Computational investigation of structural dynamics of SARS-CoV-2 methyltransferase-stimulatory factor heterodimer nsp16/nsp10 bound to the cofactor SAM

Md Fulbabu Sk[#], Nisha Amarnath Jonniya[#], Rajarshi Roy, Sayan Poddar, Parimal Kar*

Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Khandwa Road, MP 453552, India

[#]Equal contribution

*Corresponding author: Parimal Kar| Email: parimal@iiti.ac.in| Phone: +91 731 2438700 (Ext. 550)

Abstract

Recently, a highly contagious novel coronavirus (COVID-19 or SARS-CoV-2) has emerged as a global threat in people's health and global economies. Identification of the potential targets and development of a vaccine or antiviral drugs is an urgent demand. The 5'-capping mechanism of eukaryotic mRNA and some viruses such as coronaviruses (CoVs) are essential for maintaining the RNA stability, protein translation, and for viral immune escape. SARS-CoV encodes S-adenosyl-L-methionine dependent (SAM) methyltransferase (MTase) enzyme characterized by nsp16 (2'-O-MTase) for generating the capped structure. The present article highlights the binding mechanisms of nsp16 and nsp10 to identify the role of nsp10 in MTase activity. Furthermore, the conformational dynamics and energetic behind the SAM binding to nsp16 in its monomer and dimer form was analyzed by using an extensive molecular dynamics simulation along with the Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA). Our results show that the presence of nsp10 increases the favorable van der Waal and electrostatic interactions between the SAM and nsp16, thus nsp10 acts as a stimulator for its strong binding. The interaction profile suggests that hydrophobic interactions were predominately identified for protein-protein interactions. Also, the stable hydrogen bond between Ala83 (nsp16) and Tyr96 (nsp10), and between Gln87 (nsp16) and Leu45 (nsp10) plays a vital role in the nsp16-nsp10 interface. Further, Computational Alanine Scanning (CAS) mutagenesis was performed, which revealed hotspot mutants, namely I40A, V104A, and R86A for the dimer association. Therefore, the dimer interface of nsp16/nsp10 could also be a potential target to suppress the 2'-O-MTase activity of SARS-CoV-2. Overall, our study provides a comprehensive understanding of the dynamic and thermodynamic process of binding of nsp16 and nsp10 that will contribute to the novel design of peptide inhibitors based on nsp16.

Introduction

Coronaviruses are considered as an etiological agent for causing the human severe acute respiratory syndrome. In the past two decades, it was responsible for the epidemic in 2003 and 2012 by the SARS-CoV, and the middle east respiratory syndrome coronavirus (MERS-CoV), respectively. Recently, in December 2019, the outbreak of novel coronavirus designated as COVID-19 or SARS-CoV-2 emerged in the city of Wuhan, China [1-3]. The rapid worldwide spread and threat imposed by the ongoing coronavirus "COVID-19," has been declared as a global emergency by the World Health Organization (WHO). So far, SARS-CoV-2 has infected approximately 10 million people and accounts for ~ 0.5 million death in the world. Many nations are putting their best efforts in controlling and preventing the highly contagious COVID-19, which has impacted the global health and economic. Despite, neither vaccine nor any direct antiviral drugs are available for the effective treatment of the human and animal coronavirus [4-6].

Coronaviruses (CoVs) are enveloped viruses having a single-stranded positive-sense RNA genome [7, 8]. They belong to the family *Coronaviridae* of the order Nidovirales and possess the largest genome (26.4 to 31.7 kb) among the RNA viruses [9]. They are classified into four genera, namely *Alphacoronaviruses*, *Betacoronaviruses*, *Gammacoronaviruses,* and *Deltacoronaviruses* [10]. CoVs can infect both humans and animals [11, 12], and can cause diseases like Hepatitis & Pneumonitis in mouse [13] and neurologic & respiratory diseases in humans [14, 15]. Currently, seven different strains of human coronaviruses (HCoV) have been reported, including the strains of 229E and NL63 (*Alphacoronaviruses*), and OC43, HKU1, SARS-CoV (2002), MERS-CoV (2012), and SARS-CoV-2 (2019) (*Betacoronaviruses*) [16, 17]. SARS-CoV-2 is closely related to bat-CoV RaTG13 with a 96.2% genome sequence identity suggesting its possible evolution from bat CoVs [18, 19]. Among the available human CoV, SARS-CoV-2 shows nearly 80% sequence identity with SARS-CoV, whereas only 50% identity with the MERS-CoV [19-21]. Currently, most of the therapeutic options that are available for controlling the COVID-19 is based on the previous knowledge and information gathered from the SARS-CoV and MERS-CoV.

The 5' end of the eukaryotic cellular mRNA and most viral genomic mRNA possess a cap structure, which plays a vital role in mRNA splicing, translation initiation, RNA stability, and intracellular RNA transport [22]. The capping of the 5' end mRNA occurs through a sequential enzymatic process involving three enzymes: RNA triphosphatase (TPase), RNA

guanylyltransferase (GTase), and RNA (guanine-N7) -methyltransferase (N7-MTase), generating a cap-0 structure (m7GpppN). In higher eukaryotes and some viruses, the cap-0 structure (m7GpppN) is further methylated at ribose 2'-O position of mRNA by the 2'-O-methyltransferase (2'-O-MTase) to generate the cap-1(m7GpppNm) and cap-2 structure (m7GpppNmNm). This mimicking of the eukaryotic mRNA capping mechanism ultimately helped the virus to invade the host innate immune system [22, 23]. Both the MTase uses S-adenosyl-L-methionine (SAM or AdoMet), which acts as a methyl donor to the RNA substrate and generate S-adenosyl-L-homocysteine (SAH or AdoHcy) as a by-product.

The SARS-CoV-2 genome consists of 10 ORF (Open Reading Frame). ORF1ab encodes the replicase polyprotein 1ab (PP1ab), which gets cleaved by the two viral proteases PL[pro] and 3CL[pro] at N-terminus and C-terminus respectively to form all the 16 non-structural proteins (nsp1, nsp2, nsp3 by PL[pro] and nsp4-nsp16 by 3CL[pro]). It includes RNA processing enzymes such as RNA-dependent RNA polymerase (nsp12), RNA helicase and triphosphatase (nsp13), N7-MTase (nsp14), endonuclease (nsp15), and 2'-O-MTase (nsp16). The remaining ORFs are associated with encoding viral structural proteins such as spike (S) protein, envelope (E) protein, membrane (M) protein, nucleocapsid (N) protein, and other accessory proteins [19, 24-26]. Previous structural and biochemical characterization studies on SARS-CoV (2002) showed that in nsp10-nsp16/SAM complex, nsp10 acts as a stimulatory factor to execute the methylase activity of nsp16 and is a common mechanism among coronaviruses. [23] It might also increase the SAM-binding pocket stability and extend the RNA-binding groove of nsp16 [27].
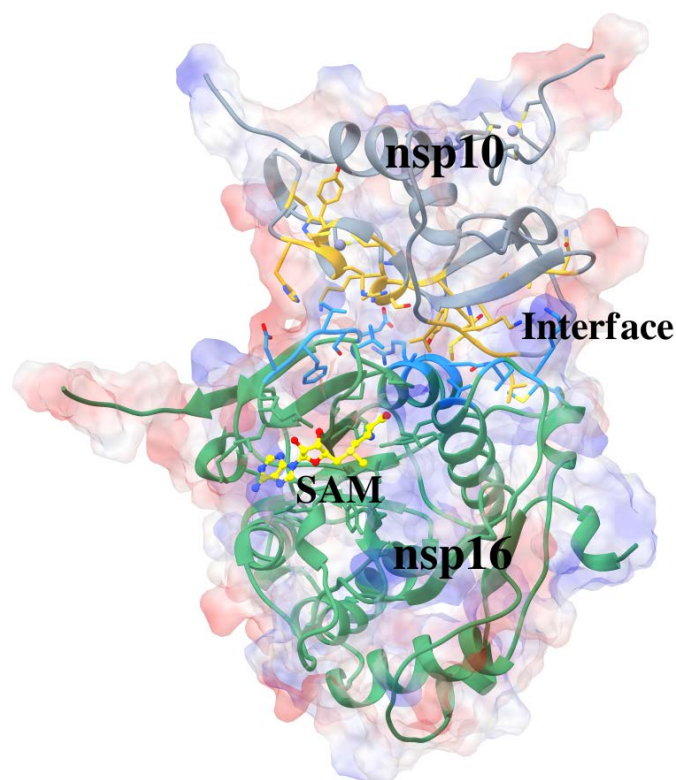
Figure 1: Crystal structure of SARS-CoV-2 nsp10-nsp16 dimer complexed with cofactor S-Adenosyl Methionine (SAM) (PDB: 6W4H). The nsp10 monomer, interacting part of nsp10, interacting part of nsp16, nsp16 monomer residues, and cofactor SAM is shown in color teal, gold, dark sky blue, light green, and yellow, respectively. The surface represents an electrostatic surface.

The present study involves the analysis of protein-protein interaction in the dimeric structure formed between nsp10 and nsp16 of SARS-CoV-2. The recently solved crystal structure of SARS-CoV-2 nsp16-nsp10 complex (PDB: 6W4H) [28] [29] was used in this study and shown in **Figure 1**. Here, we have used the Molecular Dynamics (MD) simulation technique in conjunction with Molecular-Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method to identify the critical residues involved in the complex formation. In addition, the monomeric simulation of nsp16/SAM of SARS-CoV-2 also helps in determining the structural variations and intermolecular interactions between the nsp16 and SAM molecule in the presence or absence of the nsp10, which may explain the role of nsp10 in the binding of nsp16-SAM in COVID-19. The detailed structural analysis and inter-molecular interactions revealed the interface of nsp16/ns10, which is unique for coronaviruses and thus may provide an attractive target in the development of specific antiviral drugs for controlling the pandemic SARS-CoV-2.

Simulation Methods

Structure Preparation

An experimental coordinate of the SARS-CoV-2 nsp16-nsp10 complex was retrieved from the Protein Data Bank (PDB), which is crystallized at 1.8Å (PDB ID 6W4H)[28, 29]. The complex structure includes the S-adenosyl-L-methionine (SAM) molecule. Before the simulation, protonation states were determined using PROPKA 3.1 [30] and modified the corresponding residues accordingly. Simulations were conducted to elucidate the role of nsp10 in the binding of SAM to nsp16. Also, the conformational differences in the nsp16 structure in its monomeric form were compared with the dimeric form when it binds with nsp10. The following systems were used for the study: dimer (nsp16/SAM/nsp10) and monomers such as (nsp16/SAM, i.e nsp16$_{mono}$) and monomer (nsp10$_{mono}$).

For the monomeric simulations system, nsp16 and nsp10 were separated from the complex and simulated under the same protocol as of the nsp10-nsp16 complex. Missing hydrogens in the crystal structure were built using the *Leap* module of the Ambertools19 [31], and the proper amount of Na$^+$ ions were added to neutralize all the systems. All the systems were solvated into periodic octahedron TIP3P water box [32] with a 10Å buffering distance from all directions. All the simulations were performed using *pmemd.cuda* module of the Assisted Model Building with Energy Refinement (AMBER) program (version 18) [31] in a Linux-based GPU workstation having four GTX 2080-Ti card.

Simulation Protocol

Simulations of the dimer and monomers were carried out by the amber ff14SB force field [33]. The time step in these three simulations was fixed to 2.0 fs. Bond lengths having hydrogen atoms were kept fixed using the SHAKE algorithm [34], and Particle-Mesh Ewald summation (PME) [35] was also used to compute long-range interaction with a non-bonded cut off 10 Å. At first solvated complexes were subjected to energy minimization using 500 steps of steepest descent followed by 500 steps of conjugant gradient algorithm. In these steps, amino acids were kept fixed using restraint force, having a force constant of 2.0 kcal mol^{-1}Å^{-2}. Secondly, the entire systems were minimized without any restraint on the solute using the steepest descent algorithm followed by the conjugant gradient algorithm. After that, systems were gradually heated from 0 K to 300 K in the NVT ensemble, where protein atoms are fixed using a force

constant of 2.0 kcal mol^{-1}Å^{-2}. Subsequently, systems were simulated up to 50 ps in 300 K at a constant pressure of 1 atm using Berendsen Barostat [36], having the same restraint on the solute atoms. Before the production run, we conducted 1 ns equilibration in the NPT ensemble without any restraint in the systems. Finally, all three systems were subjected to the production run for 1 μs, each using *pmemd.cuda* module of AMBER18 [31]. Overall, 100,000 snapshots were generated and used for the analysis using *Cpptraj* module [37] of AMBER18 [31].

Trajectory analysis

All the analysis, such as root means square deviation (RMSD), the radius of gyration (R$_g$), solvent accessible surface area (SASA), and other analysis were done using the *Cpptraj* module [37] of AMBER18. A distance of ≤ 3.5 Å and an angle cut-ff of ≥120º were used for the hydrogen bond calculations.

Principal component analysis

One of the widely used unsupervised data reduction schemes is the principal component analysis (PCA) [38]. Herein, the purpose of applying the PCA scheme is to identify the similarity and dissimilarity of nsp16/nsp10 dimerization and highlights the low energy conformations in the collected structural and energetics MD trajectory data [39]. This technique diagonalized the covariance matrix obtained from the atomic fluctuation of all residue present in the system and represents the motion in terms of their eigenvectors and eigenvalues [40]. These eigenvectors are also termed as principal components (PC) and indicate the direction of the movement, and the amplitude of motions is given by their respective eigenvalue. PCA analyses were done by the AMBER18 *Cpptraj* module, and the same protocol has been used in our previous studies [41, 42].

Residual network analysis of protein

The residual network analysis approach has been widely used to explore the viral fitness and resistance development of protein structure. The network analysis of protein structures (NAPS) [43] server (http://bioinf.iiit.ac.in/NAPS/) was used to identify key residue interactions in the residual network and the network-based hydrophobic contacts at different snapshots from the simulation.

Binding Free energy and alanine scanning

The interaction energy between nsp10 and nsp16, as well as for the SAM molecule to nsp16 in both monomer and dimer form, was calculated using the molecular mechanics Poisson-

Boltzmann surface area (MM-PBSA) methodology [44-50]. MMPBSA.py script available in the AmberTools19 was used for the analysis. A total of 1000 snapshots were used in all three cases for estimating the binding free energies. Details of the MMPBSA protocol were discussed in our previous studies [49, 51-57], and hence we follow the same protocol here.

In the MMPBSA scheme, binding free energy is estimated by these following equations:

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S \tag{1}$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{internal}} + \Delta E_{\text{elec}} + \Delta E_{\text{vdW}} \tag{2}$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{pol}} + \Delta G_{\text{np}} \tag{3}$$

Where ΔE_{MM}, ΔG_{solv}, $T\Delta S$ are the changes in molecular mechanics energy, desolvation free energy, and conformational entropy, respectively. A further change in molecular mechanics energy is composed of ΔE_{MM} (bond, dihedral, and angle), ΔE_{elec} (electrostatic) and ΔE_{vdW} (van der Waals) and the change in desolvation free energy is composed of polar solvation (ΔG_{pol}) and non-polar solvation energy (ΔG_{np}). Estimation of configurational entropy was calculated using normal mode analysis (NMA) method [58], and only 50 snapshots are used in the calculation due to higher computational cost. Further, decomposition of binding free energies in amino acid levels was conducted by the Molecular Mechanics Generalized Born Surface Area (MM-GBSA) scheme. Also, the Computational Alanine Scanning (CAS) was performed for some essential residues. This method was used to determine the energetic contribution of the above mention residues and the energy difference between before and after the mutation using the following equation;

$$\Delta\Delta G_{bind} = \Delta G_{mutant} - \Delta G_{wild} \tag{4}$$

Results and discussions

In order to explore the mechanisms underlying the dimerization of nsp16 and nsp10 protein and preferential binding of the cofactor SAM to the nsp16/10 dimer in comparison to nsp16 monomer, a conformational free energy landscape (FEL) and binding free energy calculations were conducted using a combined MD/MM-PBSA scheme. All simulations were stable based on the total and potential energies of these systems **(data not shown).** The molecular snapshots obtained from the MD trajectories of complexes in explicit water were used further for the binding free energy analysis.

Overall dynamic structural features

To explore the dynamic stability of systems and to ensure the rationality of the sampling method, we monitored the structural and energetic properties during the entire 1 μs production simulation. The time evolution of the root-mean-square deviations (RMSDs) for the protein backbone atoms, which can reflect the stability of the system, were calculated during the production phase relative to the initial coordinates of MD simulations for all systems (see **Figure 2(A, B))**. The RMSD plots indicate that all the studied systems had reached equilibrium after ~100 ns. It is noteworthy that the RMSD value for $nsp10_{mono}$ is higher compared to the dimer, such as $nsp16_{dimer}$, $nsp10_{dimer,}$ as well as for the monomer, i.e., $nsp16_{mono}$. The average RMSD varies from 2.1 to 4.0 Å for dimer and monomers, respectively, from 100 to 1000 ns MD simulations **(see Table 1).**

Table 1: Average protein backbone RMSD value, the radius of gyration (RadGyr), and solvent-accessible surface area (SASA). Standard deviations are given in the parentheses. The statistical error was estimated on the basis of the deviation between block averages.

System	RMSD (Å)	RadGyr (Å)	SASA (Å^2)
dimer	2.8 (0.1)	22.5 (0.0)	18928.1 (103.3)
$nsp16_{dimer}$	2.1 (0.2)	18.8 (0.1)	13005.4 (92.5)
$nsp10_{dimer}$	3.1 (0.1)	14.0 (0.1)	5922.7 (31.8)
$nsp16_{mono}$	3.1 (0.3)	18.6 (0.0)	13467.3 (140.2)
$nsp10_{mono}$	4.0 (0.2)	13.9 (0.1)	6663.8 (145.7)

The RMSD values of $nsp10_{mono}$ and $nsp16_{mono}$ increased in the first 40 ns. The stability is maintained at an average fluctuation of 4.0 and 3.1 Å, respectively. However, the RMSD value of dimer continuously increased up to the first 30 ns and after that maintained at a fluctuation of about 2.8 Å until the end of the simulation length. However, the $nsp16_{dimer}$ was more stable compared to $nsp10_{dimer}$ and had lower variations. The change in the fluctuation pattern of $nsp10_{mono}$ and $nsp16_{mono}$ monomers was greater than that of their dimeric form. Overall, it suggests that the nsp16 with the SAM molecule gets more stabilized by the interaction with the nsp10. A change of ~ 1.0 Å in the structure of nsp16 was observed from the simulations of dimer and monomer systems. It must be noted that a stable RMSD does not always provide stable energy profiles; hence we have also verified the system's potential and total energy from the simulations (**data not shown**).
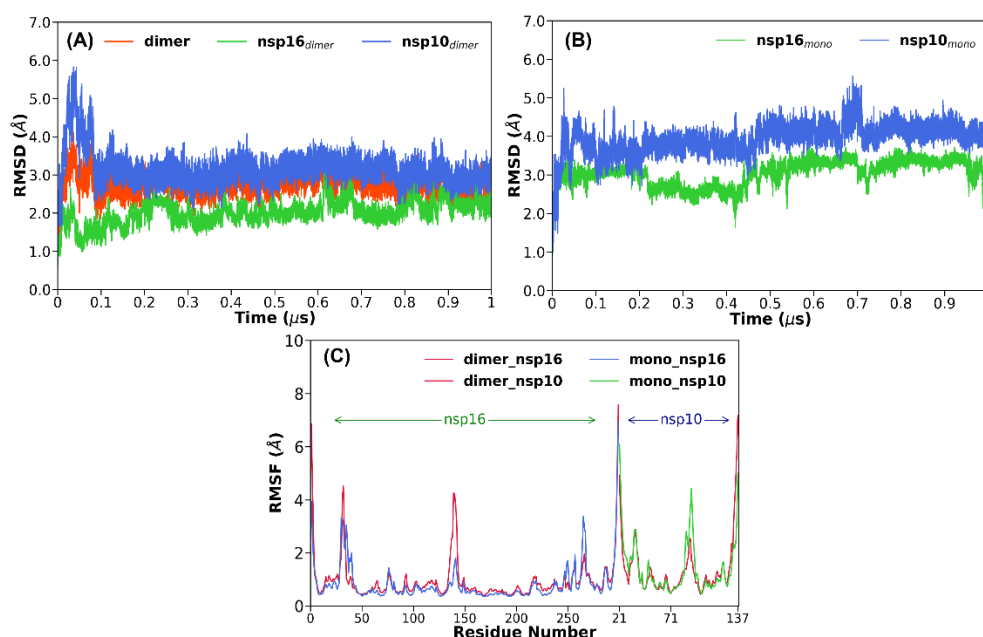
Figure 2: (**A, B**) The root-mean-square deviations (RMSDs) of the backbone atoms relative to their initial coordinates as a function of MD simulation time; (**C**) the root-mean-square fluctuations (RMSFs) of Cα atoms for each residue in dimer and monomers of nsp16/10 dimer.

Moreover, we also calculated the temporal RMSDs of heavy atoms of cofactor SAM (see **Supplementary Information Figure S1(A)**) and the backbone atoms of residues within 5 Å around cofactor SAM in the binding pocket (**see Supplementary Information Figure S1(B)**). The RMSD values of SAM and SAM binding pocket for the dimer were stable up to 850 ns the simulation and after that in the last 150 ns, fluctuations in the RMSD value was seen with an average value of 0.7 Å and 0.8 Å, respectively (**see Table S1, Supplementary Information**). It indicates that SAM binds strongly onto the cofactor binding cavity of the dimer nsp16/10. In the case of the monomer nsp16$_{mono}$ system, the RMSD values of SAM and SAM binding cavity are quite stable up to 100 ns, but after that, the RMSD value increases till 950 ns and maintains at a fluctuation of around 2.5 to 2.8 Å at the end of the simulation time. These high values of RMSDs suggested that SAM does not bind with the monomer form of nsp16$_{mono}$ in agreement with the experiment verified by Yu Chen and his coworkers for the SARS-CoV (2002) [27]. Additionally, three loops, namely loop 71-79, loop 100-108, and loop 130-148, which comprises the SAM binding pocket as shown in the crystal structure was also analyzed and found to be stable in both dimer and monomer (see **Supplementary Information S2 (A, B, C)**). However, the cap-binding groove of nsp16 in SARS-CoV as well as in SARS-CoV-2 is mainly built by two flexible loops, loop 26-38 and loop 130-148, while in case of flavivirus the NS5 MTase was quite stable with α-helices (A1, A2 and half of αD) along the

cap-binding groove.[59] Among the two loops involved in the cap-binding, only the loop 26-38 exhibited differences in the dimer and monomer of SARS-CoV-2 from our simulations. As shown in **Table S1**, the average RMSD of the monomer for this loop was higher, having a value of 3.1 Å compared to the dimer (1.4 Å). The time evolution of the RMSD of the loop 26-38 and its potential mean force (PMF) was given in **Supplementary Information Figure S3(A, and B).** From **Supplementary Information Figure S3(A)**, it was observed that RMSD of the loop 26-38 of monomer nsp16 exhibited high value but remain stable in the simulations. In contrast, the RMSD of the dimer nsp16/nsp10 showed low value but quite fluctuating in the simulations. From the PMF plot of the loop, as shown in **Supplementary Information Figure S3(B)** depicts that the primary low energy structure of $nsp16_{mono}$ was observed at ~3.2 Å due to the RMSD variation of loop 26-38. In contrast, for $nsp16_{dimer}$, the main energy minima observed at 1.4 Å. Besides, we detected two secondary minima (~1.9 Å, and ~2.8 Å), and the energy barriers for adjacent states were 1.1 and 1.2 kcal/mol, respectively. Overall, it suggests that the cap-binding groove in the dimer attains conformational changes to adopt the RNA compare to that when nsp16 is not bounded with the nsp10.

Furthermore, we have also studied the residual flexibility, root-mean-square fluctuations (RMSFs) of C_α atoms for both monomer and dimer (see **Figure 2 (B))** to understand the effect of dimerization on internal flexibility. Overall, the results suggested that dimer and monomer shared similar trends of RMSF patterns. However, we found some dynamic fluctuations located in different loop regions, including the N- and C-terminals. The large fluctuations loop regions are loop 26-38 (residues 26 to 38), loop 71-79 (residues 71 to 79), loop 100-108 (residues 100 to 108) and loop 130-148 (residues 130-148) of nsp16 monomer. The dimer nsp16 shows relatively higher flexibility than monomer. The substrate-binding site and nsp10 binding region in nsp16 showed a high degree of rigidity in the case of the dimer (residues 38 to 44, 78 to 87, 110 to 120 and around 250).

The structural compactness of each system was assessed by estimating the radius of gyration (RadGyr) from respective trajectories, and the average values of RadGyr for each system was evaluated and listed in **Table 1**. It is evident from **Table 1** that average RadGyr values for the dimer, $nsp16_{dimer}$, $nsp10_{dimer}$, $nsp16_{mono}$, and $nsp10_{mono}$ are 22.5, 18.8, 14.0, 18.6, and 13.9 Å, respectively. Herein, we also measured the solvent-accessible surface area (SASA) of all systems, and the average values were reported in **Table 1**. It can be seen from **Table 1** that the average SASA for the dimer ($nsp16_{dimer}$, $nsp10_{dimer}$) and monomer ($nsp16_{mono}$, and $nsp10_{mono}$) are 18928.1, 13005.4, 5922.7, 13467.3 and 6663.8 $Å^2$, respectively. We found that the

dimerization of nsp16 and nsp10 cover approximately 1203 Å^2 surface area in total, indicating a very stable interaction. Our simulations result also favor with experiment which showed the value of solvent exposed surface area for the heterodimer complex of nsp16/nsp10 as 19710 Å^2.[29]

Conformational free energy landscape of SAM binding

For exploring the conformational states and structural dynamics of dimer and monomer nsp16 with the cofactor SAM, 1 μs conventional MD (cMD) was performed for each case. As shown in **Supplementary Information Figure S1**, it is clear that the RMSD values of SAM are well converged after 200 ns and show noticeable two-state conformations. The dimer with SAM is well stable and shows no significant deviation from their initial state up to the end of the simulation. In the subsequent long-scale simulations, diverse conformations were sampled, thus providing scope for capturing the strong and weak bound state of SAM with nsp16. The two-dimensional free energy landscape (FEL) is performed to identify the relatively low energy states sampled in the MD simulation. Here, to explore the strong and weak bound state of the SAM molecule with nsp16 in both dimer and monomer complexes, the FEL was projected along with two collective variables (CV1 and CV2) for each system.
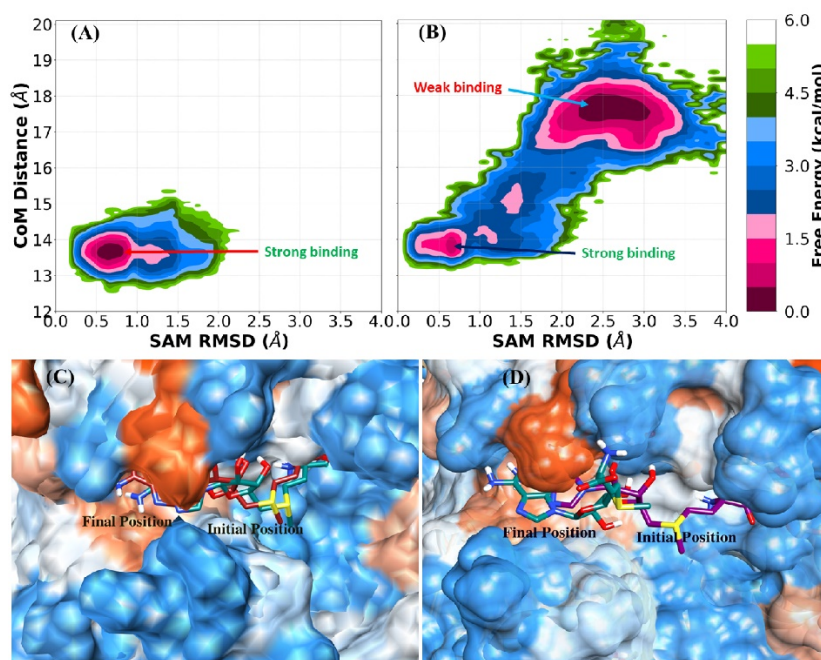

Figure 3: Free energy landscape and representative SAM position in dimer and monomer of nsp16. (**A**) FEL of SAM bound with dimer form of nsp16, (**B**) FEL of SAM bound with

monomer form of nsp16. Representative structures of (**C**) SAM with dimer nsp16 and (**D**) SAM with monomer nsp16.

The X-axis denotes as a CV1, the root-mean-square deviation (RMSD) of heavy atoms of SAM cofactor, which is able to characterize the positional deviations of heavy atoms from their respective initial position. The Y-axis denotes as a CV2 and defined as the center of mass (CoM) distance of the cofactor SAM and nsp16, which is able to characterize the displacement of the functionally vital counterpart of SAM and substrate-binding pocket of nsp16. The value of a CoM distance less than 14.5 Å denotes the strong binding of SAM toward nsp16.

The two-dimensional free energy landscapes (FEL) of dimer nsp16/nsp10/SAM and monomer nsp16/SAM are shown in **Figure 3(A), 3(B),** and the corresponding SAM position cavity is shown in **Figure 3(C), 3(D).** The FEL of dimer and monomer systems, suggesting that the conformational state of SAM in two systems prefer disparate configurations. Similar to RMSD calculations, one global minimum (strong SAM binding) was found for the dimer nsp16/nsp10/SAM system, and two local minima (strong and weak SAM binding) were discovered for the monomer nsp16/SAM complex, indicating that although the monomer nsp16/SAM has two favored conformations, in monomeric form weak binding state sampled more probable space compared to the other state. These results illustrated a stronger binding of the SAM with the dimer nsp16/nsp10 compared to the monomer nsp16, which agrees with experiments for SARS-CoV [27].

Principal component analysis of nsp16/10

The principal component analysis (PCA) was carried out for $nsp16_{dimer}$, $nsp10_{dimer}$, $nsp16_{mono}$, and $nsp10_{mono}$ systems. Each eigenvector and eigenvalue are plotted in the decreasing order, as shown in **Supplementary Information Figure S4**. For all cases, the first few eigenvalues describe the collective motion of local fluctuations. Comparing the four systems, it indicates that the first few PC's that described the properties of movements were not the same. The first two eigenvectors encapsulate for 42 %, 67 %, 46 %, and 70 % of overall movements in $nsp16_{dimer}$, $nsp10_{dimer}$, $nsp16_{mono}$, and $nsp10_{mono}$, respectively. Similarly, the first ten eigenvectors account for 81 %, 91 %, 78 %, and 90 % of the total motion in all four cases, respectively.

We also constructed free energy landscapes (FEL) using the first two principal components at 300 K for all four systems and shown in **Figure 4**. The sampling space of four systems was different, as given in **Figure 4**. The conformational sampling space of nsp16$_{dimer}$ is restricted as compared to the other three systems, which sampled wider conformational space. **Figure 4A**, we observed a global minimum of 62.7 % occupancy and a secondary minimum of 37.3 % occupancy, which suggested that the nsp16$_{dimer}$ system more stabilized in the presence of nsp10. The monomer form, nsp16$_{mono}$ sampled wider conformations, see **Figure 4C,** and the energy barriers of adjacent minimum structures are approximately 1.5 kcal/mol. On the other hand, in both monomeric and dimeric form, nsp10 protein sampled more space than nsp16, as shown in **Figure 4 B, D.** However, monomer nsp10$_{mono}$ has three distinct conformations with an energy barrier of 4.0 kcal/mol suggesting structural disparity in both nsp10$_{dimer}$ and nsp10$_{mono}$ as compared to nsp16.
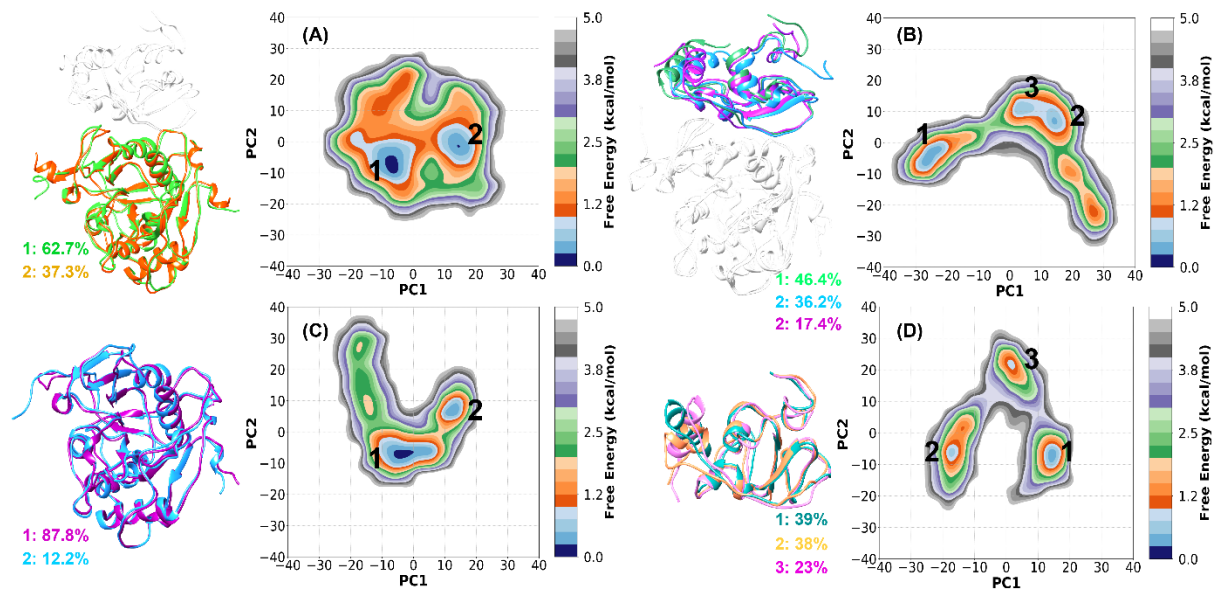
Figure 4. Two-dimensional free energy landscapes generated by projecting the first two principal components, PC1 and PC2 for (**A**) nsp16$_{dimer}$, (**B**) nsp10$_{dimer}$, (**C**) nsp16$_{monomer}$, and (**D**) nsp10$_{monomer}$. The representative structures are shown on the left panel.

Residual network analysis

The Network Analysis of Protein Structure (NAPS) server provided a visual examination of sub-network based on the physicochemical properties of the protein residues to find more

details about the interaction between two monomers after interacting with each other (see **Figure 5**). Herein, we explore the 3D interaction network of hydrophobic residues of nsp16 and nsp10 monomers. The results showed that the number of hydrophobic interaction networks between two monomers (nsp16, nsp10) was initially very high, and after 100 ns simulation, the number of networks reduced to ~2-3. The hydrophobic interaction networks such as (V104, **A71**) and (P80, **V42**) found as strong contacts and stable throughout the simulation time (see **Supplementary Information Table S2**). We also calculated the total number of interactions and noticed that the number of interaction networks is more or less conserved throughout the first 500 ns (see **Supplementary Information Figure S5**). In general, it can be concluded that the nsp16 and nsp10 interact with very high affinity.
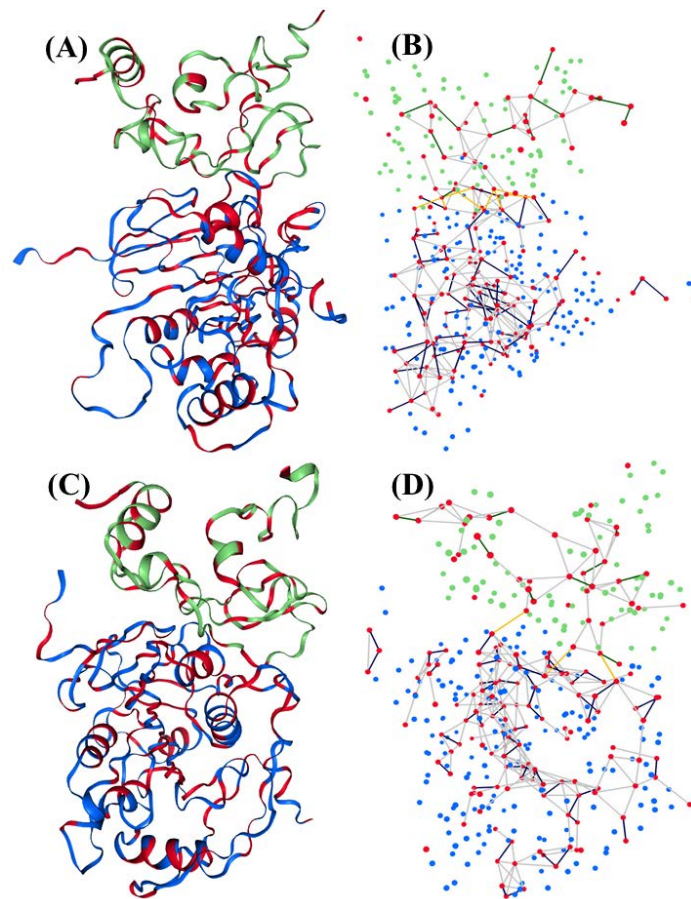

Figure 5: Sub-network representation of network 3D view showing hydrophobic residues (red) and edges (gold) between nsp16 and nsp10; (A, B) initial structure of dimer and (C, D) final structure.

Binding free energy of the SAM to nsp16 in monomer and dimer (nsp16/nsp10)

Binding free energy was calculated using the MM-PBSA method based on the equilibrated trajectories to reflect the affinity of the SAM molecule to nsp16 in its monomer form as well as when it is bound to nsp10. In coronavirus family, nsp10 acts as a stimulator in the binding of SAM molecule to nsp16 (2'-O-MTase) [23, 27]. To elucidate the role of nsp10 for the interactions of the SAM and nsp16 in COVID-19 was studied using the MD/MMPBSA approach. It provides various components contributing to the total binding energy (ΔG_{bind}) such as van der Waal interactions (ΔE_{vdW}), electrostatic interactions (ΔE_{ele}), polar solvation energy (ΔG_{pol}), non-polar solvation free energy (ΔG_{np}) and configurational entropy (TΔS). In total, 1000 frames were chosen from the stable regions for the calculation of binding free energy.

Table 2 Energetics components of binding free energy of the nsp16 protein with the SAM calculated using the MMPBSA for the systems: dimer (nsp16/nsp10) and monomer (nsp16) in kcal/mol. The standard error of the mean is provided in parentheses.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	ΔH	$-T\Delta S$	ΔG_{bind}
Dimer (nsp16/10)	-40.4 (0.1)	-181.1 (0.5)	179.0 (0.4)	-4.1 (0.0)	-46.6 (0.2)	39.8 (0.9)	-6.8 (0.9)
Monomer (nsp16)	-28.6 (0.1)	-76.8 (0.6)	84.7 (0.6)	-3.5 (0.0)	-24.2 (0.1)	23.6 (0.8)	-0.6 (0.8)

Table 2 shows the various energy components of the binding free energy of the SAM molecule to monomer (nsp16) and dimer (nsp16/nsp10). It is evident from **Table 2** that the binding affinity of the SAM molecule is higher for the dimer (ΔG_{bind} = -6.8 kcal/mol) than monomer (ΔG_{bind} = -0.6 kcal/mol). For both the cases, van der Waal interactions (ΔE_{vdW}), electrostatic interactions (ΔE_{ele}), and non-polar solvation free energy (ΔG_{np}) favors the binding of the SAM. In contrast, polar solvation energy (ΔG_{pol}) and entropy (TΔS) disfavors the binding. The electrostatic interaction energy is higher than the van der Waal interactions in both cases. In monomer (nsp16/SAM), although the ΔE_{ele} is higher (-76.8 kcal/mol) than ΔE_{vdW} (-28.6 kcal/mol). However, the disfavorable polar solvation energy (ΔG_{pol} = 84.7 kcal/mol) compensates for the electrostatic interaction energy, and hence the overall polar contribution ($\Delta E_{ele} + \Delta G_{pol}$) value is found to be 7.9 kcal/mol. In contrast, the overall non-polar contributions ($\Delta E_{vdW} + \Delta G_{np}$) value is found to be -32.1 kcal/mol. It implies that binding is driven by hydrophobic interactions. In the dimer, when nsp10 binds with nsp16, i.e. (nsp16/nsp10/SAM),

the affinity of the SAM molecule to nsp16 increases. From **Table 2**, the electrostatic interactions of the dimer (ΔE_{ele} = -181.1 kcal/mol) is higher than van der Waal interactions (ΔE_{vdW} = -40.4 kcal/mol), but the disfavouring polar contribution (ΔG_{pol}) is low (179.0 kcal/mol) than ΔE_{ele}. Hence, the total polar contributions ($\Delta E_{ele} + \Delta G_{pol}$) value is found to be negative (-2.1 kcal/mol), but it is lower than the total non-polar contribution ($\Delta E_{vdW} + \Delta G_{np}$), i.e., -44.5 kcal/mol. Hence, in dimer also the main force behind the binding of the SAM to nsp16 is hydrophobic interactions. Overall, the binding free energy analysis shows that the binding affinity of the SAM molecule to nsp16 increases with the presence of nsp10. Indicating nsp10 acts as a stimulator for binding of the SAM to nsp16 of COVID-19 in agreement with other coronavirus families such as SARS-CoV, MERS-CoV, etc. [23].

Table 3: Binding free energy decomposition of key residues calculated between the nsp16 protein and SAM for both dimer (nsp16/nsp10) and monomer (nsp16) in kcal/mol.

Residue	T_{vdW}	T_{ele}	T_{pol}	T_{np}	T_{back}	T_{side}	T_{total}
Dimer (nsp16/10)							
Asp99	0.71	-29.08	20.84	-0.14	-0.73	-6.94	-7.67
Asn43	-0.01	-9.90	5.89	-0.12	-0.30	-3.84	-4.14
Leu100	-2.42	-0.19	-0.10	-0.28	-0.87	-2.12	-2.99
Cys115	-0.74	-2.01	0.33	-0.05	-1.21	-1.26	-2.47
Met131	-2.72	0.56	-0.12	-0.13	-0.47	-1.94	-2.41
Asp114	-0.30	-8.98	7.41	-0.07	-0.77	-1.17	-1.94
Tyr47	0.07	-5.46	3.56	-0.04	-0.01	-1.86	-1.87
Gly71	-1.35	-3.27	2.84	-0.08	-1.36	-0.50	-1.86
Tyr132	-2.15	-2.23	2.86	-0.32	-0.93	-0.91	-1.84
Ser74	-0.95	-4.31	3.54	-0.08	-0.45	-1.35	-1.80
Phe149	-1.20	0.12	0.12	-0.13	-0.16	-0.93	-1.09
Monomer (nsp16)							
Leu100	-3.04	1.61	-1.17	-0.47	-0.18	-2.89	-3.07
Asp99	0.62	-19.83	16.98	-0.16	-0.15	-2.24	-2.39
Cys115	-0.70	-1.73	0.30	-0.06	-1.16	-1.03	-2.19
Met131	-1.72	-0.02	-0.05	-0.15	-0.55	-1.39	-1.94
Phe149	-1.86	-0.33	0.54	-0.18	-0.20	-1.63	-1.83
Asp114	-0.28	-7.47	6.32	-0.10	-0.73	-0.80	-1.53

Further, to explore the significant individual residues of nsp16 in the binding of SAM molecule and to evaluate the differences in the dimer (nsp16/nsp10) and monomer (nsp16), decomposition of the binding free energy of residues was performed using the MMGBSA approach and listed in **Table 3.** The residues having > 1.5 kcal/mol of energy were considered as important, and shown in **Figure 6.**
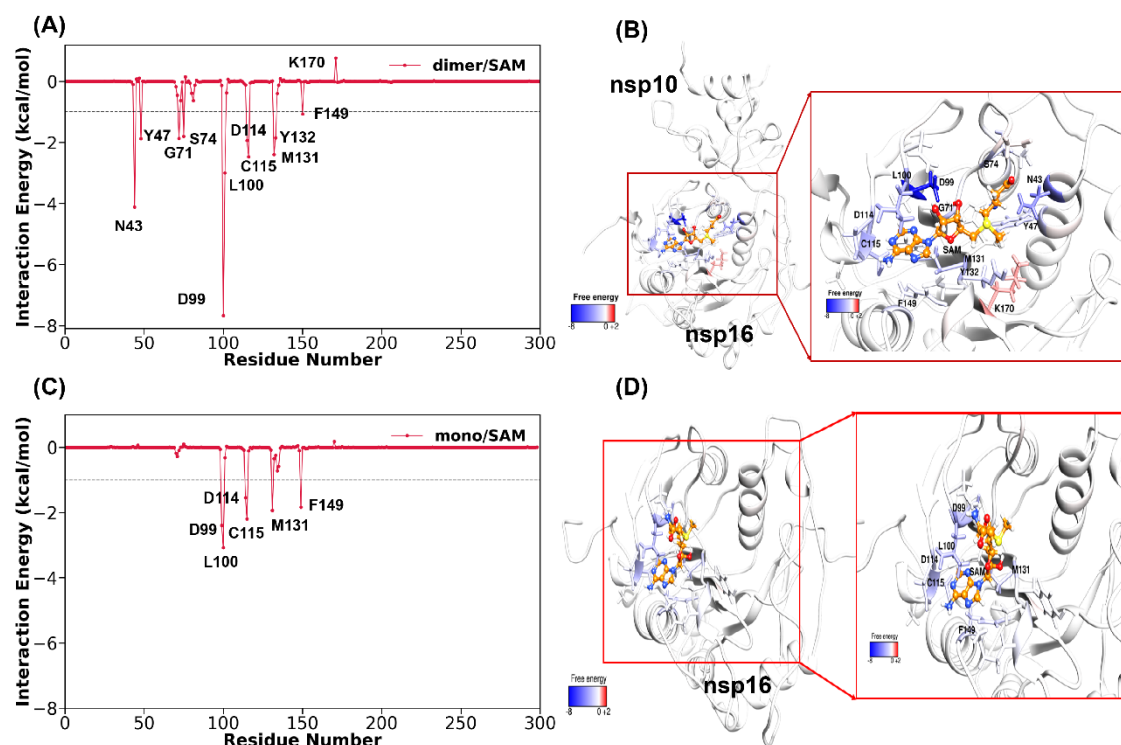

Figure 6: Per-residue decomposition free energy of nsp16 (dimer and monomer) for the binding of SAM and respective binding pocket drawn from MD snapshots. (**A, B**) dimer (nsp16/10) with SAM; (**C, D**) monomer nsp16 with SAM.

As seen in **Table 3**, the number of residues contributing to binding with SAM is high in dimer as compared to its monomer. Residues such as Leu100, Asp99, Cys115, Met131, Phe149, and Asp114 are common in both monomer and the dimer. Apart from these residues, in the dimer (nsp16/nsp10), additional residues such as Asn43, Tyr47, Gly71, Tyr132, and Ser74 also played a significant role in binding with the SAM. Hence, it suggests that binding of nsp10 to nsp16 influences the SAM binding pocket resulting in stabilization of the pocket by involving more residues in tighter binding with the SAM. These results are consistent with the higher binding energy in the dimer form.

Hydrogen bond interaction between nsp16 and SAM

The hydrogen bond (H-bond) mediated interaction plays an essential role between a protein and ligand. To investigate the interactions between the SAM and nsp16, that constituted the binding pocket, MD trajectories were used for the H-bond calculation. Furthermore, the impact of nsp10 binding to nsp16 and its influence on the H-bond pattern between SAM and nsp16 were calculated (see **Table 4**). The % occupancy reflected the stability of H-bond in the process of MD simulations. As seen in **Table 4**, the residues with > 50 % H-bond occupancy in the MD simulations are recognized for the dimer (nsp16/nsp10) compared to monomer nsp16. The residues, including Asp130, Tyr47, Gly71, and Asp99, showed more H-bond stability in the dimer form. Overall, a stable H-bond is formed between the nitrogen and oxygen atoms of the SAM with Asp130, Tyr47, Gly71, and Asp99 of nsp16, respectively. Therefore, the hydrogen bond interactions with residues Asp130, Tyr47, Gly71, and Asp99 should be promised in the design of novel SAM competitive inhibitors.

Table 4: The hydrogen bonds formed between nsp16 protein and the SAM in both dimer (nsp16/nsp10) and monomer (nsp16) form. The corresponding average distance and percent determined using the production trajectories in the MD simulations are also provided.

Binding couples		Molecular dynamics	
Acceptor	Donor…H	Distance (Å)	Occupancy[a] (%)
Dimer (nsp16/10)			
Asp130@OD2	SAM@N...HN1	2.84	71.69
SAM@N	Tyr47@OH...HH	2.79	64.95
Gly71@O	SAM@N...HN2	2.86	55.16
Asp99@OD2	SAM@O2'...HO2'	2.62	53.90
Asp99@OD1	SAM@O3'...HO3'	2.66	53.80
Asp99@OD2	SAM@O3'...HO3'	2.66	46.27
SAM@HN1	Tyr47@OH...HH	2.78	45.41
Asp99@OD1	SAM@O2'...HO2'	2.63	45.15
SAM@O	Asn43@ND2...HD22	2.84	42.20
SAM@OXT	Asn43@ND2...HD22	2.84	37.76
SAM@N1	Cys115@N...H	2.92	33.86
Asp114@OD2	SAM@N6...HN61	2.80	18.10
Asp114@OD1	SAM@N6...HN61	2.80	15.11

SAM@O2'	Asn101@ND2...HD22	2.89	12.21
Monomer (nsp16)			
Asp99@OD1	SAM@O3'...HO3'	2.64	48.80
Asp99@OD2	SAM@O3'...HO3'	2.63	45.78
SAM@N1	Cys115@N...H	2.92	28.66
Asp114@OD2	SAM@N6...HN61	2.80	26.44
Asp114@OD1	SAM@N6...HN61	2.80	24.80
Asp114@OD2	SAM@N6...HN62	2.80	19.03
Asp114@OD1	SAM@N6...HN62	2.80	17.95
Asp99@OD1	SAM@O2'...HO2'	2.65	16.42
Asp99@OD2	SAM@O2'...HO2'	2.65	14.70
SAM@N	Tyr47@OH...HH	2.92	11,31

[a]Only H-bonds with more than 10% occupancy are listed.

The changes in the distance of the atoms forming H-bond between SAM and nsp16 were also computed and shown in **Supplementary Information Figure S6.** In the dimer, the distance between the oxygen atom of Asp130, Tyr47, and Gly71 of nsp16 and nitrogen atom of SAM illustrated the average distance of 3.01, 2.89, and 3.01 Å, respectively, reflecting their stability of H-bond in the process of MD simulations. While, in the case of monomer, the average distances for these atoms were 12.48, 14.64, and 11.06 Å, respectively, suggesting no stable H-bond were formed in the monomer. However, for Asp99, the distance between its oxygen atom (OD1 and OD2) and the oxygen atom (O3) of SAM illustrated the average distance of 2.9 Å and 3.2 Å, respectively, for dimer and monomer. In contrast, the average distance between Asp99 (OD2) and SAM (O2) was higher for monomer (4.2 Å) than dimer (3.1Å). All these results were consistent with the occupancy analysis. Further, they emphasized that Asp130, Tyr47, and Gly71 of nsp16 were the key residues in the binding of SAM molecule and for the 2'-O-MTase activity of nsp16.

Binding free energy of nsp16-nsp10 complex and decomposition analysis

To evaluate further the binding free energy of the complex nsp16 and nsp10, ΔG_{bind} was obtained using the MD/MMPBSA approach. From **Table 5**, the binding free energy (ΔG_{bind}) between nsp16 and nsp10 was found to be -47.4 kcal/mol. The various energy components indicate that van der Waal interactions (ΔE_{vdW}), electrostatic interactions (ΔE_{ele}), and non-polar

solvation free energy (ΔG_{np}) favors the binding of nsp10 and nsp16. While polar solvation free energy (ΔG_{pol}) disfavor the binding. The electrostatic interactions (ΔE_{ele}) energy is higher (-429.4 kcal/mol) than the van der Waal interactions (ΔE_{vdW}) (-90.4 kcal/mol). (see **Supplementary Information Figure S7**) However, the disfavouring components of polar solvation energy (ΔG_{pol}) compensate for the ΔE_{ele} being a value of 481.7 kcal/mol. Hence, the total polar interactions ($\Delta E_{ele} + \Delta G_{pol}$) disfavor the binding between nsp10 and nsp16 with a value of 52.3 kcal/mol. In contrast, the total non-polar contributions ($\Delta E_{vdW} + \Delta G_{np}$) favor the binding between nsp16 and nsp10 (-100.0 kcal/mol). Therefore, the binding between nsp10 and nsp16 is mainly driven by hydrophobic interactions. Overall, the binding free energy analysis depicts, the binding affinity of the SAM to nsp16/nsp10 (ΔG_{bind} = -46.6 kcal/mol) is similar to that between nsp10 and nsp16 (ΔG_{bind} = -47.4 kcal/mol) of SARS-CoV-2. These observations were in agreement with its similar homology virus, SARS-CoV (2002). A similar binding affinity was seen between nsp10 and nsp16, as well as between SAM and the nsp10/nsp16 complex of SARS-CoV [27].

Table 5: Energetic components of the dimerization energy (ΔG_{bind}) between nsp16 and nsp10. Along with the Computational Alanine-Scanning (CAS) mutagenesis results of the dimer nsp16/nsp10. The binding energy is given in kcal/mol, and the standard error of the mean is provided in parentheses.

System	ΔE_{vdW}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	ΔG_{bind}	$\Delta\Delta G_{bind}^{a}$
WT	-90.4 (0.2)	-429.4 (3.0)	481.7 (3.2)	-9.6 (0.0)	-47.7 (0.4)	
I40A	-86.0 (0.2)	-429.2 (3.0)	480.5 (3.2)	-9.5 (0.0)	-44.2 (0.4)	3.5 (1.2)
V44A	-88.0 (0.2)	-429.2 (3.0)	481.4 (3.2)	-9.6 (0.0)	-45.5 (0.4)	2.1 (1.5)
V78A	-87.5 (0.2)	-429.5 (3.0)	481.0 (3.2)	-9.5 (0.0)	-45.5 (0.4)	2.2 (1.5)
R86A	-87.5 (0.2)	-522.6 (3.0)	574.5 (3.3)	-9.5 (0.0)	-45.1 (0.4)	2.6 (2.5)
Q87A	-86.8 (0.2)	-416.9 (3.0)	466.8 (3.2)	-9.4 (0.0)	-46.2 (0.4)	1.5 (1.6)
V104A	-86.9 (0.2)	-429.7 (3.0)	481.11 (3.2)	-9.4 (0.0)	-44.9 (0.4)	2.8 (1.5)

| M247A | -87.1 (0.2) | -428.5 (3.0) | 479.0 (3.2) | -9.6 (0.0) | 46.1 (0.4) | 1.5 (0.8) |

$$\Delta\Delta G_{bind}^{a} = \Delta G_{bind}^{mut} - \Delta G_{bind}^{WT}$$

To further investigate the key residues involved in the binding between nsp10 and nsp16, the binding free energy contributed were decomposed into the individual residues using the MMGBSA approach and shown in **Figure 7.**
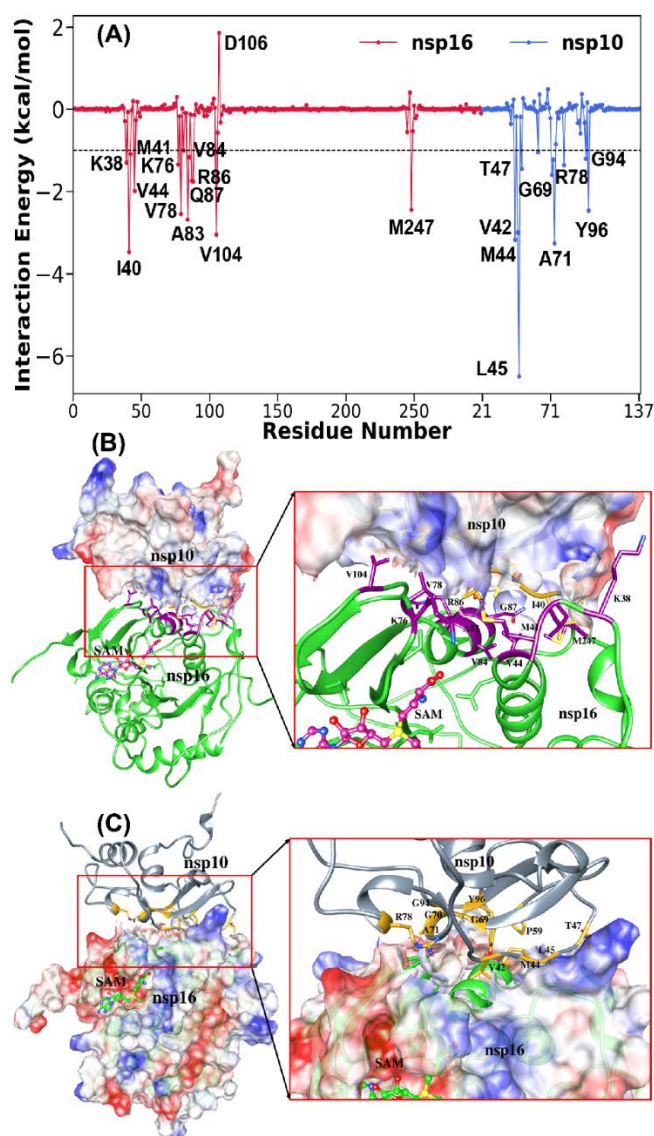

Figure 7: (A) Decomposition of ΔG on a per-residue basis for the nsp16/nsp10 dimer complex. (B) Corresponding residual position in nsp16 structure and (C) residual position in nsp10 structure.

The residues with the decomposed energy value higher than 1.0 kcal/mol were considered to be important and listed in **Table 6.** It is noted that in both systems, hydrophobic residues played a significant role. The key residues from nsp16 include Ile40, Val104, Ala83, Val78, Met247, Val44, Gln87, Arg86, Lys76, Lys38, Val84, and Met41. Similarly, the key residues from nsp10 include Leu45, Ala71, Val42, Met44, Tyr96, Gly69, Thr47, Arg78, Gly70, Gly94, and Pro59. These further reflected that hydrophobic interactions dominated the binding between nsp10 and nsp16. The greater contributing hydrophobic residues Ile40 and Leu45 from nsp16 and nsp10 respectively signify that the region occupied by these residues were important for the binding between nsp16 and nsp10.

Table 6: Binding free energy decomposition of key residues between nsp16 and nsp10 of the dimer in kcal/mol.

Residue	T_{vdW}	T_{ele}	T_{pol}	T_{np}	T_{back}	T_{side}	T_{total}
nsp16							
Ile40	-3.21	0.07	0.11	-0.44	-0.22	-3.25	-3.47
Val104	-3.40	3.00	-1.98	-0.67	0.29	-3.34	-3.05
Ala83	-1.47	-1.84	0.92	-0.29	-1.31	-1.37	-2.68
Val78	-2.66	1.61	-1.05	-0.44	0.09	-2.63	-2.54
Met247	-3.20	-1.12	2.33	-0.46	0.16	-2.61	-2.45
Val44	-1.87	-0.03	0.18	-0.27	-0.05	-1.94	-1.99
Gln87	-3.22	-4.41	6.54	-0.66	0.37	-2.12	-1.75
Arg86	-2.37	48.08	-47.12	-0.33	-0.60	-1.14	-1.74
Lys76	-1.58	27.69	-27.02	-0.43	-0.34	-1.00	-1.34
Lys38	-0.43	3.01	-3.60	-0.27	-0.93	-0.36	-1.29
Val84	-1.11	0.44	-0.46	-0.03	-0.44	-0.72	-1.16
Met41	-1.25	-0.13	0.50	-0.20	-0.02	-1.06	-1.08
nsp10							
Leu45	-5.75	-2.51	2.98	-1.22	-0.83	-5.67	-6.50
Ala71	-2.49	-5.31	4.97	-0.42	-1.21	-2.04	-3.25
Val42	-3.23	1.06	-0.38	-0.62	0.09	-3.26	-3.17
Met44	-3.86	-1.24	2.50	-0.39	0.16	-3.15	-2.99
Tyr96	-2.39	-3.63	4.00	-0.44	0.04	-2.50	-2.46
Gly69	-0.73	-3.60	2.81	-0.07	-1.36	-0.23	-1.59

Thr47	-1.87	0.06	0.77	-0.41	-0.29	-1.16	-1.45
Arg78	-2.38	-31.70	33.16	-0.44	0.14	-1.49	-1.36
Gly70	-0.64	-3.58	3.02	-0.02	-1.11	-0.11	-1.22
Gly94	-1.76	-3.85	4.75	-0.34	-0.59	-0.61	-1.20
Pro59	-1.28	-1.46	1.88	-0.17	-0.18	-0.85	-1.03

Interaction analysis between nsp16 and nsp10

In the above finding of binding free energy analysis, we have seen that nsp10 acts as a stimulator in the binding of the SAM to nsp16. Hence, to further explore the binding interactions between nsp16 and nsp10, H-bond and hydrophobicity analysis was calculated. As seen in **Table 7,** the H-bond occupancy reflects the stability of H-bond formation between protein-protein in the MD simulations. The strong H-bond was seen between the residues Ala83 (nsp16) and Tyr96 (nsp10) and between Gln87 (nsp16) and Leu45 (nsp10). Other residues include, Asp106 (nsp16) forms two H-bond with Ala71 and Gly94 of nsp10, Lys38 (nsp16) to Lys43 (nsp10), and Ser105 (nsp16) to Lys93 (nsp10). Hydrophobic interactions also played an important role in protein-protein or protein-ligand interactions [56, 60, 61]. Different H-bonds and hydrophobic interactions from the stable structure of the nsp16-nsp10 obtained by the MD simulations were plotted via Ligplot [62] and shown in **Figure 8.**

Table 7: The hydrogen bonds formed between nsp16 and nsp10 in the dimer and the corresponding average distance and percent determined using the production trajectories in the MD simulations.

Binding couples		Molecular dynamics	
Acceptor	Donor…H	Distance (Å)	Occupancy (%)
Ala83@O	**Tyr96@OH...HH**	2.73	90.80
Leu45@O	Gln87@NE2...HE21	2.85	67.83
Asp106@OD1	**Ala71@N...H**	2.87	29.23
Asp106@OD1	**Gly94@N...H**	2.84	24.23
Lys38@O	**Lys43@NZ...HZ2**	2.80	25.28
Lys38@O	**Lys43@NZ...HZ3**	2.80	25.01
Lys38@O	**Lys43@NZ...HZ1**	2.80	23.63

Asp106@OD2	**Gly94@N...H**	2.84	17.58
Ser105@O	**Lys93@NZ...HZ3**	2.79	12.65
Ser105@O	**Lys93@NZ...HZ2**	2.79	12.35
Ser105@O	**Lys93@NZ...HZ1**	2.79	12.25

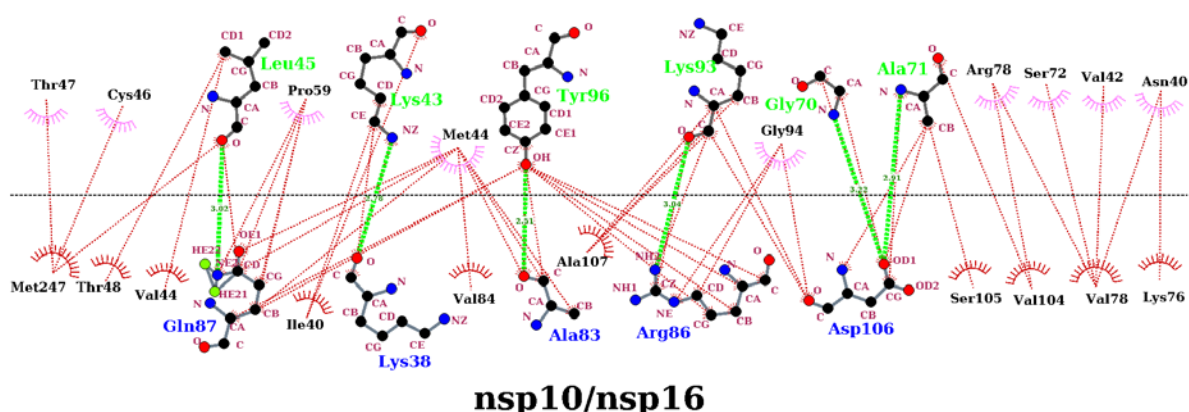**Bold** letters belong to the nsp10 structure. Only H-bonds with more than 10% occupancy are listed.

Figure 8: Protein-protein interaction diagrams for nsp16/nsp10 complex dimerization. The upper part is related to the nsp10 monomer, and the lower part is for the nsp16 monomer. The plot was generated by LigPlot+, and hydrogen bonds are shown as lime green dotted lines, and hydrophobic bonds are represented in red color.

The percentage occupancy of the residual contacts with the cut-off 3.9 Å for the protein-protein is also listed in **Supplementary Information Table S3.** Overall, these results highlighted the significant residues forming strong interactions between nsp16 and nsp10, which may help in the design of novel inhibitors blocking these protein-protein interactions against COVID-19 by inhibiting the 2'-O-MTase activity.

Computational alanine scanning mutagenesis analysis

The alanine scanning (AS) method is employed to investigate the role of a specific amino acid residue in protein-protein or protein-ligand interactions. The alanine scanning (AS) approach is used for the mutagenesis analysis. The basic principle involves in AS is the substitution of a residue with the alanine, which does not modify the main-chain conformation and hardly the side-chain beyond C_β of the complexes. In vitro, the experimental AS method has been proven

an advantageous mutagenesis method in finding the critical hot spot residues in protein interfaces. Computational alanine scanning (CAS) method has been reported as an excellent alternative approach for the in vitro experimental alanine scanning [63], [64]. Therefore, in this study, CAS was applied to investigate further the importance of specific residues except the alanine mentioned in the binding decomposition free energy of nsp16 of COVID-19.

The changes in the binding free energy were computed according to the **Equation (4)** (see method section) after replacing the residue of WT to alanine. The more the positive value of the $\Delta\Delta G_{bind}$, the heavier the effects of that single mutation in the protein-protein interactions. In our study, we have conducted the CAS for nsp16 by considering those residues with the decomposition free energy > 1.5 kcal/mol, as given in **Table 6**. The binding free energy components calculated from CAS mutagenesis for residues I40A, V104A, R86A, V78A, V44A, M247A, and Q87A of nsp16 are listed in **Table 5** and compared with the WT. As seen in **Table 6**, although the ΔG_{bind} of WT is higher than mutants, different energy components of mutants follow the same trend as of WT. The ΔE_{ele} is higher than ΔE_{vdW}, but disfavouring polar solvation energy ΔG_{pol} compensates ΔE_{ele}. As shown in **Figure 9(A),** for all the systems, the binding free energy is mainly coming from the hydrophobic interactions, where the total non-polar energy ($\Delta E_{vdW} + \Delta G_{np}$) is higher than the total polar energy ($\Delta E_{ele} + \Delta G_{pol}$).
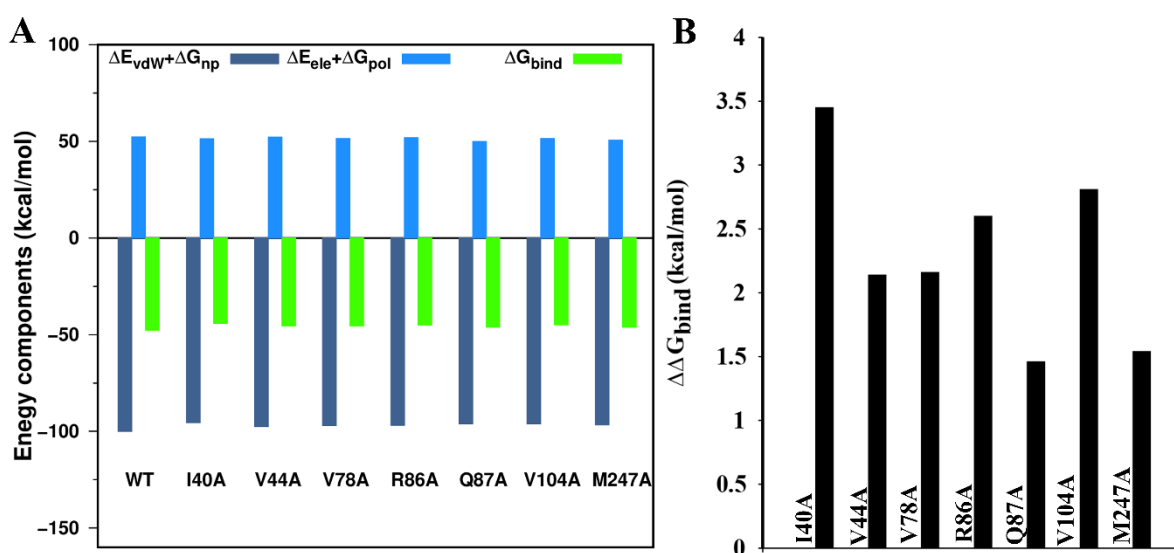

Figure 9: (**A**) Binding free energy components of the wild type and seven mutations (alanine scanning) systems. (**B**) Alanine scanning mutagenesis analysis of nsp10/16 complex.

Figure 9 (B) reflects further the significant residues in the protein interface calculated from CAS mutagenesis. It shows that $\Delta\Delta G_{bind}$ for mutants, I40A, V104A, and R86A are

comparatively high, suggesting that these residues showed significant changes in binding free energy after mutation. This observation was consistent with the above results of the decomposition of energy. CAS mutagenesis results depict that in addition to hydrophobic residues I40 and V104, hydrophilic residue R86 also played a vital role in the binding of nsp10-nsp16, whose mutations affect the binding of protein-protein.

Conclusions

The newly emerged coronavirus (SARS-CoV-2 or COVID-19) has imposed a global threat worldwide in concern with health as well as the economic status of many countries. The contagious nature and mortalities that arose by the COVID-19 globally demand an urgent need to control the pandemic situation. SARS-CoV-2 is being considered as the sister virus of SARS-CoV (2002) with ~ 80 % of sequence similarity. Previously explored information on targeting SARS-CoV can be used to study the different targets in controlling the present scenario of COVID-19. In eukaryotes and some viruses, the 5'end capping of the mRNA is an essential step for maintaining the stability of RNA, translation initiation, and for intracellular RNA transport. For targeting the 5' end capping mechanism of viral RNA such as SARS-CoV-2, one of its methyltransferase enzymes, namely 2'-O-MTase (nsp16), can be exploited in the development of antiviral drugs against COVID-19. Herein, we have employed the extensive MD simulations of 1μs along with the molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) method to study the binding mechanism of the nsp16 to nsp10 of SARS-CoV-2, and several other thermodynamic and dynamic features of protein binding. As shown previously for SARS-CoV (2002), nsp10 acts as a stimulator for the 2'-O-MTase activity of nsp16[23, 27]; we have also performed the simulations of monomer (nsp16/SAM) to reveal the role of nsp10 in SARS-CoV-2. Our study shows that the binding of nsp10 stabilizes the dimer nsp16/nsp10 structure. The FEL of monomer (nsp16) clearly exhibits two low energy state conformations, reflecting its transition from strong to weak binding of the SAM molecule with the nsp16. However, a single stable global minimum conformation was observed in the dimer form, reflecting its strong binding nature. These results further emphasized that nsp10 helps in the strong interaction between SAM and nsp16 to execute its 2'-O-MTase activity. The binding free energy between the SAM and nsp16 calculated from MM/PBSA revealed that the binding affinity is quite high for the dimer (ΔG_{bind} = -6.8 kcal/mol) than monomer (ΔG_{bind} = -0.63 kcal/mol). Overall, these results showed that nsp10 could stimulate the nsp16 to bind the methyl donor (SAM) and cap the RNA substrate for SARS-CoV-2 in agreement with other

coronaviruses such as SARS-CoV. Therefore, our simulations study predicts that SAM analog may work better in the heterodimeric form, which might help in the repurposing of drugs. Thus, apart from the active site of nsp16, the interface nsp16/nsp10 could also be considered as a potential target site in the design of antiviral drugs such as peptide inhibitors. The MD/MM-PBSA has been a widely used method for a better understanding of the mechanism between protein-protein or protein-ligand interactions. Our results suggested that hydrophobic interactions were predominant between nsp16 and nsp10 since most of the highly contributed residues were hydrophobic. It includes Ile40, Val104, Ala83, Val78, Met247, Val44, Gln87, Arg86, Lys76, Lys38, Val84, and Met41 from nsp16, and Leu45, Ala71, Val42, Met44, Tyr96, Gly69, Thr47, Arg78, Gly70, Gly94, and Pro59 from nsp10. Besides, the stable hydrogen bond between Ala83 (nsp16) and Tyr96 (nsp10), and between Gln87 (nsp16) and Leu45 (nsp10) were important in nsp16-nsp10 interface. Computational alanine scanning (CAS) study reveals that the residues I40A, V104A, R86A, V78A, V44A, M247A, and Q87A of nsp16 were considered as hot spot residues for the association of nsp16-nsp10. $\Delta\Delta G_{bind}$ for mutants, I40A, V104A, and R86A are comparatively high, suggesting that these residues showed significant changes in binding free energy after mutation. Hence, these residues can be utilized as efficient targets for developing potent inhibitors that suppress the association of nsp16-nsp10 and inhibiting the 2'-O-MTase activity of nsp16. Hence, our study provides a comprehensive understanding of the dynamic and thermodynamic process of binding of nsp16 and nsp10 that will contribute to the novel design of peptide inhibitors based on nsp16.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

1. I. I. Bogoch, A. Watts, A. Thomas-Bachli, C. Huber, M. U. Kraemer and K. Khan, *J. Travel Med.*, 2020, **27**, taaa011.
2. W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei and D. S. Hui, *New Engl. J. Med.*, 2020.
3. X. Lin, Z. Gong, Z. Xiao, J. Xiong, B. Fan and J. Liu, *Korean Journal of Radiology*, 2020, **21**, 365-368.
4. H. Lu, *Biosci. Trends*, 2020, **14**, 69-71.
5. T. Sheahan, A. Sims, S. Leist, A. Schäfer, J. Won, A. Brown, S. Montgomery, A. Hogg, D. Babusis and M. Clarke, 2020.
6. T. Pillaiyar, S. Meenakshisundaram and M. Manickam, *Drug Discov. Today*, 2020.
7. A. R. Fehr and S. Perlman, in *Coronaviruses*, Springer2015, pp. 1-23.
8. L. D. Eckerle, M. M. Becker, R. A. Halpin, K. Li, E. Venter, X. Lu, S. Scherbakova, R. L. Graham, R. S. Baric and T. B. Stockwell, *PLoS pathogens*, 2010, **6**.
9. P. C. Woo, Y. Huang, S. K. Lau and K.-Y. Yuen, *viruses*, 2010, **2**, 1804-1820.
10. A. M. King, E. Lefkowitz, M. J. Adams and E. B. Carstens, *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, Elsevier2011.
11. C. M. Coleman and M. B. Frieman, *Journal of virology*, 2014, **88**, 5209-5212.
12. Z.-R. Lun and L.-H. Qu, 2004.
13. S. R. Weiss and J. L. Leibowitz, in *Advances in virus research*, Elsevier2011, vol. 81, pp. 85-164.
14. Y. Arabi, A. Harthi, J. Hussein, A. Bouchama, S. Johani, A. Hajeer, B. Saeed, A. Wahbi, A. Saedy and T. AlDabbagh, *Infection*, 2015, **43**, 495-501.
15. C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu and X. Gu, *The Lancet*, 2020, **395**, 497-506.
16. S. R. Weiss and S. Navas-Martin, *Microbiol. Mol. Biol. Rev.*, 2005, **69**, 635-664.
17. Z.-Q. Zeng, D.-H. Chen, W.-P. Tan, S.-Y. Qiu, D. Xu, H.-X. Liang, M.-X. Chen, X. Li, Z.-S. Lin and W.-K. Liu, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2018, **37**, 363-369.
18. P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li and C.-L. Huang, *Nature*, 2020, **579**, 270-273.
19. C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li and X. Li, *Acta Pharmaceutica Sinica B*, 2020.
20. Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin and F. Cheng, *Cell discovery*, 2020, **6**, 1-18.
21. R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang and N. Zhu, *The Lancet*, 2020, **395**, 565-574.
22. Y. Furuichi and A. J. Shatkin, *Adv. Virus Res.*, 2000, **55**, 135-184.
23. Y. Wang, Y. Sun, A. Wu, S. Xu, R. Pan, C. Zeng, X. Jin, X. Ge, Z. Shi and T. Ahola, *J. Virol.*, 2015, **89**, 8416-8427.
24. Y. W. Chen, C.-P. B. Yiu and K.-Y. Wong, *F1000Research*, 2020, **9**.
25. B. H. Harcourt, D. Jukneliene, A. Kanjanahaluethai, J. Bechill, K. M. Severson, C. M. Smith, P. A. Rota and S. C. Baker, *Journal of virology*, 2004, **78**, 13600-13612.
26. H. Yang, W. Xie, X. Xue, K. Yang, J. Ma, W. Liang, Q. Zhao, Z. Zhou, D. Pei and J. Ziebuhr, *PLoS biology*, 2005, **3**.
27. Y. Chen, C. Su, M. Ke, X. Jin, L. Xu, Z. Zhang, A. Wu, Y. Sun, Z. Yang and P. Tien, *PLoS Pathog.*, 2011, **7**.
28. *1.80 Angstrom Resolution Crystal Structure of NSP16 - NSP10 Complex from SARS-CoV-2*, http://www.rcsb.org/structure/6W4H.

29.   M. Rosas-Lemus, G. Minasov, L. Shuvalova, N. L. Inniss, O. Kiryukhina, G. Wiersum, Y. Kim, R. Jedrzejczak, N. I. Maltseva and M. Endres, *bioRxiv*, 2020.

30.   M. H. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525-537.

31.   D. A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, I. T.E. Cheatham, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J.Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R.Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao and D. M. Y. a. P. A. Kollman, 2018, **University of California, San Francisco**.

32.   D. J. Price and C. L. Brooks III, *The Journal of chemical physics*, 2004, **121**, 10096-10103.

33.   J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696-3713.

34.   V. Kräutler, W. F. Van Gunsteren and P. H. Hünenberger, *J. Comput. Chem.*, 2001, **22**, 501-508.

35.   T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.

36.   H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola and J. Haak, *The Journal of chemical physics*, 1984, **81**, 3684-3690.

37.   D. R. Roe and T. E. Cheatham III, *J. Chem. Theory Comput.*, 2013, **9**, 3084-3095.

38.   T. Ichiye and M. Karplus, *Proteins: Struct. Funct. Bioinform.*, 1991, **11**, 205-217.

39.   S. Wold, K. Esbensen and P. Geladi, *Chemometrics Intellig. Lab. Syst.*, 1987, **2**, 37-52.

40.   A. Amadei, A. Linssen, B. De Groot, D. Van Aalten and H. Berendsen, *J. Biomol. Struct. Dyn.*, 1996, **13**, 615-625.

41.   N. A. Jonniya, M. F. Sk and P. Kar, *ACS Omega*, 2019, **4**, 17404-17416.

42.   M. F. Sk, R. Roy, N. A. Jonniya, S. Poddar and P. Kar, *J. Biomol. Struct. Dyn.*, 2020, https://doi.org/10.1080/07391102.07392020.01768149.

43.   B. Chakrabarty, V. Naganathan, K. Garg, Y. Agarwal and N. Parekh, *Nucleic Acids Res.*, 2019, **47**, W462-W470.

44.   P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan and W. Wang, *Acc. Chem. Res.*, 2000, **33**, 889-897.

45.   J. Wang, W. Wang, P. A. Kollman and D. A. Case, *Journal of molecular graphics and modelling*, 2006, **25**, 247-260.

46.   P. Kar, M. Seel, U. H. Hansmann and S. Höfinger, *Journal of Physical Chemistry B*, 2007, **111**, 8910-8918.

47.   P. Kar, Y. Wei, U. H. Hansmann and S. Höfinger, *J. Comput. Chem.*, 2007, **28**, 2538-2544.

48.   P. Kar, R. Lipowsky and V. Knecht, *Journal of Physical Chemistry B*, 2011, **115**, 7661-7669.

49.   P. Kar, R. Lipowsky and V. Knecht, *J. Phys. Chem. B.*, 2013, **117**, 5793-5805.

50.   T. Hou, J. Wang, Y. Li and W. Wang, *J. Chem. Inf. Model.*, 2011, **51**, 69-82.

51.   P. Kar and V. Knecht, *Journal of Physical Chemistry B*, 2012, **116**, 2605-2614.

52.   P. Kar and V. Knecht, *Journal of computer-aided molecular design*, 2012, **26**, 215-232.

53.   P. Kar and V. Knecht, *Journal of Physical Chemistry B*, 2012, **116**, 6137-6149.

54.   N. A. Jonniya, M. F. Sk and P. Kar, *ACS omega*, 2019.

55.   N. A. Jonniya and P. Kar, *J. Biomol. Struct. Dyn.*, 2020, **38**, 1306-1321.

56.   R. Roy, B. Ghosh and P. Kar, *ACS omega*, 2020, **5**, 3932-3942.

57. M. F. Sk, R. Roy and P. Kar, *Journal of Biomolecular Structure and Dynamics*, 2020, DOI: 10.1080/07391102.07392020.01724196.
58. B. Xu, H. Shen, X. Zhu and G. Li, *J. Comput. Chem.*, 2011, **32**, 3188-3193.
59. M. P. Egloff, D. Benarroch, B. Selisko, J. L. Romette and B. Canard, *EMBO J.*, 2002, **21**, 2757-2768.
60. N. A. Jonniya and P. Kar, *J. Biomol. Struct. Dyn.*, 2019, DOI: 10.1080/07391102.07392019.01602079.
61. M. F. Sk, R. Roy and P. Kar, *J. Biomol. Struct. Dyn.*, 2020, 1-16, DOI:10.1080/07391102.07392020.01724196.
62. A. C. Wallace, R. A. Laskowski and J. M. Thornton, *Protein engineering, design and selection*, 1995, **8**, 127-134.
63. I. S. Moreira, P. A. Fernandes and M. J. Ramos, *J. Comput. Chem.*, 2007, **28**, 644-654.
64. I. Massova and P. A. Kollman, *J. Am. Chem. Soc.*, 1999, **121**, 8133-8143.