins (Figure S1B,D).

In order to address the Nsp1 protein sequence similarities and differences among SARS-CoV-2, SARS-CoV, and MERS, the genomic sequences were retrieved from the curated NCBI virus database (https://www.ncbi.nlm.nih.gov-genome/viruses/) with accession numbers YP_009725297.1, NP_828860.2, and YP_009047213.1, respectively, and multiple sequence alignments were performed using ClustalOmega.

Figure 1. (A–C) EMSA measurement of Nsp1 affinity for SL1 of 5′UTR. Unlabeled SL1 RNA was tracked in native PAGE after incubation with increasing amounts of (A) Nsp1-FL or (B) Nsp1ΔCTD to compare their binding affinities. The $K_{D,app}$ values were determined as the protein concentration to affect half of the maximal binding to the SL1 RNA. (C) The binding curves showing a higher binding affinity of $K_{D,app} = 0.18 \mu M$ was observed for Nsp1-FL, and a lower affinity of $K_{D,app} = 0.31 \mu M$ was observed for Nsp1ΔC. The positions of free SL1 RNA and the respective Nsp1/SL1 complex are indicated. (D–F) Surface (white) and cartoon (blue) diagram showing the structure of the Nsp1/SL1 RNA complex from SARS-CoV-2. The bound RNA is shown in gray nucleotide base pairs with an orange phosphate backbone. The N- and C-terminal regions of Nsp1 and 5′ and 3′ of SL1 RNA are labeled accordingly. (G–I) Surface charge distribution of the Nsp1/SL1 RNA complex from SARS-CoV-2 (blue, negatively charged area; red, positively charged area). Front, orthogonal, and bottom views of the complex structure are shown for both the cartoon representation and the surface charge distribution.
The MSA of the Nsp1 protein sequence of SARS-CoV-2 shows a high similarity of 84.4% with that of SARS-CoV-1 (Figure S1A), suggesting that the Nsp1 protein of both the BetaCoVs has a higher chance of sharing the same physical and biological function and structure. Surprisingly, unlike the above S’UTR region compassion, only 17.7% sequence similarity of Nsp1 was found with MERS (Figure S1A) in comparison with SARS-CoV-1 and -2, indicating that the structure, function, and physical and chemical binding properties of Nsp1 in SARS-CoV-2 are shared with SARS-CoV-1 of the same clade and would most likely be different from that of the MERS.

The observed high sequence similarity of Nsp1 between SARS-CoV-2 and SARS-CoV-1 directly implicates the possible structural similarity, but to date there is no full-length (FL) atomic structure or model of any of the coronavirus structures available, which might be due to higher flexibility of the protein. This conceals the key molecular and functional aspects to better understand the viral and host molecules. The only available Nsp1 structure is from SARS-CoV-1 (NMR structure of low-resolution, PDB: 2HSX), which lacks the complete 60 amino acid residues of the C-terminal region (Nsp1-1_115).

Hence, our first step is building the robust and validated (energy minimum and optimized physical binding forces) virtual structure of full-length Nsp1 with the aid of homology modeling via SWISS-MODEL (www.swissmodel.expasy.org) using PDB: 2HSX as a template and denovo ab initio structure building through I-TASSER (www.zhanglab.ccmb.med.umich.edu) to obtain the full-length Nsp1 structure. The RMSD of the previously published partial structure of Nsp1 from SARS-CoV-1 and the Nsp1-FL structure of SARS-CoV-1 was 0.38, which suggests overall structural accuracy of the N-terminal domain. The Nsp1 structure (Figure S2) consists of two α-helices and five β-sheets with several loops. The N-terminal region has more rigid structures consisting of α-helices and two main antiparallel β-sheets and presumes to have an RNA binding activity due to the observed surface positive charge (Figure S2B–D). However, the C-terminal region is more discordant and consists of several long intermotif loops leading to the higher flexibility of the protein. The extreme C-terminal consists of two short α-helices connected with a long interdomain loop of 28a, making this region highly dynamic, and was very recently shown to play a role in binding to the 40S subunit of the host ribosome. The overall domains of Nsp1 arrange in forming a small groove or pocket-like structure encompassing the two antiparallel β-sheets (15aa–20aa and 117aa–122aa), which can accommodate the large molecule (Figure S2D). Interestingly, the surface charge distribution and physical properties of Nsp1 (Figure S2E to S2G) reveal that this region is highly positively charged, and we extend our supposition to the potential target region for RNA binding (negatively charged). Furthermore, the C-terminal or the 40S binding region of Nsp1 is in the far vicinity and negatively charged. Key residues 123–130aa of this region of Nsp1 also proved to be important in SARS-CoV-1 toward the viral RNA physical interaction, which is also highly aligned in our MSA (Figure S2A). Added to that, the presence of a β-sheet in the RNA binding sequence indicates that the RNA binding region is very stable. These observations next directed us to experimentally test the RNA binding activity of Nsp1 of SARS-CoV-2, underlining its physical interaction.

The affinity of viral Nsp1 homologues from different clades for different RNA segments (stem–loop regions) binding with a range affinities, some of which are dependent on the nucleotide lengths and some having sequence specificity, was considered. Meanwhile, which RNA segment of the 5′UTR region of SARS-CoV-2 binds with Nsp1 is elusive. As mentioned earlier, Nsp1 from both SARS-CoV-2 and SARS-CoV-1 have the highest homology along with the surface charge distribution and are anticipated to behave similarly. Concerning the RNA segment, 10–28 nucleotides corresponding to the SL1 region of SARS-CoV-1 and MERS-CoV were experimentally proven to interact with Nsp1, and in our alignment, that particular region of SARS-CoV-2 aligned well with SARS-CoV-1 (99% identity) and MERS-CoV (47% identity), suggesting that the 10–28 nucleotides of SARS-CoV-2 RNA potentially interact with Nsp1. We therefore sought to assess the potential interaction between the Nsp1 of SARS-CoV-2 and the SL1 region of RNA using the enzyme mobility shift assay (EMSA). (Refer to the Supporting Information for the experimental procedure). As expected, the in vitro translated RNA segment of SL1 of S’UTR directly and physically binds to Nsp1 with a greater affinity of 0.18 μM with full-length purified protein (Figure 1A,C). This further corroborates that Nsp1 of SARS-CoV-2 directly binds to the SL1 region of viral RNA, which is relatively similar as observed in SARS-CoV-1.

We next used the C-terminal truncated construct of Nsp1 (1–133) (Nsp1ΔC) to test its binding affinity with the same RNA segment because the C-terminal domain may not be essential for RNA binding. However, we notice a nearly 1 order reduced binding affinity between Nsp1ΔC and SL1, and its binding affinity was found to be 0.31 μM with saturated binding at a 1:2 molar concentration (Figure 1B,C). Less smearing of the RNA/protein complex was noticed with Nsp1ΔC, which could be the reason that the truncated form is more stable because the C-terminal domain connected to the linker region is highly flexible and may result in a lower binding affinity.

To better understand the structural and molecular mode of physical interactions between SARS-CoV-2 Nsp1 and SL1, we performed virtual structural studies using molecular dynamics and computational-model-based selective docking and simulation. To this end, we used our validated and resolved structure of Nsp1-FL (Figure S2B–D) and modeled RNA structure SL1 of S’UTR (Figure S1D). With these individual structures, we used three independent servers—ModelX (RNA–protein interactions suite) (www.modelx.crg.es/models#max), MDock (www.hdock.phys.hust.edu.cn), and HADDOCK 2.4 (www.wenmr.science.uu.nl/haddock2.4) for further validation and precision of the docking mode and interaction (Figure S3). Among the four possible clusters, clusters 1 and 2 of the docked complex have the highest HADDOCK score and a larger reproducible cluster size with the lowest possible RMSD, suggesting the large possibility of the true structure. This confidence was also further enhanced by the observed lowest binding free energies (hydrogen bonding, van der Waals, and ionic interactions) in cluster/model 1, which makes us consider the selection of the best possible model (Figures S3 and S4). Cluster 2 was found to be the docking of RNA to Nsp1 in the reverse direction with the same conserved interface. Further validation and refinement were completed by ensuring that the residues occupied Ramachandran favored positions using Coot (www.mrc-imb.cam.ac.uk/). The final docked complex structure was then compared with the initial Nsp1-alone modeled structure, and their overall RMSD was found to be 0.32 Å for Cα (central
carbon) atoms. This suggests that there are no large conformational changes upon docking. All 3D structures were visualized, and figures were generated using PyMol software.

The overall docked Nsp1−RNA complex structure shows Nsp1 directly binding to SL1 RNA in a clap-like fashion (Figure 1D−F), where the two long antiparallel b-sheet interfaces sit over the RNA helix of SL1. The binding of Nsp1 is also observed to induce the physical bending of RNA (14.3°), which could have helped in enhanced physical and direct interaction with higher affinity. The apical/loop region of SL1 is oriented toward the side of N- and C-termini (Figure 1F). Furthermore, the binding of Nsp1 over the SL1 RNA creates a large burred interface of ∼943 Å², as calculated from the PISA server (https://www.ebi.ac.uk/pdbe/pisa/) (Figure S5). This suggests a bona fide tight interaction between the Nsp1 and SL1 RNA of 5′UTR, which is also evidenced by functional binding studies (Figure 1A,B).

On the other hand, the surface charge distribution of the Nsp1/SL1 complex shows that most of the RNA is bound to the positively charged region of the protein (Figure 1G−I). The C-terminal region comprising a negatively charged surface is least associated and positioned away from the RNA binding groove, implying less chance for physical interaction. The observed positively charged region of Nsp1 and its binding to a groove/pocket in a clap-like fashion with the experimental binding studies corroborates the mode of viral RNA recognition by the Nsp1 toward its translational regulation.

We next looked closely at the key residues involved in building physical interaction between Nsp1 and SL1 (Figure 2). The overall interaction between the molecules is mediated by several hydrogen bonding, van der Waals, and ionic interactions. It is evident from the complex structure that the RNA molecule intercalates into the groove-like structure of the Nsp1 protein and makes its direct interaction through the residues mainly from aa11 to aa17, aa118 to 130, and aa144 to 148 (Figure 2 and Figure S5). Among these interacting residues, T12, Y118, R124, K125, N128, K129, L141, and D147 are found to hydrogen bond with the SL1 RNA, and several other residues of Nsp1 are found to be in a close proximity to the RNA, causing strong ionic interactions. In particular, the positively charged long-chain Arg and Lys amino acid residues (K11, R124, K125, and K141) intercalate into the RNA groove (Figure 2) and build strong ionic interactions with the negatively charged phosphate backbone of the RNA, thus causing the interaction to have greater affinity as observed in Figure 1A,C. Previous studies concerning SARS-CoV-1 also experimentally showed that R124 and K125 play a crucial role in viral RNA recognition,29,31 which is also consistent as observed in our complex structure in addition to other potential residues involved in the physical interaction.

To unravel the binding affinity and importance of key hydrogen-bonding residues involved in the interaction, we performed binding energy calculations virtually by mutating individual amino acid residues of Nsp1 and docking or simulating to measure the overall binding free energy (kcal) for...
Figure 3. (A and B) Intermolecular connectivity of the Nsp1/SL1 complex in the real-time dynamic state. Nodes are located at the positions of Cα atoms (for amino acids of Nsp1), and phosphate backbone C4′ and C2 atoms (for nucleotides of SL1 RNA) are taken to build the node connectivity map within the cutoff distance of 7.3 Å. Nsp1 is shown by dark-green spears, and SL1 RNA is shown by light-green spears. (A) The interacting region or interphase during the dynamic state is shown by cyan intermolecular lines. (B) The amino acid pair connections of Nsp1 are shown by light-gray dots, those between nucleotide pairs of 5′UTR RNA are shown as yellow, and inter-amino acid—nucleotide paired clusters are shown by cyan dots. Each dot represent one amino acid or nucleotide pairing at a particular location. (C) Molecular dynamics simulation studies showing the oscillation and B-factor (stability factor, as low a value as stable) profiles of the Nsp1/SL1 complex. The amino acid residue position is shown on the X axis, and the degree of movement of amino acids as a B-factor is shown on the Y axis. (D) Domain separation dynamics of the Nsp1/SL1 complex. Low and studied eigenvectors for RNA and Nsp1 (aa1-aa120) are noticeable, indicating the greater stability of the complex for up to aa120 of Nsp1. Abrupt and higher eigenvectors were observed with the C-terminal domain (aa121-aa180). Increased eigenvectors are directly linked with oscillation or lead to the disjoining conformation of Nsp1-CTD. (F) Mobility scale and stable-complex-forming region. Highly stable and less stable residues are shown in blue and red, respectively.
the individual mutant. As shown in Figure 2D, the binding kinetics and biophysical characterization show that wild-type Nsp1 binds to SL1 RNA with a free energy of $-68$ kcal and that most of the mutants showed a noticeable reduction in the binding affinity with the RNA (Figure 2D). Specifically, the mutations on Y118, R124, K125, and K141 showed a 2- to 3-fold depletion in the binding affinity and its DG ranged from $-24$ to $-40$ kcal, suggesting the weaker interaction or disruption of the physical interaction with the mutant protein. This further corroborates the true binding affinity and efficacy of Nsp1 with RNA and highlights the key amino acid residues involved in this physical interaction in the Nsp1/SL1-RNA complex.

To validate the overall complex structure and real-time in-solution behavior of the Nsp1-SL1 RNA complex, we performed virtual biophysical experiments using molecular dynamics and simulations using the DynOomics 30 server (www.gnm.csb.pitt.edu) and LARMD (www.chemyang.ccnu.edu). The time course molecular simulations for 10 ns of dynamics were recorded. B-factor profiles (thermal stability factor, with the lowest number indicating high stability) and domain separation analysis combined with simulation studies were performed using the DynOomics server and validated with Schrodinger molecular dynamics tertiaries. Our extended biophysical molecular dynamics and simulation studies also principally suggest that the overall Nsp1-SL1 RNA complex is
Figure 3D, as expected that very small eigenvectors were observed for the whole complex-forming region (eigenvector score 0.01) and the N-terminal region of Nsp1 (aa1-aa115) alone and most used drugs were either unrationalized or very small molecules, which have less specificity. Structural and functional analyses of potential and specific drugs that could block the function of RNA interaction using Nsp1-FL from SARS-CoV-2 are lacking and warrants further study. Hence, we next investigated the most possible and specific drugs or inhibitors which could not only block the Nsp1 function but also structurally block the interaction with the RNA (based on the binding modes observed in our current work). On the basis of the chemistry of the inhibitor, the charge and surface distribution of the Nsp1 site, and the observed binding mode of RNA to Nsp1, we first screened the binding of high potential inhibitors by docking (arbitrary docking was also performed to overcome the docking bias) with Nsp1-FL (Figure 4). Among the several screened drugs, glycyrrhizic acid, lobarian acid, garcinolic acid, and tirilazad are found to bind Nsp1 with greater affinity and also are found to structurally impede the physical interaction between Nsp1 and SL1 (Figure 4). The interaction between potential Nsp1 and SL1 complex. The time-course simulation of the complex for 10 ns was recorded, and the mobility scale (Figure 3F and Movie S1 and S2) also shows that the interaction interphase is less mobile and the outer residues of the complex show higher mobility, thus providing high confidence in Nsp1 and SL1 complex formation and physical stability.

Inhibiting or blocking the function of Nsp1 in any virus has several implications for the therapeutic in regulating viral replication. In this regard, some studies have experimentally demonstrated the binding affinity of drugs with the Nsp1 of different viruses belonging to Coronaviridae, Togaviridae, and others33,34 but the mode of action and structural studies are elusive. Recently, in-silico studies were reported concerning potential herbal drugs or plant extracts that could randomly bind to Nsp1 of SARS-CoV-2.35,36 However, these in-silico studies conducted using the short N-terminal domain (aa1-aa115) alone and most used drugs were either unrationalized or very small molecules, which have less specificity. Structural and functional analyses of potential and specific drugs that could block the function of RNA interaction using Nsp1-FL from SARS-CoV-2 are lacking and warrants further study. Hence, we next investigated the most possible and specific drugs or inhibitors which could not only block the Nsp1 function but also structurally block the interaction with the RNA (based on the binding modes observed in our current work). On the basis of the chemistry of the inhibitor, the charge and surface distribution of the Nsp1 site, and the observed binding mode of RNA to Nsp1, we first screened the binding of high potential inhibitors by docking (arbitrary docking was also performed to overcome the docking bias) with Nsp1-FL (Figure 4). Among the several screened drugs, glycyrrhizic acid, lobarian acid, garcinolic acid, and tirilazad are found to bind Nsp1 with greater affinity and also are found to structurally impede the physical interaction between Nsp1 and SL1 (Figure 4). The interaction between potential Nsp1
inhibitors is mediated by several polar, hydrophobic, and salt-bridge interactions. It is quite promising that the candidate drugs and target RNA share the same binding region at very high affinity. The binding ΔG values of these drugs to Nsp1-FL range from −10.4 to −8.6 kcal/mol, as estimated computationally. This suggests that these drugs could potentially act by inhibiting the Nsp1 interaction with the SL1 of 5′UTR in the SARS-CoV-2 genome or any other target host or viral RNA.

In summary, our reported first full-length structure of Nsp1 from SARS-CoV-2 extends our understanding of the viral leader protein, which plays several vital roles in the virulence and in combating with host cell factors. The structural docking, biological experimental results with molecular dynamics studies demonstrate that Nsp1 from SARS-CoV-2 directly interacts with viral SL1 RNA of 5′UTR at higher affinity (ΔG−70 kcal and 0.18 μM) and forms a stable complex for further translational regulation. The drug-binding studies analysis shows that glycyrrhizic acid, lobarian acid, garcinolic acid, and tirilazad bind to Nsp1 with higher affinity and can impede the Nsp1 physical interaction with viral RNA. Meanwhile, it is now established that Nsp1 physically binding to viral RNA in SARS-CoV-2 has implications in translation regulation with the 40S ribosome and that this specific interaction enhances viral gene replication.\(^{30}\) It has been shown that the partial mutation in Nsp1 of mouse hepatitis virus (MHV) produces an attenuated virus,\(^ {30}\) which has key implications in vaccine development. Understanding the physical and chemical aspects of viral RNA and Nsp1 interaction via various interaction forces (hydrogen bonding, van der Waals, electrostatic, etc.) and their binding affinities has great implications for potential drug targets. Furthermore, the structural and mechanism of SARS-CoV-2 Nsp1 and its interplay with viral RNA and the ribosome will open the arena of exploring the development of live attenuated vaccines (Figure 5). This could be similar to the situation for the Sabin vaccine for poliovirus with mutations in the 5′UTR region resulting in abolishing viral translation and thus leading to the development of the current polio vaccine.\(^ {31}\) Since Nsp1 also physically and directly binds to its viral RNA in the 5′UTR region, possible mutations in the 5′UTR region or Nsp1 of SARS could hold potential for viral attenuation and might aid the development of a vaccine. This strategy can be exploited to develop a SARS-CoV-2 vaccine and needs further study for the further development of potential therapeutics. In support, we also notice that mutations in Nsp1 had a decrease binding affinity with 5′UTR (Figure 2D and Figure S5). On the basis of our structural, functional, and molecular dynamics studies along with a screening of potential drugs, we designed the schematic representation addressing the potential role of the Nsp1/SL1 complex in translational regulation and the dissection of physical and chemical aspects (Figure 5).

**Associated Content**

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcl.0c02818.

Sequence alignment and structure of SARS-CoV-2 5′UTR and SARS-CoV-2 Nsp1; validation report for the predicted structure of Nsp1 in complex with the SL1 region of 5′ UTR of SARS-CoV-2 RNA; validation report of model scores and binding energy of the predicted complex structure; and intermolecular interactions in the Nsp1/SL1 complex and buried interface (PDF)

Dynamic simulation and stability of the Nsp1/SL1 RNA complex, front view (MOV)

Dynamic simulation and stability of the Nsp1/SL1 RNA complex, lateral view (MOV)

**Author Information**

Corresponding Author

Naveen Vankadari — Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology and Faculty of Medicine, Nursing and Health Science, Monash University, Clayton, Victoria 3800, Australia; orcid.org/0000-0001-9363-080X; Email: Naveen.vankadari@monash.edu

**Authors**

Nandhani Nisha Jeyasankar — Faculty of Medicine, Nursing and Health Science, Monash University, Clayton, Victoria 3800, Australia

Wilma Jerom Lopes — Faculty of Medicine, Nursing and Health Science, Monash University, Clayton, Victoria 3800, Australia

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcl.0c02818

**Author Contributions**

N.N.J. and W.J.L. contributed equally.

**Author Contributions**

N.V. conceived the study, led the work, and prepared the figures for structural and drug interaction studies. N.V., W.J.L., and N.N.J. cowrote the paper and performed molecular docking, modeling, and dynamics studies. W.J.L. and N.N.J. prepared figures for sequence alignment. N.V. produced the RNA for the studies, formed a protein/RNA complex for analysis, and performed the EMSA experiments. N.V. performed data validation, and all authors analyzed the results and drafted the study.

**Notes**

The authors declare no competing financial interest.

**Acknowledgments**

We thank the Monash University Software Platform for license access to the concerned software. We thank Dene Litter of Monash University for providing the purified proteins of Nsp1-FL and Nsp1AΔC.

**References**

1. Chaw, S. M.; Tai, J. H.; Chen, S. L.; Hsieh, C. H.; Chang, S. Y.; Yeh, S. H.; Yang, W. S.; Chen, P. J.; Wang, H. Y. The origin and underlying driving forces of the SARS-CoV-2 outbreak. *J. Biomed. Sci.* 2020, 27 (1), 73.

2. Andersen, K. G.; Rambaut, A.; Lipkin, W. I.; Holmes, E. C.; Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* 2020, 26 (4), 450–452.

3. Astuti, I.; Ysrafil. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. *Diabetes Metab Syndr* 2020, 14 (4), 407–412.

4. Diaz, J. SARS-CoV-2 Molecular Network Structure. *Front. Physiol.* 2020, 11, 870.

5. Chan, J. F.; Kok, K. H.; Zhu, Z.; Chu, H.; To, K. K.; Yuan, S.; Yuen, K. Y. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical
pneumonia after visiting Wuhan. *Emerging Microbes Infect.* 2020, 9 (1), 221–236.

(6) Huston, N. C.; Wan, H.; Araujo Tavares, R. C.; Wilen, C.; Pyle, A. M. Comprehensive in-vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *bioRxiv* 2020, DOI: 10.1101/2020.07.10.197079.

(7) Zhang, Y.; Kutateladze, T. G. Molecular structure analyses suggest strategies to therapeutically target SARS-CoV-2. *Nat. Commun.* 2020, 11 (1), 2920.

(8) Masters, P. S. The molecular biology of coronaviruses. *Adv. Virus Res.* 2006, 66, 193–292.

(9) Chen, Y.; Liu, Q.; Guo, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J. Med. Virol.* 2020, 92 (4), 418–433.

(10) Wang, H.; Li, X.; Li, T.; Zhang, S.; Wang, L.; Wu, X.; Liu, J. The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur. J. Clin. Microbiol. Infect. Dis.* 2020, 39, 1629.

(11) Vankadari, N. Overwhelming Mutations or SNPs of SARS-CoV-2: A Point of Caution. *Gene* 2020, 752, 144792.

(12) Vankadari, N.; Wilce, J. A. Emerging WuHAn (COVID-19) coronavirus: glycanshield and structure prediction of spike glycoprotein and its interaction with human CD26. *Emerging Microbes Infect.* 2020, 9 (1), 601–604.

(13) Hasan, A.; Paray, B. A.; Hussain, A.; Qadir, F. A.; Attar, F.; Aziz, F. M.; Sharifi, M.; Derakhshankhah, H.; et al. A review on the cleavage priming of the spike protein on coronavirus by angiotensin-converting enzyme-2 and furin. *J. Biomol. Struct. Dyn.* 2020, 1–9.

(14) Vankadari, N. Structure of Furin Protease Binding to SARS-CoV-2 Spike Glycoprotein and Implications for Potential Targets and Virulence. *J. Phys. Chem. Lett.* 2020, 11 (16), 6655–6663.

(15) Amawi, H.; Abu Deibah, G. I.; AA, A. A.; Dua, K.; Tambuwala, M. M. COVID-19 pandemic: an overview of epidemiology, pathogenesis, diagnostics and potential vaccines and therapeutics. *Ther. Delivery* 2020, 11 (4), 245–268.

(16) Vankadari, N. Arbidol: A potential antiviral drug for the treatment of SARS-CoV-2 blocking trimerization of the spike glycoprotein. *Int. J. Antimicrob. Agents* 2020, 56, 105998.

(17) Shaffer, L. 15 drugs being tested to treat COVID-19 and how they would work. *Nat. Med.* 2020, DOI: 10.1038/s41591-020-00019-9.

(18) Calina, D.; Docea, A. O.; Petracis, D.; Egorov, A. M.; Ishmukhametov, A. A.; Gabibov, A. G.; Shiltiman, M. I.; Kostoff, R.; et al. A. Towards effective COVID19 vaccines: Updates, perspectives and challenges (Review). *Int. J. Mol. Med.* 2020, 46 (1), 3–16.

(19) Yoshimoto, F. K. The Proteins of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2 or n-COV19), the Cause of COVID-19. *Protein J.* 2020, 39 (3), 198–216.

(20) Mirza, M. U.; Froeyen, M. Structural elucidation of SARS-CoV-2 vital proteins: Computational methods reveal potential drug candidates against main protease, Nsp12 polymerase and Nsp13 helicase. *J. Pharm. Anal.* 2020, 10, 320.

(21) Calligari, P.; Bobone, S.; Ricci, G.; Bocedi, A. Molecular Investigation of SARS-CoV-2 Proteins and Their Interactions with Antiviral Drugs. *Viruses* 2020, 12 (4), 445.

(22) Bianchi, M.; Benvenuto, D.; Giovannetti, M.; Angeletti, S.; Ciccozzi, M.; Pascarella, S. SARS-CoV-2 Envelope and Membrane Proteins: Structural Differences Linked to Virus Characteristics? *BioMed Res. Int.* 2020, 2020, 4389089.

(23) Kamitani, W.; Huang, C.; Narayanan, K.; Lokugamage, K. G.; Makino, S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.* 2009, 16 (11), 1134–40.

(24) Tohya, Y.; Narayanan, K.; Kamitani, W.; Huang, C.; Lokugamage, K.; Makino, S. Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. *J. Virol.* 2009, 83 (10), 5282–8.

(25) Narayanan, K.; Huang, C.; Lokugamage, K.; Kamitani, W.; Ikegami, T.; Tseng, C. T.; Makino, S. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.* 2008, 82 (9), 4471–9.

(26) Huang, C.; Lokugamage, K. G.; Rozovics, J. M.; Narayanan, K.; Semler, B. L.; Makino, S. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.* 2011, 7 (12), No. e1002433.

(27) Lokugamage, K. G.; Narayanan, K.; Huang, C.; Makino, S. Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. *J. Virol.* 2012, 86 (24), 13598–608.

(28) Thoms, M.; Buschauer, R.; Ameisimeier, M.; Koepke, L.; Denk, T.; Hirschenberger, M.; Kratzat, H.; Hayn, M. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* 2020, 369, 1249.

(29) Tanaka, T.; Kamitani, W.; DeDiego, M. L.; Enjuanes, L.; Matsuura, Y. Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.* 2012, 86 (20), 11128–37.

(30) Zost, R.; Cervantes-Barragan, L.; Kuri, T.; Blakgori, G.; Weber, F.; Ludewig, B.; Thiel, V. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. *PLoS Pathog.* 2007, 3 (8), No. e109.

(31) Terada, Y.; Kawachi, K.; Matsuura, Y.; Kamitani, W. MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA. *Virology* 2017, 511, 95–105.

(32) Tahir ul Qamar, M.; Alqahtani, S. M.; Alami, M. A.; Chen, L-L. Structural basis of SARS-CoV-2 3CL(pro) and anti-COVID-19 drug discovery from medicinal plants. *J. Pharm. Anal.* 2020, 10, 313.

(33) Feibelman, K. M.; Fuller, B. P.; Li, L.; LaBarbera, D. V.; Geiss, B. J. Identification of small molecule inhibitors of the Chikungunya virus nsP1 RNA capping enzyme. *Antiviral Res.* 2018, 154, 124–131.

(34) Ferreira-Ramos, A. S.; Li, C.; Eydoux, C.; Contreras, J. M.; Morice, C.; Querat, G.; Gigante, A.; Perez Perez, M. J.; Jung, M. L.; et al. Approved drugs screening against the nsP1 capping enzyme of Venezuelan equine encephalitis virus using an immuno-based assay. *Antiviral Res.* 2019, 163, 59–69.

(35) Sharma, A.; Tiwari, V.; Sowdhamini, R. Computational search for potential COVID-19 drugs from FDA-approved drugs and small molecules of natural origin identifies several anti-virals and plant products. *J. Biosci.* 2020, 45 (1), 100.

(36) de Lima Menezes, G.; da Silva, R. A. Identification of potential drugs against SARS-CoV-2 non-structural protein 1 (nsp1). *J. Biomol. Struct. Dyn.* 2020, 1–11.

(37) Gutierrez, A. L.; Denova-Ocampo, M.; Racaniello, V. R.; del Angel, R. M. Attenuating mutations in the poliovirus S’ untranslated region alter its interaction with polypyrimidine tract-binding protein. *J. Virol.* 1997, 71 (5), 3826–33.