Potential drugs targeting Nsp16 protein may corroborates a promising approach to combat SARS-CoV-2 virus

Uma Shankar[1±], Neha Jain[1±], Prativa Majee[1±], Subodh K. Mishra[2], Brijesh Rathi[3], Amit Kumar[1]*

[1]Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Simrol, Indore, 453552, India

[2]Department of Biological Sciences, RNA Institute, College of Arts and Sciences, University at Albany-SUNY, Albany, New York, 12222, USA

[3]Laboratory for Translational Chemistry and Drug Discovery, Hansraj College, University of Delhi, Delhi-110007 India

[±] These authors contributed equally

*Correspondence should be addressed to Amit Kumar. Email: amitk@iiti.ac.in, amitkdbs@gmail.com.

KEYWORDS: Nsp16, COVID-19, FDA approved, LOPAC, Drug designing

ABSTRACT

The recent ongoing pandemic caused by SARS-CoV-2 continues to impose devastating impacts and is accountable for the loss of more than 250,000 human lives within a short span of four months. This urges immediate therapeutic measures to control the impact of this disease. One of the most conserved and potentially druggable sites is the Nsp16 active site that performs the 2'-O-methyltransferase activity and puts a 5' cap on the viral RNA molecules. This allow them to mimic endogenous transcripts for the efficient translation of viral proteins and evasion of the

immune response. Three libraries of compounds (>5500) with chemical diversity were screened to identify hits against Nsp16 active site of SARS-CoV-2. From each library a top hit was identified, namely Velpatasvir (ΔG = -11.76 kcal/mol; LibDock score = 201.708) from the FDA compounds; JFD00244 (ΔG = -10.86 kcal/mol; LibDock score = 162.105) from the LOPAC library and compound **6** ((ΔG = -11.2 kcal/mol; LibDock score = 163.321) from the SAM based analog library. Interestingly, all three hits showed higher affinity than the positive controls. Velpatasvir is a known anti-viral drug used against Hepatitis C virus, and JFD00244 is a SIRT2 inhibitor. Molecular simulation studies showed all three molecules to have stable and energetically favourable interactions with the active site of Nsp16. In summary, this investigation identified three potential drug candidates that are predicted to be potent Nsp16 inhibitors and could be pursued further in cell-based studies.

INTRODUCTION

The SARS-CoV-2 causing the COVID-19 disease is the biggest outbreak of this decade, taking a heavy toll on the human population. More than one lakh deaths are reported worldwide, while the number of infected cases has reached over 2,800,000. Initially, reported from the city of Wuhan, China, this pandemic has hit almost all corners of the world, including 213 countries. The SARS-CoV-2 shares the common lineage with the other two epidemic causing viruses, Severe Acute Respiratory Syndrome-related coronavirus (SARS-CoV) and Middle East Respiratory Syndrome-related coronavirus (MERS-CoV), all belonging to the group of β-coronaviruses. SARS-CoV outbreak in 2003-2004 and MERS-CoV outbreak in 2012 have together caused more than 10,000 cases across the world till date, but surprisingly, the SARS-CoV-2 infection cases have extended to a million in a short span. Seeing the present condition, WHO has declared it a global pandemic situation, and the whole world urges for an immediate solution to combat this disease[1].

The SARS-CoV-2 viral genome consists of 29.9 kb single-stranded positive-sense RNA, which contains 15 ORFs region encoding several non-structural, structural, and accessory proteins. Other than the important structural proteins like Spike protein, Envelope protein, Membrane protein, and Nucleocapsid protein, the essential non-structural proteins required for the viral replication and proliferation are coded by the ORF1a and ORF1b[2]. ORF1a and ORF1b together make up almost two-third of the total genome at the 5'end, the first coding for Nsp1 to 11 and the later coding for Nsp12 to 16, respectively[3]. The non-structural proteins particularly embrace the virus specific enzymes like RNA dependent RNA polymerase and viral proteases, which assist in the viral replication and translation. But the most striking feature of the long RNA containing viruses is that they possess the 5'cap structure in their RNA like the eukaryotic RNAs. This helps the virus system to evade the host immune system by mimicking the host 5' cap structure of RNA as well as helps in effective viral protein production using the host machinery. This tactic of outwitting host defense mechanism is also practiced by other viruses belonging to families like Rhabdoviridae, Orthomyxoviridae, Alphaviridae, etc.[4] Conventionally, the eukaryotic RNA cap primarily consists of the N7-methylated guanine nucleotide attached to the first transcribed nucleotide (majorly consisting of an adenine) through a 5'-5' triphosphate bond. These steps are coordinated by RNA triphosphatase, guanylyltransferase, and N7 methyltransferase enzymes. The next step that succeeds this cap-0 structure (^{7Me}GpppA..) is the 2'-O-methylation by 2'-O-methyltransferase enzyme resulting in the cap-1 structure (^{7Me}GpppA$_{2'-O-Me}$...), which is required to prevent the decapping or degradation of the RNA[5]. A somewhat same strategy is utilized by the coronaviruses and the 2'-O-methylation process is executed exclusively by the Nsp16 protein, which is the AdoMet dependent methyltransferase enzyme in case of SARS-CoV-2 (Figure 1a). The S-adenosylmethionine (SAM) dependent 2'-O-

methyltransferase activity of Nsp16 was first reported in the feline coronavirus[6]. The addition of SAM recruits Nsp10 protein in this event, and this Nsp10-Nsp16 complex is quintessential for the viral RNA replication and transcription process[7-10]. The zinc finger protein Nsp10 acts as a key regulator in the viral replication as it modulates both the 3'-5' exoribonuclease and 2'-O-methyltransferase activities by interacting with Nsp14 and Nsp16 proteins, respectively. The Nsp10 bears an overlapping, interacting surface for both Nsp14 and Nsp16 enzymes for recruiting them into the transcription-replication complex[11]. The Nsp16 comprises the KDKE catalytic tetrad and is made up of central β-sheets with seven strands, encompassed by five α-helices, thereby lacking two of the seven α-helices from the conventional SAM-dependent 2'-O-methyltransferases making up the canonical fold. Nsp10 compensates this structural glitch thereby, stabilizing the Nsp16 protein and leads to the proper alignment of the first lysine (K46 catalytic residue) of the KDKE tetrad involved in the methylation reaction[12, 13]. Blocking the binding of SAM to Nsp16 may result in the formation of inactive 2'-O-methyltransferase enzyme in the absence of Nsp10 resulting in the failure of Cap1 structure formation which in turn leads to the inefficient viral translation and generation of host immune response against SARS-CoV-2 (Figure 1b). This makes SAM binding site of Nsp16 an attractive target to combat COVID-19 disease.

With no licensed drugs or vaccine against deadly COVID-19 disease, the situation gets aggravated with time and thereby, demands an urgent therapeutic solution. Drug respuposing approaches have suggested Chloroquine and Remdesivir effective against SARS-CoV-2[14-16]. Recently, SAM binding site of Nsp16 protein has emerged to be an interesting drug target site against this virus. Researchers have reported drugs showing significant inhibitory effect against the 2'-O-methyltransferase i.e., the Nsp16 in other viruses including SARS-CoV[17-19], but this

vital protein interface have been left untouched yet in case of SARS-CoV-2. The availability of Nsp10-Nsp16 complex crystal structure makes the screening of the drugs easier[20]. In this study, we utilized the Nsp10-Nsp16 complex as the potent druggable site for SAR-CoV-2 and screened three respective libraries against it. Virtual screening and molecular simulation analysis revealed Velpatasvir, an already known anti-viral drug against Hepatitis C virus, JFD00244, a SIRT2 inhibitor and compound **6**, a synthetic adenosyl analog from the FDA library, LOPAC library and SAM based analog library respectively as the best hits against the Nsp16 protein. This study thereby comes up with potent Nsp16 inhibitors for the treatment of SAR-CoV-2.

METHODS

1. **Sequence and structural alignment of Nsp16 for SARS-CoV-2, SARS-CoV, and MERS-CoV**

The Nsp16 nucleotide and protein sequences of SARS-CoV-2 (NC_045512), SARS-CoV (NC_004718) and MERS-CoV (NC_038294) were downloaded from the GenBank database available at the National Institute of Biotechnology Information (NCBI). The sequence alignment was performed using the MAFFT multiple alignment tool (https://mafft.cbrc.jp/alignment/server/) [21] and visualized using JalView.

For structural alignment, crystal structure files of the three viruses, SARS-CoV-2 (PDB: 6W4H), SARS-CoV (PDB: 3R24), and MERS-CoV (PDB: 5YN6) having S-adenosylmethionine (SAM) in the bound form were downloaded from the RCSB protein data bank. All three structures were superimposed to each other using Chimera tool, and motif analysis was performed.

2. **Construction of ligand libraries**

The ligands used in this study were compiled from the FDA approved drugs and the commercially available Library Of Pharmacologically Active Compounds (LOPAC). The list for 2400 FDA approved drugs was downloaded from the DrugBank in SDF format and 1280 bioactive small molecules from the LOPAC library were retrieved from Sigma (https://www.sigmaaldrich.com/life-science/cell-biology/bioactive-small_molecules/lopac1280_-navigator.html). The FDA approved drugs library included almost all the major classes of drugs targeting kinases, GPCRs, neurotransmission, gene-regulation, cancer, viruses, etc. Another library of 2000 molecules was constructed based on SAM by bioisosteric replacement and chain elongation methods in the SPARK (Cresset, UK) tool. Molecules in all three libraries were converted to 3D conformations and minimized to generate minimum energy conformers by applying the MMFF force field using the Discovery Studio 4.0[22].

3. Preparation of protein receptor

The available crystal structure of 1.8 Angstrom resolution of the Nsp10-Nsp16 complex of SARS-CoV-2 was fetched from the RCSB protein data bank (PDB: 6W4H)[20]. The attached ligands and ions were removed, missing loops were inserted and polar hydrogens were added using the Discovery Studio 4.0. Further, the protein refinement was done by using loop definition based on CHARMM force field at pH 7.4 parameter, keeping other parameters as default.

4. Virtual screening and molecular docking of the ligand libraries against the Nsp16 protein of SARS-CoV-2

Molecular docking of the three ligand libraries were performed against the SAM binding site of the Nsp16 viral protein that was executed on two platforms, namely, LibDock program available

in the Discovery Studio v4.0 (Accelrys, USA)[23] and AutoDock (http://autodock.scripps.edu/)[24]. The Discovery Studio docking module, LibDock consists of a high throughput docking algorithm that places the ligands into the active site of the protein based on the polarity of the interaction sites, while AutoDock performs the docking of the ligands inside a set of pre-calculated grids representing the target site of the protein.

LibDock performs flexible docking and uses polar and non-polar sites on the receptor termed hotspots. For LibDock docking, the residues of Nsp16 involved in the interaction with SAM were manually selected to define the active site sphere and docked using the FAST conformation method. Maximum 100 conformers of a ligand were allowed and filtered with a minimum LibDock score of 100. The resultant ligands were arranged in descending order of the LibDock score, and the top 20 hits were considered for further analysis.

For AutoDock analysis, initially, PDB file of the receptor protein was refined, Gasteir charges and polar hydrogen atoms were added, and non-polar hydrogens were merged using the AutoDockTools version (1.5.6) followed by converting the PDB files into PDBQT. The same tool was explored to prepare the ligand library into PDBQT format and Kollman charges were added to it. 3D grid box with the grid center size X=83.779, Y=15.102 and Z=26.856 and spacing of 0.425 encompassing the SAM binding sites of the protein and the essential residues situated in the binding pocket was constructed by the AutoGrid4. Docking was performed using the Lamarckian Genetic algorithm, and 10 conformers for each ligand were constructed using the AutoDock4. The resultant docked log file (.dlg) were analyzed, and conformations were ranked based on the binding energy (kcal/mol) of the protein-ligand complex.

5. **Analog generation**

For SAM based analogs generation, bioisosteric fragment replacement and ligand growing were performed using the SPARK tool. Various databases including ChEMBL_23 having very common, common, rare and very rare fragments observed in available approved drugs and ring systems from the VEHICLe databases, were considered for Bioisosteric replacement and chain growing of the ligand. The resultant 2000 conformers were analyzed for their binding affinity to Nsp16 using LibDock and AutoDock tools.

6. ADMET and drug-likeness prediction of the predicted analogs

ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties including the molecular weight, toxicity, total polar surface area, H-bond acceptors, H-bond donors, LD50, etc. were determined using the Discovery Studio v4.0. Lipinski rule of five was evaluated using DruLiTo tool. Aggregation probability was calculated using ChemAGG webserver (http://admet.scbdd.com/ChemAGG/index/).

7. Molecular dynamic (MD) simulation

The molecular dynamic simulation was performed to obtain the refined modeled structure of the ligand with the target to estimate the stability of interaction in the real environment. The simulation was carried on in a water sphere with the help of NAMD (Nanoscale Molecular Dynamics) standard molecular dynamics tool (https://www.ks.uiuc.edu/Research/namd/). The Visual Molecular Dynamics (VMD) tool v.1.9.3 was used to generate the required structure files (.psf) by exploiting the CHARMM force fields for proteins. The topology and parameter files for the ligands were generated using the CHARMM-GUI server. The protein-ligand system was first minimized for 10,000 steps by using the conjugate gradient algorithm and then subsequently heated from 0 K to 310 K. Later, 100 ns standard molecular dynamics was setup, and the

trajectory DCD file was generated that was further utilized to evaluate RMSD value of the docked complex.

RESULTS and DISCUSSION

Sequence and structural similarities revealed conserved active site of the Nsp16 (2'-O-methyltransferase) in SARS-CoV family

S-adenosylmethionine dependent 2′-O-methyltransferase enzyme of coronaviruses is involved in mRNA capping that facilitates the viral RNA to mimic the cellular RNA, thereby assisting the viruses to hide from the immune system cells and helps in the efficient viral protein translation. This enzyme is also involved in metabolism, chromatin remodeling, detoxification, and signal transduction[17], and thus crucial for the virus survival, representing an interesting target. To check the conservation at the sequence level, nucleotide and amino acid sequence of Nsp16 of SARS-CoV-2, SARS-CoV, and MERS-CoV were retrieved and aligned. Nucleotide analysis of SARS-CoV-2 revealed 89.45 % similarity with SARS-CoV and 67.24 % identity with MERS-CoV (Supplementary Information, Figure S1). While the protein alignment depicted 93.29 % and 66.11 % identity with SARS-CoV and MERS-CoV, respectively (Supplementary Information, Figure S2).

Nsp16 has a binding motif for S-adenosylmethionine and Nsp10 (Figure 2a-b). Coronavirus Nsp16 cannot function alone and requires the binding of Nsp10 to form a functional protein complex, and the recruitment of Nsp10 is facilitated by the binding of SAM to the Nsp16 protein[10]. This association differentiates coronavirus methyltransferase from other classical methyltransferase members of the viruses (Figure 2a-b)[25]. The structural analysis of SARS-CoV-2 revealed the presence of SAM binding site that comprises of ASN6841, TYR6845, LEU6855,

GLY6869, GLY6871, PRO6878, GLY6879, ASP6897, LEU6898, ASN6899, ASP6912, CYS6913, ASP6928, MET6929, and TYR6930 (Figure 2b). Structural alignment of SARS-CoV-2 Nsp16 with that of SARS-CoV shows high conservation with an RMSD of 0.201, while RMSD of 0.412 was observed for MERS-CoV depicting the closure identity between the SARS-CoV Nsp16 as compared to the MERS-CoV (Figure 2c-d). Structural alignment depicted the conserved SAM binding motif in all the three viruses depicting the critical role of these residues in the methyltransferase domain and other functional activities (Figure 2e).

Virtual screening of FDA approved and LOPAC libraries against SARS-CoV-2 Nsp16

Initially, FDA approved drugs library and LOPAC library were screened against Nsp16 using LibDock program in Discovery Studio v4.0, and the results were filtered based on the LibDock score, which depends upon the combined score of van der Waals forces, hydrogen bonds (H-bonds), and pi-pi interactions. Higher LibDock score indicates higher chances of receptor-ligand binding. SAM acts as an essential co-factor for Nsp16. If a drug can replace SAM from its binding site with higher affinity, the function of Nsp16 can be abolished. With this hypothesis, the SAM binding site was used as the ligand binding site in the LibDock program, and docking analysis was performed. Earlier, Sinefungin, a SAM analog, has shown an inhibitory effect on MERS-CoV methyltransferase [10]. As SAM binding sites of MERS-CoV and SARS-CoV are found to be conserved, SAM and Sinefungin both were taken as positive controls. Docking of SAM on Nsp16 by LibDock module showed the RMSD of 0.793 as compared to the co-crystalized ligand depicting the high efficiency of the LibDock (Supplementary Information, Figure S3). SAM provided a LibDock score of 136, while Sinefungin gave 146 (Supplementary Information, Figure S4 and Table S1). Therefore, to screen the FDA approved and LOPAC compounds libraries were having a higher affinity than that of SAM and Sinefungin, the docking

results were filtered on the basis of LibDock score > 146. In FDA approved drug library, a total of 156 drug molecules were received that showed the LibDock score >146 (Supplementary Information, Table S2). Similarly, 39 hits were observed in the LOPAC library that had a higher binding affinity to Nsp16 as compared to Sinefungin and SAM (Supplementary Information, Table S3). The top 20 hits of both FDA and LOPAC compound libraries were considered for further analysis (Supplementary Information, Figures 5a-b & 6a-b and Supplementary Information, Tables S4 & S5). The lead molecules of FDA approved drugs library consist of various anti-viral, anti-cancer, metabolic disorders, and neurotransmission drugs. The highest LibDock score of 233.11 was observed for Goserelin which is an anti-cancer drug used for the treatment of breast and prostate cancer while the second-ranked Tenapanor is used as therapeutic against irritable bowel syndrome. Anti-viral drugs included Velpatasvir, Ritonavir, Ledipasvir, Atazanavir, and Indinavir. Velpatasvir and Ledipadvir act as a substrate for NS5A and are used against chronic Hepatitis C Virus (HCV) infection. While Ritonavir, Atazanavir, and Indinavir work by inhibiting HIV protease and regulate the developmental cycle of HIV. The strong interaction of these anti-viral drugs, suggests the usage of Nsp16 as an alternate target for these anti-viral drugs. Another hit molecule, Cobicistat with a LibDock score of 203.37, is a known pharmaco-enhancer used in combination with Darunavir for the treatment of HIV. Flavin adenine dinucleotide, a supplement used in Vitamin B2 deficiency and peptide drugs, Glutathione disulfide, and Octreotide, were also among the lead molecules.

In the LOPAC compound library, the top rank with a LibDock score of 188.917 was observed for D-151, a di-adenosyl polyphosphate used for inhibiting ADP-induced platelet aggregations. Second-ranked Nelfinavir is known for its multi-target mechanism and reported to exhibit anti-cancer and anti-viral activity. It is also reported for reducing the inflammation by

blocking the MAPK signaling pathway in macrophages. In COVID-19, inflammation acts as one of the most critical factors in deciding the efficacy of the recovery in the patient, and thus Nelfinavir might act as an efficient therapeutic option by targeting Nsp16 and thereby reducing the inflammation. Other anti-cancer drugs in the top lead molecules included Fulvestrant, JFD00244, Lometrexol hydrate, and Dequalinium chloride. Antibiotic Paromomycin sulfate and anti-fungal agent Ketoconazole also showed an efficient binding with Nsp16.

Docking analysis revealed the interaction of top hits with the conserved residues of the active site of Nsp16

To analyze the interaction of top hits of FDA approved drugs library and LOPAC library and get a glimpse of the binding energy of the ligand-protein interaction, docking analysis using AutoDock program was performed and 20 different conformers based on the orientation of the ligand in the active site of Nsp16 were generated for each ligand. The interaction between ligand and Nsp16 was a result of H-bonds, van der Waals interaction, and pi-pi interaction that directly influence the binding energy of the complex.

At first, the grid was prepared by selecting the co-crystalized ligand, and then the docking analysis was performed for SAM and Sinefungin. Upon analysis, SAM interacted with various active site residues with the binding energy of -7.85 kcal/mol, while Sinefungin interaction resulted in the binding energy of -7.92 kcal/mol (Supplementary Information, Figure S4 and Supplementary Information, Table S1). The AutoDock result of the top 20 hits of FDA and LOPAC libraries were filtered by keeping the binding energy threshold of -7.92 kcal/mol. The molecules giving the higher binding energy than the threshold were considered as better ligands for Nsp16 in comparison to SAM and Sinefungin.

Upon analyzing the docked log files of the complexes, out of 20 FDA approved drugs and LOPAC library, 09 and 11 compounds respectively showed higher affinity towards Nsp16 as compared to the controls (Figure 3 and 4).

Velpatasvir (DB11613) from the FDA approved drugs scored the best binding energy of -11.76 kcal/mol (Figure 5a-d) while JFD00244 with a binding energy of -10.86 kcal/mol topped the LOPAC library (Figure 6a-d). Velpatasvir, in the combination with Sofosbuvir inhibits NS5A and NS5B proteins that are involved in the Hepatitis C virus replication, assembly, and host immune response modulation[26]. The close similarity of the replication mechanism in the Hepatitis C virus and SARS-CoV-2 makes it an interesting therapeutic option for drug development. Recently, it was reported to interact with 3C like protease (3CLpro) of COVID-19 [27]. The higher binding energy and LibDock score with Nsp16 protein of SARS-CoV-2 support a new protein target of Velpatasvir in the virus, and if supported by in vitro mode, it can be a potential drug by binding to multiple protein targets in the same pathogen which can tackle the multi-drug resistance in the virus. Cobicistat depicted the second lowest binding energy of -11.59 kcal/mol making it a promising drug candidate for Nsp16. Ritonavir, Ledipasvir, Indinavir, and Flavin adenine dinucleotide also depicted to have binding energy to be higher than the threshold (Figure 3 and Supplementary Information, Table S4).

In the LOPAC compound library, JFD00244 ranked first with the lowest binding energy of -10.86 kcal/mol (Figure 6a-d). JFD00244 is a sirtuin 2 (SIRT2) inhibitor and is reported to have an anti-cancer activity [28]. S-(p-Azidophenacyl) glutathione stood second, followed by Nelfinavir mesylate hydrate with a binding energy of -10.22 kcal/mol and -9.35 kcal/mol respectively. Lometrexol hydrate, Paromomycin sulfate, Astaxanthin, Fulvestrant, Ketoconazole, Calmidazolium chloride, Suramin, and Dequalinium chloride from the LOPAC library also

showed higher binding affinity than SAM and Sinefungin and had higher binding energies than the two controls (Figure 4 and Supplementary Information, Table S5).

Simulation analysis revealed a stable and energetically favored interaction of Velpatasvir and JFD00244 with Nsp16.

To analyze the stability and further refinement of the Nsp16 and lead molecule complexes, Velpatasvir from the FDA approved drugs library with an AutoDock binding energy of -11.76 kcal/mol and LibDock score 201.708 and JFD00244 from the LOPAC library with binding energy -10.86 kcal/mol and LibDock score of 162.105 were selected for molecular dynamics simulation analysis. 10,000 steps of conjugate gradient energy minimization of Velpatasvir resulted into the change in potential energy of the docked complex from -61.23 kcal/mol to – 101.775 kcal/mol (Supplementary Information, Figure S7). Thereafter the potential energy, kinetic energy and total energy of the system remained constant throughout the simulation process at constant temperature and pressure (Supplementary Information, Figure S7). Root mean square deviation (RMSD) analysis of alpha carbon (Cα) of the backbone of Nsp16 revealed an initial fluctuation from 0 Å to 1.18 Å w.r.t. the reference initial state of the complex in the first 50 ps. Thereafter, it remained constant till 100 ns throughout the simulation process depicting the stability of the docked complex (Figure 5e).

Energy minimization of JFD00244 showed the decrease in the system energy from -159.3 kcal/mol to -186.83 kcal/mol leading to a generation of a more stabled docked complex (Supplementary Information, Figure S8). Trajectory analysis of the Nsp16-JFD00244 revealed the fluctuations in the initial 17 ns, and thereafter the system was equilibrated till the end (Figure 6e).

Analog generation resulted into more efficient Nsp16 binders

Majority of top FDA approved drugs and LOPAC library hits shared a complex 9-12 membered ring structures like Goserelin, Tenapanor, Leuprolide, Nelfinavir mesylate, JFD00211, Velpatasvir, Indinavir, Lometrexol hydrate, Dequalinium chloride hydrate and Methotrexate etc. (Supplementary Information, Figure S9). Also, many top hits, for example, Flavin adenine dinucleotide, alpha beta-Methyleneadenosine, D-151, Thio-NADP sodium, NADPH tetrasodium, shared adenosyl moiety similar to that of SAM and Sinefungin. On analyzing the docked complexes of all these ligands, it was observed that the heterocyclic rings of these ligands fit into the groove similar to that of the adenosyl group of SAM and Sinefungin making the site an attractive option to build the analog library by fixing the heterocyclic ring at one hand (Figures 3 & 4, Supplementary Information, Figure S9). The heterocyclic ring structure of both the top leads, Velpatasvir and JFD00244 also interacted with the heterocyclic ring pocket similar to that of SAM while the side chains interacted with the rest of the cavity (Figures 5a and 6a).In all the ligands, the change in binding affinity was majorly based on the side chain attached. Also, the binding of SAM to the Nsp16 leads to the recruitment of Nsp10, an essential protein required for the proper functioning of 2'-O-methyltransferase making it a potential site for lead drug development and optimization. Therefore, the adenosyl (Adenine + ribose sugar) group was used as a starting point to generate 2000 number of analogs first by selecting the Sulphur atom and then by selecting the terminal carbon by fragment replacement and chain extension algorithms. This SAM based analog library was then screened using the LibDock tool and finally by AutoDock tool. The top 20 molecules with the LibDock score greater than that of Sinefungin were analyzed by AutoDock tool. 14 compounds out of 20 showed higher binding energy as compared to SAM (Figure 7 and Supplementary Information, Table S6). The highest ranked

analog, Compound 6 formed four hydrogen bonds in the SAM binding pocket of Nsp16 with a binding affinity of -11.2 kcal/mol, which was significantly higher than that of SAM (-7.85 kcal/mol) depicting ~1.5 folds higher binding affinity of the analog with the Nsp16. The second (Compound 15) and third (Compound 2) ranked molecules had a binding energy of -10.33 kcal/mol and -10.29 kcal/mol with the Nsp16, respectively (Supplementary Information, Table S6). Interestingly, on analyzing the binding poses of the analogs on Nsp16, it was observed that the adenosyl group of almost all the analogs was restricted to the same binding cavity as observed in SAM and Sinefungin complex depicting the importance of the pocket (Figure 7b). The pocket may provide the stability to the protein-ligand complex by providing stacking stabilization through pi-pi interactions. Furthermore, the total binding energy of the analogs was depended on the hydrogen and van der Waals interaction made by the varying side chains of each molecule. In our analysis, it was observed that replacement of the linear chain at the termini results into a better binding pose, eventually leading to a lower binding energy. SAM and Sinefungin differ from each other just by a Sulphur atom resulting into a better binding affinity of the later with Nsp16 protein. Similarly, we also observed that replacing the Sulphur (S) with either Nitrogen (N) or Carbon (C) results into a better ligand binding than SAM (Figure 7).

To analyze the stability of the predicted analog and Nsp16, the Nsp16-lead analog complex with the lowest binding energy, i.e., Compound 6 was analyzed in a water sphere having 0.15 M KCl ions using molecular dynamic simulation. The RMSD of the complex as predicted by analyzing the trajectory analysis of the system for 100 ns depicted the initial fluctuation from 0 to 1.5 Å till 700 ps, which thereafter remained constant throughout the simulation. The RMSD analysis revealed the formation of a stable Nsp16-Compound 6 complex. (Figure 8e and Supplementary Information, Figure S10).

Physiochemical properties analysis of the analogs affirms their drug-likeness

The generation of a safe and effective drug, depends upon various parameters, which mainly can be summarized into its ADMET (Absorption, Distribution, metabolism, Excretion, and Toxicity) analysis. Also, there are various parameter ranges like molecular weight, number of hydrogen bond acceptors and donors, lipophilicity etc. that can help in a better generation of a successful drug. Lipinski rule of five is one of the classical filters that can be applied for filtering out the best lead candidates from synthetic constructs (Supplementary Information, Table S7). As per the rule, a molecule with a molecular weight < 500 Dalton, maximum 5 hydrogen bond donors, and 10 hydrogen bond acceptors with logP value < 5 have a better drug-likeness in comparison to others that fail these parameters. The top 20 lead analogs were analyzed by using Discovery Studio v4.0 to check whether they follow the Lipinski rule of five. Out of 20 analogs, all the analogs have the molecular weight and LogP value within the threshold range, but only 12 molecules have less than 10 hydrogen bond donor and/or 5 hydrogen bond acceptors making them more likeable to act as lead molecules. Compound 6has a molecular weight of 417.2 Dalton, LogP -1.724, and 9 hydrogen bond acceptors and 4 hydrogen bond donors, making it a preferable lead molecule against Nsp16. Mutagenic analysis and Ames test analysis of the analogs revealed the non-mutagenic nature of all the analogs except compound 5. Log S that depicts the solubility of the drugs was in the range of -0.7 – 3.0 supporting the solubility of the analogs except Compound 1 that was moderately soluble with the value of -4.02. Drug aggregation inside the cells/blood stream is one of the major setbacks of high throughput screened drugs; therefore we analyzed the aggregation probability of the analogs that depicted the non-aggregating nature of these analogs supporting their usage as lead molecules against Nsp16 (Supplementary Information, Tables S7 & S8).

CONCLUSION

The uncontrolled outbreak of SARS-CoV-2 in more than 200 countries worldwide have put a major challenge to the healthcare sector for the identification and generation of novel drug molecules that can be used against the viral infection. Nsp16 of SARS-CoV-2 plays a critically essential role in viral replication and host immune response invasion by capping the viral genomic RNA similar to that of the host, thereby making it a novel and a potential target to be explored against the virus. Extensive virtual screening and docking analysis identified an anti-viral drug, Velpatasvir and an anti-cancer SIRT2 inhibitor agent, JFD00244 as lead molecules in the FDA approved drug library and LOPAC library, respectively. Molecular dynamics simulation analysis revealed the stability of the Nsp16-ligand complex. Furthermore, a synthetic library was constructed based on S-adenosylmethionine by fixing the adenosyl group and replacing its aliphatic side chain. The analog generation resulted into various lead molecules that showed a higher binding affinity towards Nsp16 than SAM and already known inhibitor, Sinefungin. The lead analog, **6** showed the highest binding affinity with the lowest binding energy. ADMET and physiochemical property analysis supported drug-likeness, non-mutagenic, non-toxic and non-aggregating nature of the lead analog, 6. In summary, the work leads to the conclusion that Velpatasvir, an already known anti-viral drug, and SIRT2 inhibitor JFD00244, could be repurposed for SARS-CoV-2 virus targeting Nsp16. Hence, Velpatasvir could be tested for in-vitro and in-vivo inhibitory effects on SARS-CoV-2 virus proliferation. The conserved nature of the binding pocket residues of Nsp16 in coronavirus family may lead to the development of effective therapeutics against the previously evolved strains like MERS-CoV and SARS-CoV and as well as against the newly emerging endemic causing viruses of the same family in the future.

ASSOCIATED CONTENT

Supporting Information. Figures S1-S10 and Tables S1-S8

AUTHOR INFORMATION

Corresponding Author

*Amit Kumar, Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore, Simrol, Indore, 453552, India; Orcid ID: 0000-0002-5913-4308; Email – amitk@iiti.ac.in, amitkbds@gmail.com

Author Contributions

AK conceived the idea and methodologies. US, NJ, PM and SKM run the computations and analyzed data with AK. BR performed the critical review. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no competing financial interest.

ABBREVIATIONS

SARS-CoV - Severe Acute Respiratory Syndrome-related coronavirus; FDA - Food & Drug

Administration; LOPAC - Library of Pharmacologically Active Compounds; SAM - S-

adenosylmethionine

REFERENCES.

1.   Sohrabi, C.; Alsafi, Z.; O'Neill, N.; Khan, M.; Kerwan, A.; Al-Jabir, A.; Iosifidis, C.; Agha, R., World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International journal of surgery (London, England)* **2020**, 76, 71-76.
2.   Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z., A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, 579, 265-269.
3.   Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* **2020**, 395, 565-574.
4.   Decroly, E.; Ferron, F.; Lescar, J.; Canard, B. J. N. R. M., Conventional and unconventional mechanisms for capping viral mRNA. **2012**, 10, 51-65.
5.   Picard-Jean, F.; Brand, C.; Tremblay-Létourneau, M.; Allaire, A.; Beaudoin, M. C.; Boudreault, S.; Duval, C.; Rainville-Sirois, J.; Robert, F.; Pelletier, J.; Geiss, B. J.; Bisaillon, M., 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. *PloS one* **2018**, 13, e0193804-e0193804.
6.   Decroly, E.; Imbert, I.; Coutard, B.; Bouvet, M.; Selisko, B.; Alvarez, K.; Gorbalenya, A. E.; Snijder, E. J.; Canard, B. J. J. o. v., Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2′ O)-methyltransferase activity. **2008**, 82, 8071-8084.
7.   Decroly, E.; Debarnot, C.; Ferron, F.; Bouvet, M.; Coutard, B.; Imbert, I.; Gluais, L.; Papageorgiou, N.; Sharff, A.; Bricogne, G. J. P. p., Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2′-O-methyltransferase nsp10/nsp16 complex. **2011**, 7.
8.   Chen, Y.; Su, C.; Ke, M.; Jin, X.; Xu, L.; Zhang, Z.; Wu, A.; Sun, Y.; Yang, Z.; Tien, P.; Ahola, T.; Liang, Y.; Liu, X.; Guo, D., Biochemical and Structural Insights into the Mechanisms of SARS Coronavirus RNA Ribose 2′-O-Methylation by nsp16/nsp10 Protein Complex. *PLoS pathogens* **2011**, 7, e1002294.
9.   Debarnot, C.; Imbert, I.; Ferron, F.; Gluais, L.; Varlet, I.; Papageorgiou, N.; Bouvet, M.; Lescar, J.; Decroly, E.; Canard, B., Crystallization and diffraction analysis of the SARS coronavirus nsp10-nsp16 complex. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2011**, 67, 404-408.
10.   Aouadi, W.; Blanjoie, A.; Vasseur, J. J.; Debart, F.; Canard, B.; Decroly, E., Binding of the Methyl Donor S-Adenosyl-l-Methionine to Middle East Respiratory Syndrome Coronavirus 2'-O-Methyltransferase nsp16 Promotes Recruitment of the Allosteric Activator nsp10. *Journal of virology* **2017**, 91.

11.   Bouvet, M.; Lugari, A.; Posthuma, C. C.; Zevenhoven, J. C.; Bernard, S.; Betzi, S.; Imbert, I.; Canard, B.; Guillemot, J.-C.; Lécine, P.; Pfefferle, S.; Drosten, C.; Snijder, E. J.; Decroly, E.; Morelli, X., Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. *J Biol Chem* **2014**, 289, 25783-25796.

12.   Lugari, A.; Betzi, S.; Decroly, E.; Bonnaud, E.; Hermant, A.; Guillemot, J.-C.; Debarnot, C.; Borg, J.-P.; Bouvet, M.; Canard, B. J. J. o. B. C., Molecular mapping of the RNA Cap 2′-O-methyltransferase activation interface between severe acute respiratory syndrome coronavirus nsp10 and nsp16. **2010**, 285, 33230-33241.

13.   Subissi, L.; Imbert, I.; Ferron, F.; Collet, A.; Coutard, B.; Decroly, E.; Canard, B. J. A. r., SARS-CoV ORF1b-encoded nonstructural proteins 12–16: replicative enzymes as antiviral targets. **2014**, 101, 122-130.

14.   Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res* **2020**, 30, 269-271.

15.   Madrid, P. B.; Chopra, S.; Manger, I. D.; Gilfillan, L.; Keepers, T. R.; Shurtleff, A. C.; Green, C. E.; Iyer, L. V.; Dilks, H. H.; Davey, R. A.; Kolokoltsov, A. A.; Carrion, R., Jr.; Patterson, J. L.; Bavari, S.; Panchal, R. G.; Warren, T. K.; Wells, J. B.; Moos, W. H.; Burke, R. L.; Tanga, M. J., A Systematic Screen of FDA-Approved Drugs for Inhibitors of Biological Threat Agents. *PLOS ONE* **2013**, 8, e60579.

16.   Gao, J.; Tian, Z.; Yang, X., Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Bioscience trends* **2020**, 14, 72-73.

17.   Decroly, E.; Debarnot, C.; Ferron, F.; Bouvet, M.; Coutard, B.; Imbert, I.; Gluais, L.; Papageorgiou, N.; Sharff, A.; Bricogne, G.; Ortiz-Lombardia, M.; Lescar, J.; Canard, B., Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog* **2011**, 7, e1002059-e1002059.

18.   Pugh, C. S.; Borchardt, R. T.; Stone, H. O. J. B., Inhibition of Newcastle disease virion messenger RNA (guanine-7-)-methyltransferase by analogs of S-adenosylhomocysteine. **1977**, 16, 3928-3932.

19.   Dong, H.; Zhang, B.; Shi, P.-Y. J. A. r., Flavivirus methyltransferase: a novel antiviral target. **2008**, 80, 1-10.

20.   Monsen, R. C.; Trent, J. O., G-quadruplex virtual drug screening: A review. *Biochimie* **2018**, 152, 134-148.

21.   Katoh, K.; Rozewicki, J.; Yamada, K. D., MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* **2017**, 20, 1160-1166.

22.   Systemes, D. J. S. D., CA: Dassault Systemes, BIOVIA Discovery Studio Visualizer 2016. **2016**.

23.   Systèmes, D. J. D. S. B. S. D., CA, USA, BIOVIA, Discovery Studio Modeling Environment. **2016**.

24.   Huey, R.; Morris, G. M. J. T. S. R. I., USA, Using AutoDock 4 with AutoDocktools: a tutorial. **2008**, 54-56.

25.   Menachery, V. D.; Debbink, K.; Baric, R. S., Coronavirus non-structural protein 16: evasion, attenuation, and possible treatments. *Virus research* **2014**, 194, 191-199.

26.   Zignego, A. L.; Monti, M.; Gragnani, L., Sofosbuvir/Velpatasvir for the treatment of Hepatitis C Virus infection. *Acta Biomed* **2018**, 89, 321-331.

27.   Chen, Y. W.; Yiu, C.-P. B.; Wong, K.-Y., Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CL (pro)) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. *F1000Res* **2020**, 9, 129-129.

28.   Tervo, A. J.; Kyrylenko, S.; Niskanen, P.; Salminen, A.; Leppänen, J.; Nyrönen, T. H.; Järvinen, T.; Poso, A., An In Silico Approach to Discovering Novel Inhibitors of Human Sirtuin Type 2. *Journal of Medicinal Chemistry* **2004**, 47, 6292-6298.

Figure Legends.

Figure 1. Schematic representation of the function and hypothesis of the work. a. Normal function of Nsp16 in complex with SAM and Nsp10 that leads to efficient viral translation and helps the virus to hide from the host immune system. **b.** Proposed hypothesis of targeting the SAM binding site leading to failure of attachment of SAM and recruitment of Nsp10 leading to the inactivation of Nsp16 thereby leading to the inefficient viral translation and generation of host immune response.

Figure 2. Nsp16 domains and conservation analysis. a. Surface representation of SARS-CoV-2 Nsp16 (green) in complex with Nsp10 (cyan) along with SAM (yellow). **b.** SAM binding site of SARS-CoV-2 Nsp16 along with the ligand-interacting residues **c & d**. Structural alignment of Nsp16 of SARS-CoV-2 (red), SARS-CoV (green), and MERS-CoV (blue) in cartoon and ribbon form, respectively. **e.** Aligned SAM along with the conserved residues of the Nsp16 binding site in the SARS-CoV-2 (red), SARS-CoV (green), and MERS-CoV (blue) depicting the conserved binding site in the coronavirus family.

Figure 3. Nsp16-FDA approved lead molecule complex structures. Receptor Nsp16 (wheatish grey) in complex with the lead FDA approved molecules (blue, stick representation) that showed higher affinity in comparison to SAM and Sinefungin in both LibDock and AutoDock analysis.

Figure 4. Nsp16-LOPAC library lead molecule complex structures. Receptor Nsp16 (wheatish grey) in complex with the lead LOPAC library compounds that showed higher affinity in comparison to SAM and Sinefungin in both LibDock and AutoDock analysis.

Figure 5. Nsp16-Velpatasvir interaction. a. Docking site of Velpatasvir (red) on the Nsp16 (wheatish grey) protein. **b.** Binding pose of Velpatasvir on the SAM binding site of Nsp16. **c& d.** Interaction of Velpatasvir with the Nsp16 residues in ribbon and graphical form. e. 100 ns RMSD graph obtained by the trajectory analysis of Nsp16–Velpatasvir.

Figure 6. Nsp16–JFD00244 interaction. a. Docking site of JFD00244 (green) on the Nsp16 (wheatish grey) protein. **b.** Binding pose of JFD00244 on the SAM binding site of Nsp16. **c & d.** Interaction of JFD00244 with the Nsp16 residues in ribbon and graphical form. **e.** 100 ns RMSD graph obtained by the trajectory analysis of Nsp16–JFD00244.

Figure 7. SAM based analog library generation and their interaction analysis with Nsp16. **a.** Structure of SAM depicting the sites (*i* and *ii*) where replacement and/or chain extension, were

performed resulting in the generation of 2000 ligands. **b.** Top 14 lead analogs with their LibDock score and AutoDock binding energy (in kcal/mol).

Figure 8. Nsp16–Compound 6 interaction. a. 2D structure of Compound 6 analog**. b.** Binding cavity of Compound 6 on Nsp16. **c & d.** Interaction of Compound 6 with the Nsp16 residues in ribbon and graphical form **e.** 100 ns RMSD graph obtained by the trajectory analysis of Nsp16–Compound 6.

Figure 1.

Figure 2.

Velpatasvir
LibDock Score – 201.708
Binding Energy – -11.76

Cobicistat
LibDock Score – 203.375
Binding Energy – -11.59

Abarelix
LibDock Score – 213.948
Binding Energy – -10.48

Terlipressin
LibDock Score – 189.871
Binding Energy – -10.3

Ritonavir
LibDock Score – 192.147
Binding Energy – -10.86

Ledipasvir
LibDock Score – 188.506
Binding Energy – -9.46

Flavin adenine dinucleotide
LibDock Score – 190.744
Binding Energy – -9.05

Indinavir
LibDock Score – 182.363
Binding Energy – -8.71

Octreotide
LibDock Score – 191.251
Binding Energy – -8.2

Figure 3.

JFD00244
LibDock Score – 162.105
Binding Energy – -10.86

S-(p-Azidophenacyl)glutathione
LibDock Score – 163.059
Binding Energy – -10.22

Nelfinavir
LibDock Score – 180.409
Binding Energy – -9.35

Lometrexol
LibDock Score – 157.561
Binding Energy – -9.34

Paromomycin sulfate
LibDock Score – 169.767
Binding Energy – -8.86

Astaxanthine
LibDock Score – 164.58
Binding Energy – -8.63

Fulvestrant
LibDock Score – 162.787
Binding Energy – -8.35

Ketoconazole
LibDock Score – 153.794
Binding Energy – -8.22

Calmidazolium chloride
LibDock Score – 177.89
Binding Energy – -8.16

Suramin
LibDock Score – 177.89
Binding Energy – -8.16

Dequalinium chloride
LibDock Score – 158.01
Binding Energy – -7.2

Figure 4.

26

Figure 5.

Figure 6.

a

b

Compound 1
LibDock Score - 164.493
Binding Energy - -8.27

Compound 2
LibDock Score - 164.291
Binding Energy - -10.29

Compound 3
LibDock Score - 159.475
Binding Energy - -8.43

Compound 4
LibDock Score - 157.091
Binding Energy - -8.11

Compound 5
LibDock Score - 167.836
Binding Energy - -8.85

Compound 6
LibDock Score - 163.321
Binding Energy - -11.2

Compound 7
LibDock Score - 157.091
Binding Energy - -7.99

Compound 8
LibDock Score - 165.566
Binding Energy - -7.95

Compound 9
LibDock Score - 165.257
Binding Energy - -9.41

Compound 11
LibDock Score - 171.007
Binding Energy - -8.54

Compound 14
LibDock Score - 165.314
Binding Energy - -8.33

Compound 15
LibDock Score - 163.974
Binding Energy - -10.33

Compound 20
LibDock Score - 166.324
Binding Energy - -9.54

Figure 7.

Figure 8.

Active 2'-O-MTase

Inactive 2'-O-MTase

No capping

Velpatasvir

JFD00244

Decrease Virus Efficiency and prevent immunoevasion