

## Identification of Potent Drugs and Antiviral Agents for the Treatment of the SARS-CoV-2 Infection

N.R. Jena\*

Discipline of Natural Sciences, Indian Institute of Information Technology, Design, and Manufacturing

Dumna Airport Road, Jabalpur-482005, India

---

\*Corresponding author's email address: [nrjena@iiitdmj.ac.in](mailto:nrjena@iiitdmj.ac.in)

### Abstract

The recent outbreak of the SARS-CoV-2 infection has affected the lives and economy of more than 200 countries. The unavailability of vaccines and the virus-specific drugs has created an opportunity to repurpose existing drugs to examine their efficacy in controlling the virus activities. Here, the inhibition of the RdRp viral protein responsible for the replication of the virus in host cells is examined by evaluating the binding patterns of various approved and investigational antiviral, antibacterial, and anti-cancer drugs by using the molecular docking approach. Some of these drugs have been proposed to be effective against the SARS-Cov-2 virus infection. In an attempt to discover new antiviral agents, artificially expanded genetic alphabets (AEGIS) such as dP, dZ, dJ, dV, dX, dK, dB, dS, dP4, and dZ5 and different sequences of these nucleotides were also docked into the active site of the RdRp protein. It is found that among the various approved and investigational drugs, the Clevudine, N4-hydroxycytidine (EIDD-1931), 2'-C-methylcytidine, EIDD-2801, Uprifosbuvir, Balapiravir, Acalabrutinib, BMS-986094, Remdesivir (full drug), GS-6620, and Ceforanide would act as potent inhibitors of the RdRp protein. Similarly, among the AEGIS nucleotides, dB, dJ, dP, dK, dS, dV, and dZ are found to inhibit the RdRp protein efficiently. It is further found that sequences containing up to three artificial nucleotides can also inhibit the RdRp viral protein. As these unnatural nucleotides are foreign molecules for the cells, their replication in host cells may not occur naturally, and hence these may act as potent drugs for the treatment of the SARS-CoV-2 infection. However, *in vivo* evaluation of these drugs and

artificial nucleotides are required against SARS-CoV-2 before prescribing them to the infected patients.

## **1. Introduction**

The recent outbreak of the Coronavirus infection originated from the Wuhan City of China in December 2019 (Wuhan virus-2019 or COVID-19) [1,2] has led to severe damage to human lives and the economy of more than 200 countries worldwide. The unavailability of vaccines and potent drugs specific to COVID-19 has created a challenge and opportunity to identify or discover potent antiviral agents for the treatment of this pandemic. Although continuous attempts are going on to design vaccines [3,4] and antiviral agents [5-12], not a single vaccine has cleared human trials successfully to date. Hence, it is imperative to repurpose available drugs for the treatment of COVID-19.

It is found that COVID-19 is a positive-strand RNA virus, and the virus protein is structurally similar to the severe acute respiratory syndrome coronavirus (SARS-CoV-1) [13], which originated from China in 2002 and several other beta coronaviruses, such as the bat and pangolin coronaviruses. Due to its structural similarity with the SARS-CoV-1, it is also popularly known as SARS-CoV-2. The SARS-CoV-2 protein consists of a multi-protein cascade comprising of several structural and non-structural proteins [14]. Among these proteins, the RNA-dependent RNA polymerase (RdRp) non-structural protein (NS) is mainly responsible for the replication of the virus inside host cells [15]. It is recently proposed that the RdRp requires other non-structural proteins such as NS7 and NS8 to process replication of the incoming RNA strand [16]. Hence, to block the replication of the SARS-CoV-2, it is desirable to block the active site of the RdRp NS protein by potent nucleotide inhibitors.

Recent cryo-EM studies have resolved the structure of the RdRp NS protein in complex with NS7 and NS8 [16,17]. These studies have revealed that it contains several vital domains, such as an N-terminal beta-hairpin (residues 31-50), nidovirus RdRp-associated nucleotidyltransferase domain (NiRAN, residues 117-250), interface domain (residues 251-365), and RdRp domain (residues 367-920) [16,17]. The active site is located in the RdRp domain formed by the conserved motifs ranging from A to G (residues 397-815). Among these motifs, motif A contains residues 611-626, and motif C includes residues 753-767, including the catalytically important residues Ser759, Asp760, and Asp 762, which play essential roles in the synthesis of RNA. The metal ions found in most of the viral RdRps are coordinated by Asp760 and Asp762. Motif F (residues 397-581) and G (residues 621-679) interact with the template strand RNA and direct it into the active site [16,17]. The positively charged hydrophilic residues, such as Lys545, Arg553, and Arg555 of the motif F, helps to stabilize the +1 base of the incoming strand, thereby putting the RNA strand in the correct position for catalysis [16]. The residues involved in the RNA binding and catalysis are highly conserved and hence crucial for RdRp activities. Therefore, to inhibit replication of the viral RdRp NS protein, antiviral agents should disrupt the binding of the RNA template with the above residues.

Recently, the structure of the RdRp, complexed with NS7, NS8, a 50 base template-primer RNA, and Remdesivir monophosphate (RMP), was resolved by the cryo-EM study [16]. The RMP is an antiviral prodrug, which is believed to inhibit the RdRp activity by non-obligate RNA chain termination. It was found that the RMP interacts with the RdRp protein by interacting with the side chains of Lys545 and Arg555 [16]. The phosphate group of the RMP was found to interact with the  $Mg^{+2}$ . Other than RMP, Favipiravir, Ribavirin, Galidesivir, EIDD-2801, EIDD-1931 (N4-hydroxycytidine), etc have also been proposed to inhibit RdRp

activity by cell-based assays [18,19]. As these are nucleoside- or nucleotide-based drugs, the mechanism of inhibition of RdRp activity is presumed to be the same as that of RMP. However, the binding modes of these antiviral drugs are not yet known. Further, the binding mode of the full drug of Remdesivir with the modified backbone as deposited in the DRUGBANK database [20] is not known.

To understand the detailed binding modes of various nucleoside and nucleotide-based drugs and to identify the more potent drugs based on the strong interactions with the active site of the RdRp, binding of 28 approved and ten investigational drugs were studied herein by using molecular docking approach. The binding modes of several artificially expanded genetic information systems (AEGIS) [21-29] with RdRp were also examined to unravel their possible roles in inhibiting RdRp activity. These unnatural nucleotides were proposed to make the artificial gene to create artificial life.

## **2. Computational Methodology**

### **2.1 System Preparation**

Two Cryo-EM structures of the RdRp protein complexed with NS7 and NS8 are available in the protein data bank [15,16]. In one of the structures, the RdRp-NS7-NS8 complex was not bound to metal ions (pdb ID 6M71) [17]. In contrast, in the other one, the RdRp-NS7-NS8 complex was bound to three  $Mg^{+2}$  ions (coordinated by Asp760, Asp761, Asp618, and pyrophosphate), a double-stranded RNA (14-base template strand and 11-base primer strand), and Remdesivir in its monophosphate form (RMP) (pdb id 7BV2) [16]. However, in the latter structure, Arg555 was protruding away from the active site, unlike in the former structure (Fig. 1a). The superposition of RdRp without metal ions onto the RdRp with metal ions by considering the  $C_{\alpha}$  atoms produced an RMSD of 0.583 Å (Fig. 1a). As the latter

structure contains metal ions, which are conserved in almost all of the RdRp viral proteins, this structure was considered for docking after removing coordinates of NS7, NS8, pyrophosphate, and RMP. Before docking, hydrogen atoms were added to the RdRp protein of SARS-CoV-2 by UCSF Chimera to maintain the neutral pH.

The three-dimensional structures of various nucleobase, nucleoside, and nucleotide drugs were extracted from the DRUGBANK database (by downloading the SDF files) [20]. Wherever three-dimensional structures were unavailable, two-dimensional structures were obtained from the DRUGBANK database. Subsequently, three-dimensional structures of these drugs were generated by using the GaussView 5.0 program [30] as per the standard geometrical parameters. Subsequently, these molecules were minimized by using the UCSF Chimera. The coordinates of artificial nucleotides, such as P and Z, were extracted from the protein databank (pdb id 4XNO) [31]. Subsequently, hydrogen atoms were added to these nucleotides by using the GaussView program. These nucleotides were modified to create other nucleotides such as P4, Z5, J, V, B, S, X, and K. Subsequently, these nucleotides were energy-minimized by using the Gaussian 09 program [32]. Consequently, the minimized structures were utilized for molecule docking.

## 2.2 Molecular Docking

Docking of various nucleobase, nucleoside, and nucleotide drugs into the active site of the RdRp-Mg<sup>+2</sup> complex was carried out by using the GOLD 5.0 program [32-34]. The genetic algorithm [32] was used for the docking purpose to create ten different conformations of the drug molecule by keeping the protein rigid. The binding site was considered to be situated within a radius of 10Å from the conserved Thr680. The Chemscore was used for docking, and the PLP score was used to rescore the binding of different drugs with the RdRp-Mg<sup>+2</sup>

complex. Out of the ten different poses, the one whose purine or pyrimidine interacts with the residues of motif F (Lys545, Arg553, Arg553, etc) and the sugar-phosphate backbone interacts with the  $Mg^{+2}$  and residues of motif C (Asp760, Asp761, etc) like the RdRp-RMP (pdb ID 7BV2) [16] and Hepatitis C virus protein (HCV)-Sofosbuvir monophosphate (SMP) (pdb id 4WTG) [35] complexes were considered for further analysis.

To examine if the above docking protocol can reproduce experimentally observed binding modes, the RMP was docked into the binding site of the RdRp- $Mg^{+2}$  protein of the SARS-CoV-2 (pdb id 7BV2) [16] after deleting all cofactors. As illustrated in Fig.1b, the above docking protocol generated the experimental binding mode accurately. It was found that the purine ring of docked RMP can make two hydrogen bonds with the Arg553 (3.6 Å) and Lys545 (3.8 Å), while the sugar group can make a hydrogen bond with Thr680 (3.8 Å) as was observed in the cryo-EM structure (pdb ID 7BV2) (Fig. 1a,b). To re-verify the docking protocol, the SMP was docked into the active site of the HCV protein (pdb id 4WTG) [35] after removing all cofactors. Interestingly, the exact experimental binding mode was reproduced (Fig 1c). These results motivated us to use the same protocol for docking of various drugs into the active site of the RdRp- $Mg^{+2}$  complex of the SARS-CoV-2.

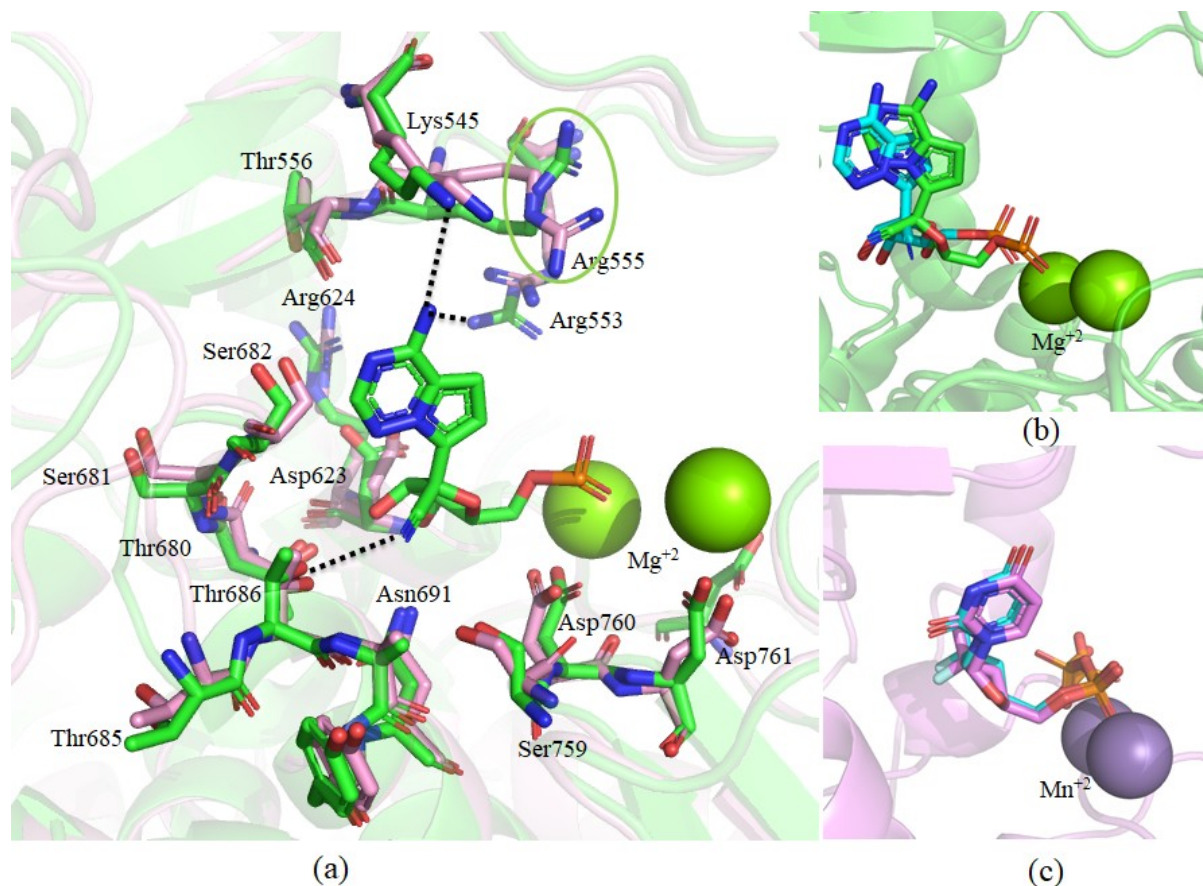


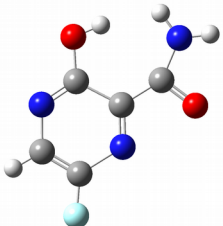

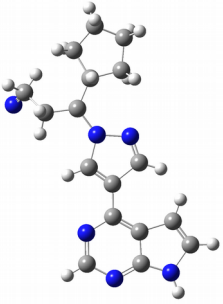
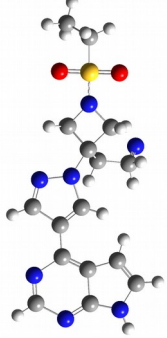
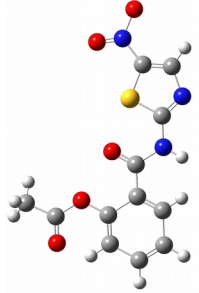
Fig. 1: (a) Superposition of the RdRp without metal ions (in pink, pdb id 6M71 ) onto the RdRp-Mg<sup>2+</sup>-RMP complex (in green, pdb id 7BV2). The Arg555 moving away from the active site in the latter structure is marked by a green circle. The hydrogen-bonding interactions (dotted lines) of RMP with different residues of the RdRp protein are also depicted to explain its binding mode. The comparison of docked (in cyan) and experimental binding modes of (b) AMP (in green, pdb id 7BV2), and (c) SMP (in violet, pdb id 4WTG) are also shown.

### 3. Results and Discussions


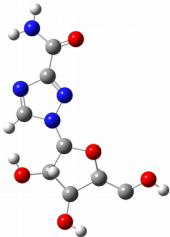
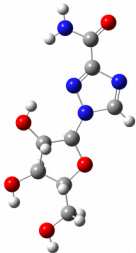

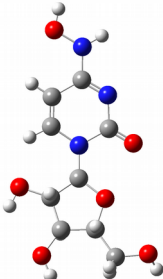
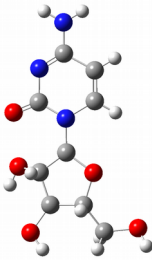
#### 3.1 Binding modes of nucleobase-based drugs

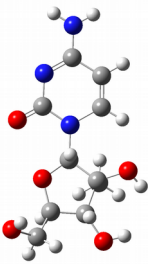
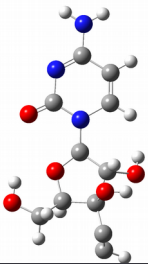
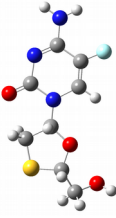
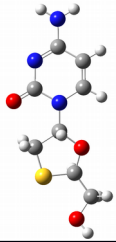
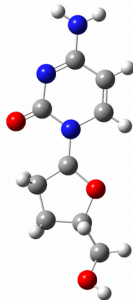
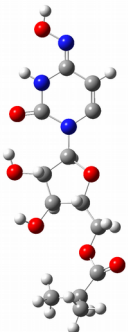
Among the different nucleobase drugs, the Favipiravir, Triazavirin, Ruxolitinib, Baricitinib, and Nitazoxanide were considered for docking to examine their binding patterns with the RdRp protein. The first one is a pyrimidine-based drug, while others are purine-based. The structure of these drugs, their docking fitness, and hydrogen bonding residues are presented in Table 1. Their docked binding modes are depicted in Fig. 2.

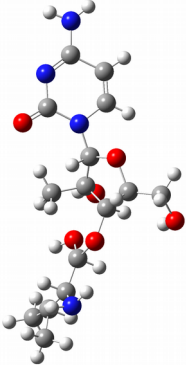
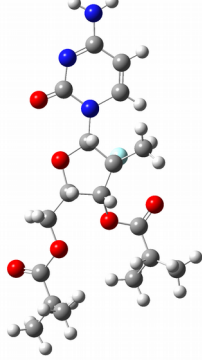

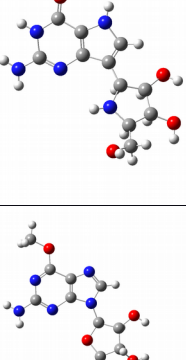
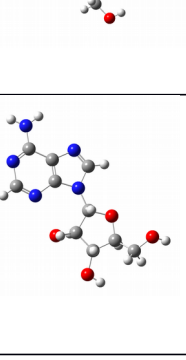

Table 1: The list of repurposed drugs, their DRUGBANK accession number, structure, docking fitness, and the residues of the RdRp protein with which they can make direct hydrogen bonds. The protein target of these drugs and their current use are also mentioned against their name.

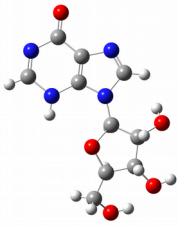
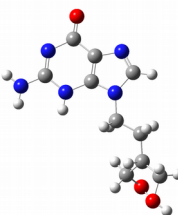
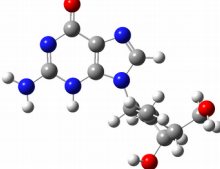
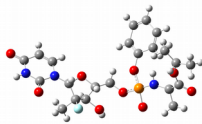
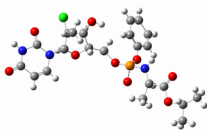
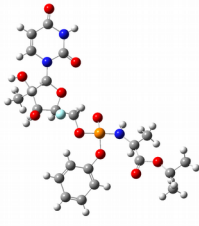
S. no	Drug Name	Accession number	Structure	Docking fitness	H-Bonds
1	Favipiravir  (inhibits replication of influenza A and B)	DB12466		27.98	Arg553, Asp623, Asp760
2	Triazavirin  (inhibits replication of influenza A and B)	DB15622		20.00	Arg553, Cys622
3	Ruxolitinib  (inhibits janus-associated kinase to treat bone marrow cancer)	DB08877		32.36	Lys545, Asn691, Thr680, Thr685
4	Baricitinib  (inhibits Janus kinase 1 (JAK1) and 2 (JAK2) to cure autoimmune disorders, such as rheumatoid arthritis)	DB11817		37.97	Lys545, Arg553, Asn691, Thr680, Pro618
5	Nitazoxanide  (anti-infective drug that modulates the survival, growth, and proliferation extracellular and intracellular protozoa, helminths, anaerobic and microaerophilic bacteria, and viruses)	DB00507		32.27	Cys622, Thr687, Gly683

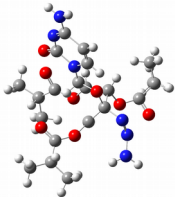
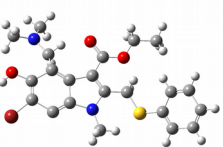
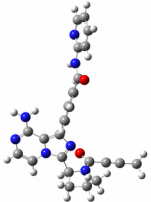
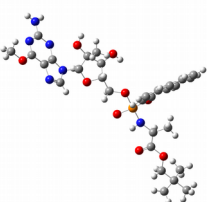



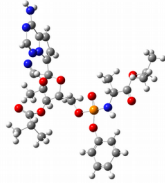
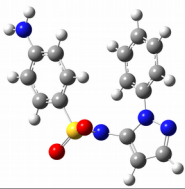
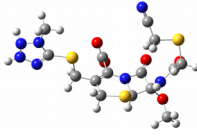
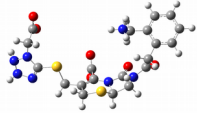
6	Clevudine  (investigated for the treatment of hepatitis virus B infection)	DB06683		40.74	Thr556, Tyr619, Arg553, Asp760, Cys622, Lys621
7	Ribavirin  (antiviral agent that interferes with the synthesis of viral mRNA)	DB00811		35.36	Asp623, Cys622, Tyr619, Asp760 Arg624, Lys621
8	Levovirin  (investigated for use/treatment in hepatitis virus C)	DB06523		32.69	Asp623, Cys622, Asp760
9	Taribavirin  (investigated for the treatment of hepatitis virus C).	DB06408		31.79	Asp623, Arg553
10	N4-hydroxycytidine  (induces mutations in RNA virions)	DB15660		42.78	Thr556, Arg553, Tyr619, Asp760, Asp623
11	Cytarabine  (used for the treatment of leukemia, especially acute non-lymphoblastic leukemia)	DB00987		40.97	Thr556, Arg553, Tyr619, Asp760, Asp623

12	2'-C-methylcytidine  (inhibits hepatitis C virus (HCV) activity)	DB13921		41.13	Thr556, Asp623, Asp760, Tyr619
13	TAS-106  (possesses antitumor activity)	DB06656		25.41	Arg553, Asp760
14	Emtricitabine  (prevents HIV I infection by inhibiting the reverse transcriptase)	DB00879		37.76	Thr556, Arg553, Tyr619
15	Lamivudine  (prevents HIV I infection by inhibiting the reverse transcriptase)	DB00709		37.36	Arg553, Thr556, Tyr619, Cys622,
16	Zalcitabine  (inhibits HIV replication by acting as a chain-terminator of viral DNA)	DB00943		39.69	Thr556, Arg553, Arg555, Tyr619
17	EIDD-2801  (inhibits replication of human and bat coronaviruses, including SARS-CoV-2, in mice and human airway epithelial cells)	DB15661		41.04	Arg555, Arg553, Asp760, Lys621, Tyr619

18	<b>Valopicitabine</b>  (inhibits viral RNA chain elongation and viral RNA-dependent RNA polymerase activity)	DB13920		38.77	Thr556, Arg553, Asp623, Tyr619
19	<b>Mericitabine</b>  (inhibits HCV RNA polymerase, an enzyme that is necessary for hepatitis C viral replication)	DB12045		25.50	Thr556, Arg553, Ser759
20	<b>Galidesivir</b>  (possesses antiviral activity against various negative- and positive-sense RNA viruses including coronaviruses, filoviruses, and arenaviruses)	DB11676		31.45	Thr556, Arg553, Asp623, Asp760
21	<b>Immucillin-G</b>  (investigational nucleoside)	DB02230		36.63	Arg553, Asp623, Tyr619, Asp760
22	<b>Nelarabine</b>  (inhibits DNA synthesis and cytotoxicity and is used to treat acute T-cell lymphoblastic leukemia)	DB01280		38.02	Thr556, Arg553, Asp623, Asp760, Tyr619
23	<b>Vidarabine</b>  (possess antiviral activity against infections caused by a variety of viruses such as the herpes viruses, the vaccinia	DB00194		35.93	Arg553, Asp760, Lys545, Tyr619

	virus and varicella zoster virus)				
24	Inosine  (investigational molecule and may possess neurorestorative, anti-inflammatory, immunomodulatory and cardioprotective effects)	DB04335		39.61	Lys545, Arg553, Asp760, Tyr619
25	Penciclovir  (possesses antiviral activity against herpes simplex virus (HSV) infections)	DB00299		34.08	Thr556, Asp760, Arg624, Arg553, Tyr619
26	Entecavir  (antiviral drug used in the treatment of hepatitis B infection)	DB00442		36.92	Lys545, Arg553, Lys624, Asp623
27	Sofosbuvir  (antiviral drug used as part of combination therapy to treat chronic Hepatitis C, an infectious liver disease caused by infection with Hepatitis C Virus (HCV))	DB08934		39.60	Lys621, Asp760, Cys622, Lys551
28	Uprifosbuvir  (under investigation for the treatment of Chronic Hepatitis C Genotype (GT)1 and GT2 Infection)	DB15206		47.55	Thr556, Arg553, Lys621, Asp760
29	Adafosbuvir  (under investigation to evaluate the effect of renal impairment on the pharmacokinetics of AL-335)	DB14906		39.90	Lys545, Arg553, Asp760, Lys551

30	Balapiravir  (under in trials for the treatment of Dengue and Hepatitis C viral infections)	DB05060		41.30	Thr556, Arg553, Arg555, Lys621, Asp760
31	Umifenovir  (antiviral agent used for the treatment and prophylaxis of influenza and other respiratory infections) investigation into its use for a variety of enveloped and non-enveloped RNA and DNA viruses, including <i>Flavivirus</i> , Zika virus, foot-and-mouth disease, Lassa virus, Ebola virus, herpes simplex, hepatitis B and C viruses, chikungunya virus, reovirus, Hantaan virus, and coxsackie virus B5)	DB13609		35.03	Arg555, Arg553, Ser682, Thr687, Asn691
32	Acalabrutinib  (inhibits the Bruton Tyrosine Kinase (BTK) for the treatment of chronic lymphocytic leukemia, small lymphocytic lymphoma)	DB11703		46.49	Thr556, Thr687, Asp760, Lys551, Asn691
33	BMS-986094  (in trials studying the treatment of Hepatitis C, HCV (Genotype 1), and Hepatitis C Virus)	DB11966		41.76	Lys545, Lys621, Arg553, Asp760, Ser759
34	Remdesivir  (used for the treatment of Ebola and coronavirus family of viruses including the SARS-Cov-2)	DB14761		46.73	Ser682, Arg553, Asp760

35	GS-6620  (under trial for antiviral activity for the treatment of chronic Hepatitis C virus infection).	DB15222		42.90	Ser622, Arg553, Arg555
36	Sulfaphenazole  (an antibacterial agent)	DB06729		34.39	Asp760
37	Cefmetazole  (Possesses antibacterial activity against both gram-positive and gram-negative microorganisms. It has a high rate of efficacy in many types of infection and to date no severe side effects have been noted).	DB00274		31.61	Lys545, Arg553, Ser682
38	Ceforanide  (anti-bacterial agent)	DB00923		45.33	Ser682, Arg553, Lys551, Thr687

As evident from Fig. 2, Favipiravir, an inhibitor of influenza A and B viruses [36] binds with the RdRp protein by making strong hydrogen-bonding interactions with Arg553, Asp623, and Asp760. Its binding position in the RdRp active site is the same as the sugar moiety of the RMP (Fig. 2c) [16]. This may be the reason for which it was observed to inhibit the RdRp activity in the cell-based assays. Interestingly, all of the purine-based drugs are found to bind with the RdRp protein like the RMP. In these drugs, the purine ring makes hydrogen bonding interactions mainly with the Lys545, while the extended five-membered rings make hydrogen bonding interactions with Thr680, Thr685, Asn691, Cys622, and Pro618 (Table 1). It is further found that the CO, S-CH<sub>3</sub>, SO<sub>2</sub>, and NO<sub>2</sub> groups in Favipiravir, Triazavirin,



Baricitinib, and Nitazoxanide, respectively, will make interactions with one of the  $Mg^{+2}$  ions. As Baricitinib makes more number of hydrogen bonding interactions and its docking score is more than that of other drugs (Table 1), it may be proposed that the Baricitinib, which is an inhibitor of the JANUS kinase and is used for the treatment of rheumatoid arthritis [37] would serve as a better inhibitor of the RdRp protein than that of Fivapiravir and other nucleobase drugs studied here.

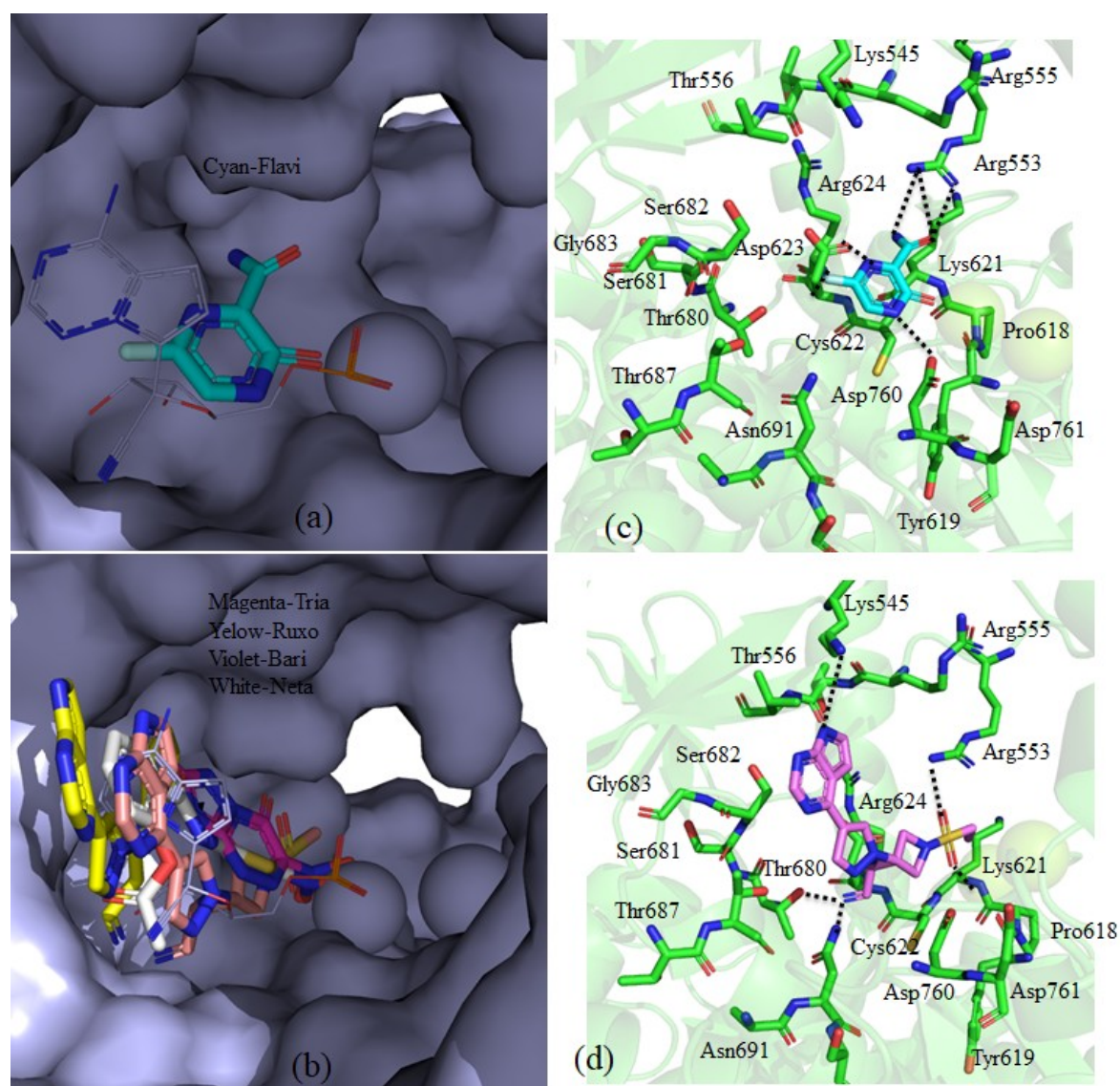


Fig. 2: The binding modes of different nucleobase drugs (represented by sticks) with the RdRp protein (expressed by surface). These drugs include (a) Favipiravir and (b) Triazavirin, Ruxolitinib, Baricitinib, and Nitazoxanide. The first four letters of each drug are mentioned against the color code used to represent them. The binding of (c) Favipiravir and (d) Baricitinib with different amino acids of the RdRp protein are also depicted. The dotted lines

indicate hydrogen-bonding interactions. For comparison, the crystal structure of the RMP (in light blue) is also shown by the line representation.

### 3.2 Binding modes of pyrimidine-based nucleoside drugs

Nucleoside-based drugs can be divided into pyrimidine-based and purine-based. It is found that most of the pyrimidine based drugs, such as Clevudine, Taribavirin, N4-hydroxycytidine (EIDD-1931), Cytarabine, 2'-C-methylcytidine, Emtricitabine, Lamivudine, Zalcitabine, EIDD-2801, and Valopicitabine (Table 1) bind to the RdRp protein in a similar manner (Fig. 3a). Among these drugs, the clevudine, N4-hydroxycytidine (Fig. 3b), EIDD-2801 (Fig. 3c) and 2'-C-methylcytidine are found to bind tightly with the RdRp protein by making strong hydrogen bonds with mainly Thr556, Arg553, Asp623, Asp760 and Tyr619 and possess the highest docking scores (Table 1). They also interact strongly with the  $Mg^{+2}$  ions. Hence, these drugs will inhibit RdRp activity very strongly. This is in agreement with the cell-based assays, where EIDD-2180 was found to inhibit RdRp about 3-6 times more than that of the Remdesivir [19]. Similarly, the N4-hydroxycytidine has been found to inhibit the Venezuelan equine encephalitis virus, human coronavirus HCoV-NL63 [38], bat coronaviruses and the SARS-CoV-2 in mice [39] and human airway epithelial cells [39]. Other pyrimidine-based drugs, such as Ribavirin, Levovirin, and TAS-106, are found to bind RdRp slightly differently than the former pyrimidine-based drugs (Fig. 3d). Among these three drugs, Ribavirin is found to be more promiscuous (Fig. 3e). However, its pyrimidine ring is unable to make any hydrogen bonding interactions with the F-block residues (Fig. 3e). Interestingly, the five-membered ring of the Mericitabine binds like the Cytarabine, while the extended carbon rings connected to the sugar moiety helps it to point toward Lys551 and Ser814 (Fig. 3f). Due to this reason, their interaction with the  $Mg^{+2}$  ions is weaker compared to other pyrimidine-based drugs. If we compare the binding of all of these drugs with RdRp, it is clear that the clevudine, N4-hydroxycytidine, EIDD-2801, and 2'-C-methylcytidine would be more



effective than that of other drugs including Favipiravir and Ribavirin that are proposed to be useful for the treatment of the SARS-CoV-2 infection.

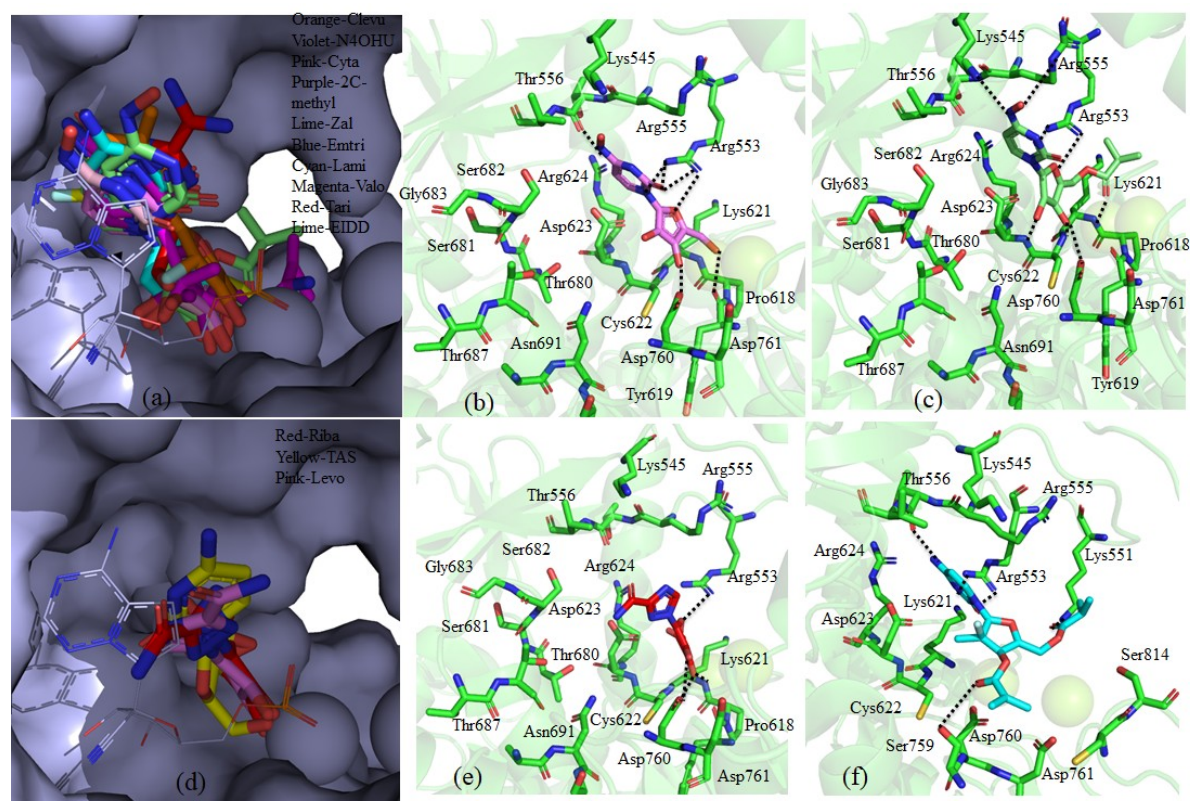


Fig. 3: The binding modes of different pyrimidine-based nucleoside drugs (represented by sticks) with the RdRp protein (represented by surface). These drugs include (a) Clevudine, Taribavirin, N4-hydroxycytidine, Cytarabine, 2'-C-methylcytidine, Emtricitabine, Lamivudine, Zalcitabine, EIDD-2801, and Valopicitabine. The interactions of (b) N4-hydroxycytidine and (c) EIDD-2801 with different amino acids are illustrated. (d) The binding modes of Ribavirin, Levovirin, and TAS-106 (in the sticks) with RdRp (in the surface) are also shown. The interactions of (e) Ribavirin and (f) Mericitabine with different important amino acids are depicted. The dotted lines show important hydrogen bonds. Color codes used for different drugs are mentioned in (a) and (d). The line representation shows the crystal structure of the RMP (in light blue) for the comparison of different binding modes.

### 3.3 Binding modes of purine-based nucleoside drugs

Binding modes of Galidesivir, Immucillin-G, Nelarabine, Vidarabine, Penciclovir, Entecavir, and Inosine are shown in Fig. 4. Except Immucillin-G, Entecavir, and Inosine all of these drugs bind to the RdRp protein almost similarly. The vital amino acids making hydrogen bonds with these drugs are summarized in Table 1. Among these drugs, Nelabarine and Inosine have the highest docking fitness and can form strong hydrogen bonds with the

protein. Hence, these purine-based nucleosides will act as better inhibitors. However, if we compare these purine-based nucleoside drugs with those of pyrimidine-based nucleoside drugs, it is found that the latter drugs, in particular, N4-hydroxycytidine and EIDD-2801 would act as better antiviral drugs.

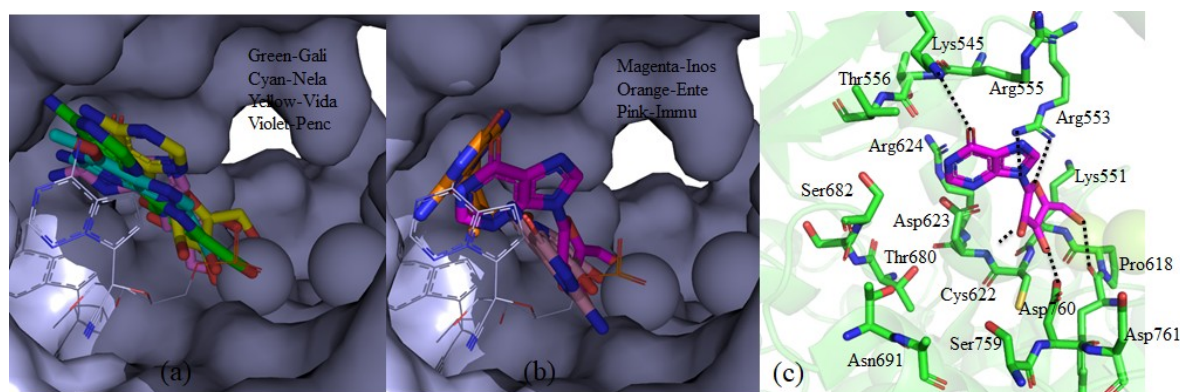


Fig. 4: The binding modes of (a) Galidesivir, Nelarabine, Vidarabine, and Penciclovir, (represented by sticks) and (b) Immucillin-G, Entecavir, and Inosine (represented by sticks) with the RdRp protein (represented by surface). (c) The dotted lines show important hydrogen bonds made by Inosine with different amino acids of the RdRp protein. Color codes used for different drugs are mentioned in (a) and (b). The crystal structure of the RMP (in light blue, pdb id 7BV2)) is shown by the line representation for comparison of different binding modes.

### 3.4: Binding modes of Purine- and Pyrimidine-based drugs with modified backbones

There are several purine- and pyrimidine-based antiviral drugs available, whose sugar-phosphate backbones are replaced with extended synthetic rings to prohibit their replication in cells, such as Sofosbuvir, Uprifosbuvir, Remdesivir, etc (Table 1). It should be mentioned that the recent cryo-EM study of the RdRp-Remdesivir monophosphate complex [16] does not account for the full structure of the drug mentioned in the DRUGBANK database. Similarly, the structure of the HCV-Sofosbuvir monophosphate (pdb id 4WTG) [35] does not include the full structure of the Sofosbuvir stored in the DRUGBANK database. Hence, this study is expected to unravel the detailed binding modes of these full drugs with the RdRp protein. It is found that among these drugs, the pyrimidine-based drugs, such as Sofosbuvir, Uprifosbuvir, Adafosbuvir, and Balapiravir follow a similar binding pattern with the RdRp

protein (Fig. 5a). In contrast, purine-based drugs, such as Umifenovir, Acalabrutinib, BMS-986094, Remdesivir, and GS-6620 follow a slightly different binding pattern (Fig. 5b,c). The detailed interactions of Balapiravir (pyrimidine-based) and Remdesivir (purine-based) with various amino acids of the RdRp protein are depicted in Fig. 6. As can be found from this figure, Balapiravir makes hydrogen bonding interactions with Thr556, Arg553, and Asp760, while the extended ring containing CH<sub>3</sub> groups stack opposite Lys551, Asp761, and Cys813 (Fig. 6a). Similarly, the purine ring of the Remdesivir binds somehow similarly as observed in the Cryo-EM study. The purine ring makes hydrogen bonds with Ser682 and Arg553. The amino acid residues Thr556, Lys545, Arg555, and Thr687 are located within a distance of 4-6 Å. The 6-membered ring connected to the phosphate group is making stacking interaction with Arg553. Similarly, the end ring containing two CH<sub>3</sub> groups stack opposite Phe812 and Cys813. Due to more interactions with the RdRp protein, Uprifosbuvir, Balapiravir, Acalabrutinib, BMS-986094, Remdesivir, and GS-6620 may act as a potent inhibitor of the RdRp protein.

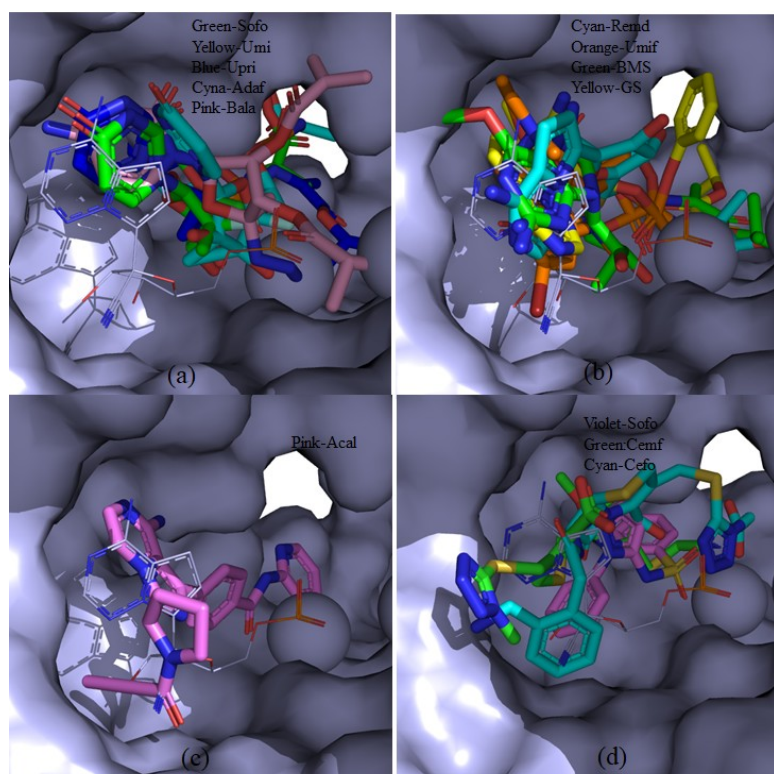




Fig. 5: The binding modes of (a) Sofosbuvir, Uprifosbuvir, Adafosbuvir, Balapiravir, (b) Umifenovir, BMS-986094, Remdesivir, GS-6620, (c) Acalabrutinib, and (d) Sulfaphenazole, Cefmetazole, and Ceforanide. The drug molecules are shown in the stick format and the protein is shown in the surface representation. The color codes are mentioned against the drug molecules (first four letters are mentioned). The experimental structure of the RMP (in light blue, pdb id 7BV2) is shown in the line representation for comparison.

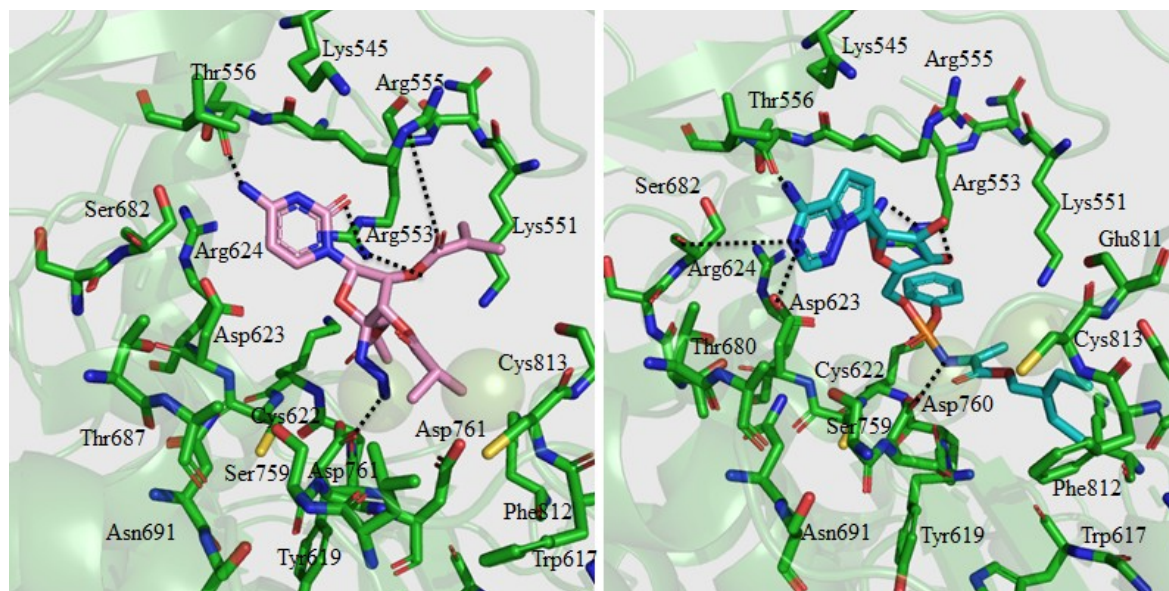


Fig. 6: The interactions of (a) Balapiravir and (b) Remdesivir with different amino acids with the RdRp protein. The hydrogen-bonding interactions are shown by dotted lines.

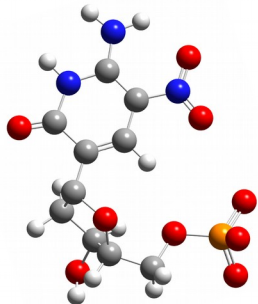
To examine if non-nucleoside-based drugs can bind with RdRp like the nucleoside-based drugs, Sulfaphenazole, Cefmetazole, and Ceforanide, which are semi-synthetic antibacterial and anti-infective drugs [40] were docked into the binding sites of the RdRp protein. It is found that among these drugs, Ceforanide, binds tightly with the protein and makes strong hydrogen bonds with Ser682, Arg553, Lys551, and Thr687. Hence, it may also inhibit the RdRp viral protein.

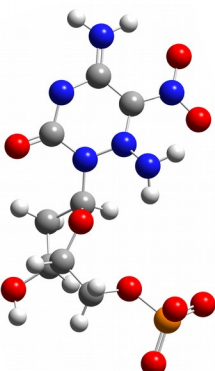
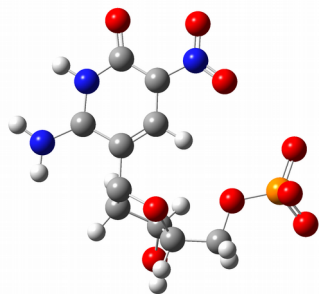
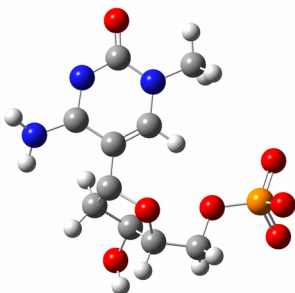
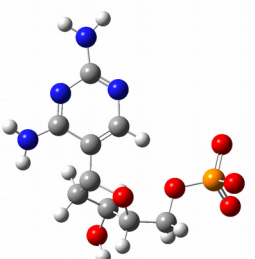
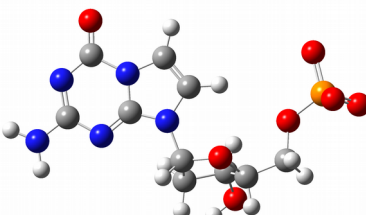
### 3.5: Binding modes of artificial nucleotides

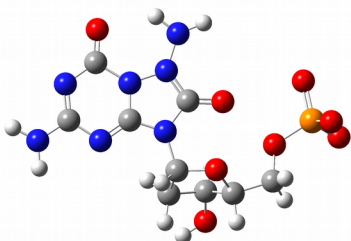
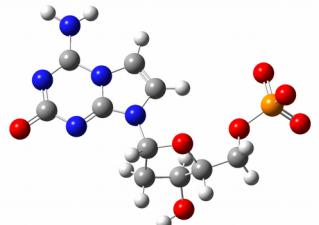
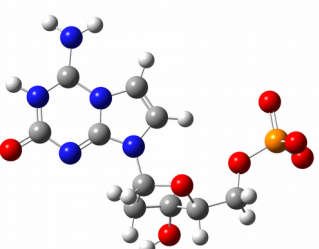
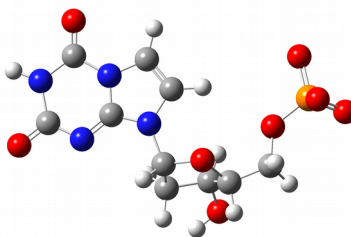
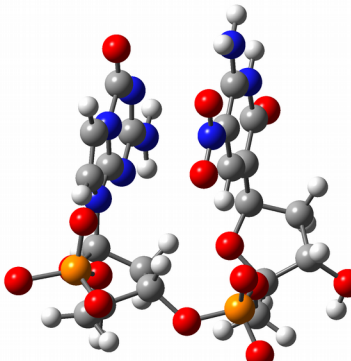
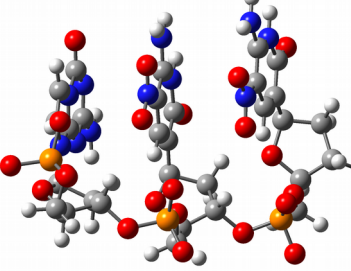
Artificial nucleotides are important nucleotides that are currently examined for their usefulness in forming functional genes [21-28]. To investigate if these nucleotides can inhibit the RdRp protein, their binding patterns with the RdRp protein were evaluated. The structures, docking fitness, and hydrogen bonding residues involving these nucleotides are

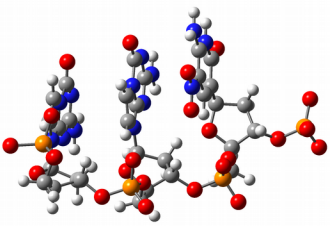
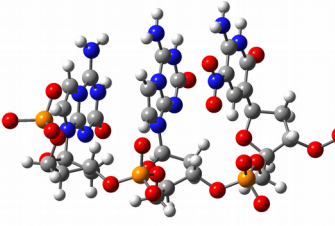
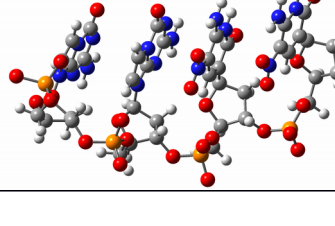
presented in Table 2. The binding modes of these artificial nucleotides are illustrated in Fig. 7. As can be seen from this figure, the monophosphate forms of different pyrimidines, such as dZ, dZ5, dV, dS, and dK possess a similar binding mode. Similarly, the monophosphate forms of purines, such as dP, dP4, dJ, dB, and dX follow a similar binding mode. As depicted in Fig. 7b,c, dZ and dP, the most studied AEGIS nucleotides can bind with the F-block residues almost identically. The same trend is also found in the case of other artificial nucleotides. The residues that make hydrogen bonds with these unnatural nucleotides are Lys545, Arg553, Arg555, Lys621, Cys622, Asp623, Thr680, Ser682, Thr687, Asn691, Ser759, and Asp760 (Fig. 7c,d, Table 2). The phosphate groups are also found to make tight ionic interactions with the  $Mg^{+2}$  ions. If we compare binding modes, number of hydrogen bonds, salt-bridge, and other interactions and docking fitness of these nucleotides with those of the antiviral drugs studied here, it is evident that these artificial nucleotides would act as better inhibitors of the RdRp viral protein.

Table 2: Table 1: The list of various AEGIS nucleotides, their structure, docking fitness, and the residues of the RdRp protein with which they can make direct hydrogen bonds.

S. No.	Sequence	Structure	Docking Fitness	H-bonds
1	dZ		46.54	Lys545, Thr556, Arg553, Lys612 Cys222

2	dZ5		27.34	Lys545, Thr556, Arg553, Ser682
3	dV		48.40	Lys545, Arg553, Arg624, Arg555, Ser759, Thr680, Asp623
4	dS		53.43	Lys545, Arg553, Asp623, Cys622, Lys621
5	dK		44.30	Lys545, Arg553, Asp623, Cys622, Lys621
6	dP		56.56	Lys545, Arg553, Cys622, Lys621, Asp621

7	dP4		43.90	Arg553, Cys622, Cys622, Lys621
8	dJ		54.43	Lys545, Thr556, Arg553, Cys622, Lys621
9	dB		52.06	Lys545, Arg553, Cys622, Asp623, Ser682, Lys621, Cys622
10	dX		35.13	Lys545, Thr680, Asn691, Ser759
11	5'-dPdZ-3'		59.50	Arg553, Thr556, Thr680, Thr687, Ser682, Ser759,
12	5'-dPdZdZ-3'		53.71	Lys545, Arg553, Arg555, Ser682, Cys622, Asp760, Ser759

13	5'- dPdPdZ- 3'		55.40	Lys545, Ser682, Arg553, Ser759, Lys621
14	5'- dBdBdZ- 3'		35.12	Lys545, Thr680, Arg553, Thr687, Asn691, Ser759
15	5'- dPdPdZd Z-3'		21.14	Lys545, Arg553, Arg555, Asp760, Lys551, Asp760, Tyr619,

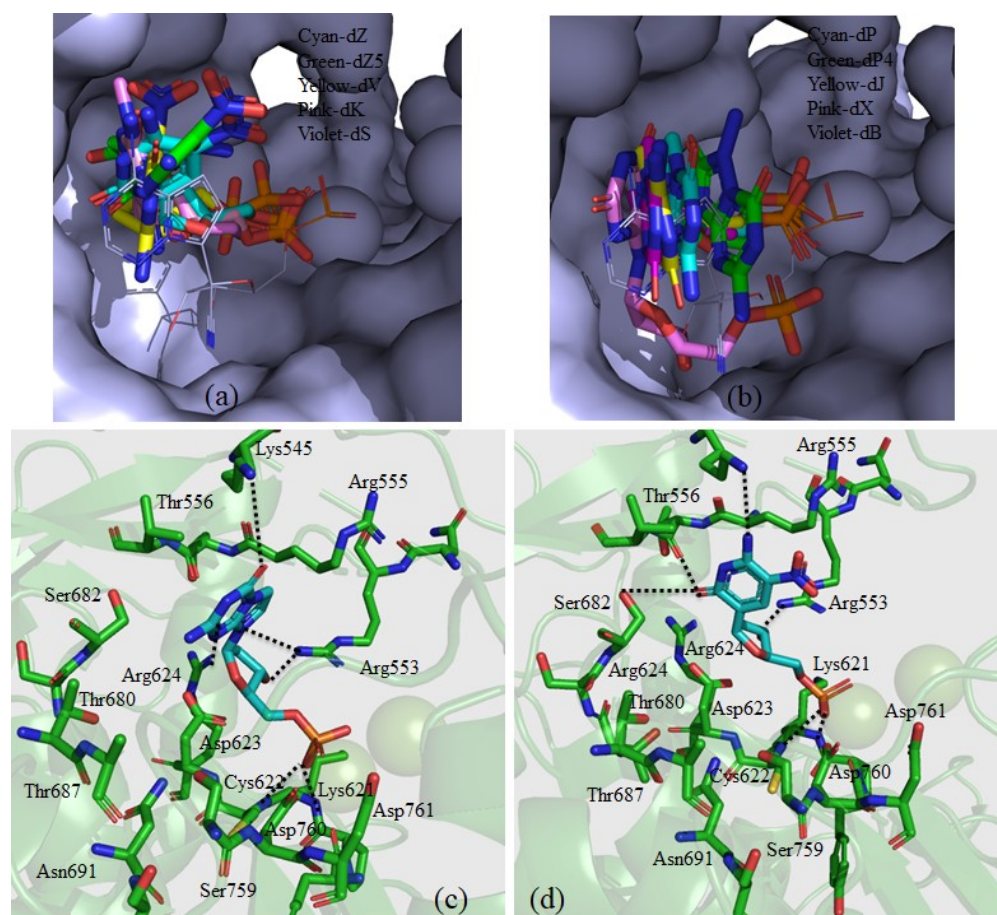


Fig. 7: The binding modes of different (a) pyrimidine-based (in the stick) and (b) purine-based (in the stick) artificial nucleotides with the RdRp protein (in surface representation).



The detailed interactions of (c) dP and (d) dZ with different amino acid residues of the RdRp protein are also shown. Dotted lines show the hydrogen bonding interactions.

To further evaluate the binding of more than one artificial nucleotide with the RdRp protein (Table 2), different sequences of these nucleotides were docked into the binding site of the RdRp protein. The docking fitness and hydrogen bonding residues are presented in Table 2. The binding modes of these nucleotides are shown in Fig. 8. As can be found from this figure, the 5'-dPdZ-3' can make tight interactions with the F-block residues and the phosphate group can make stable salt-bridge interactions with the  $Mg^{+2}$  ion (Table 2, Fig. 8a). The dP also makes stacking interactions with dZ. However, the interbase stacking interaction gets lost with the increasing size of the sequence (Fig. 8 b-e). Although nucleobases can make more number of contacts with the protein in large sequences, the phosphate- $Mg^{+2}$  interaction becomes weak. This is evident from the docking fitness (Table 2). These results suggest that more than three nucleotide sequences may reduce the inhibitory activity of the artificial nucleotides.

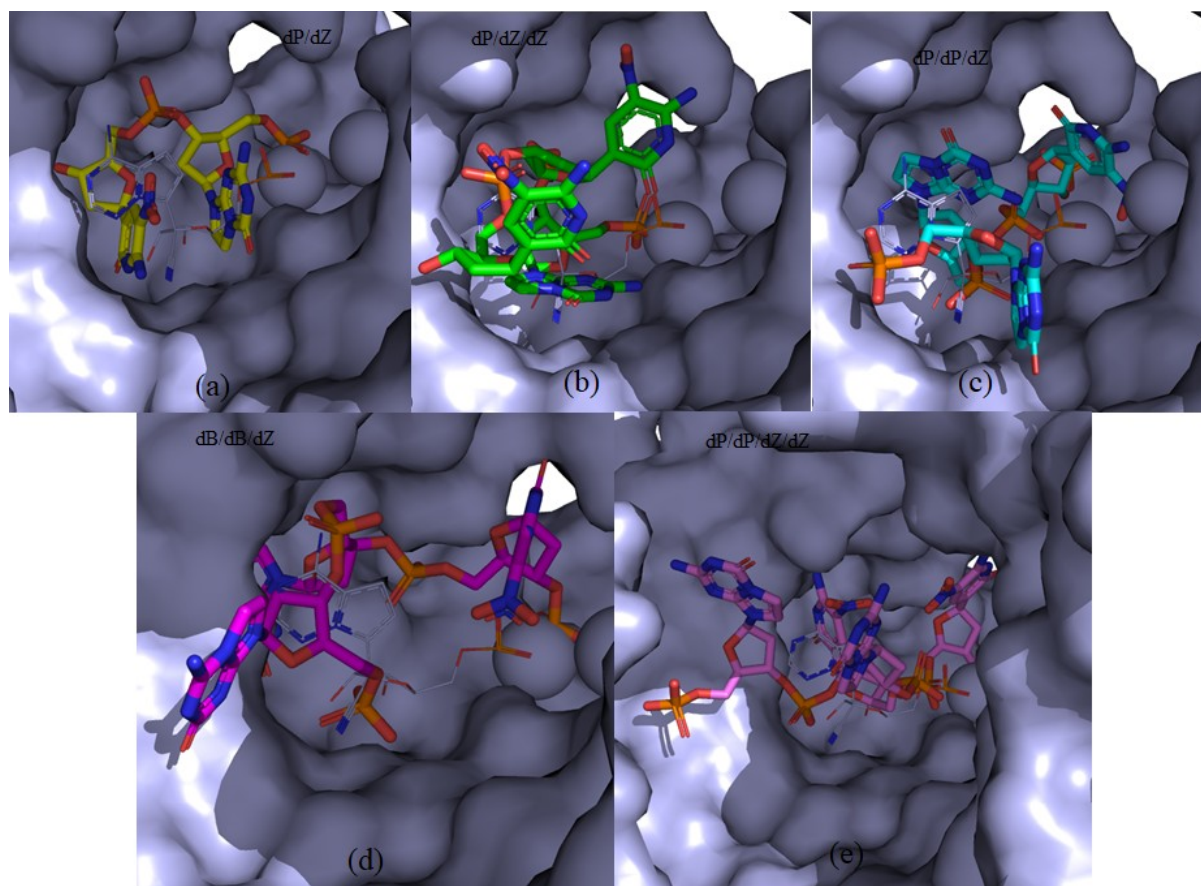


Fig. 8: The binding modes of (a) 5'-dPdZ-3', (b) 5'-dPdZdZ-3', (c) 5'-dPdPdZ-3', (d) 5'-dBdBdZ-3', and (e) 5'-dPdPdZdZ-3' with the RdRp protein (in surface representation).

#### 4. Conclusions

It is revealed that all the approved and investigational drugs studied here can bind to the RdRp viral protein of the SARS-CoV-2. However, due to the higher binding affinity, some of the drugs would have more inhibitory activity and can be immediately used for the treatment of the Wuhan virus infections after a careful clinical trial. These drugs include cleavudine, N4-hydroxycytidine, 2'-C-methylcytidine, EIDD-2801, Uprifosbuvir, Balapiravir, Acalabrutinib, BMS-986094, Remdesivir, GS-6620, and Ceforanide. It is also found that the artificial nucleotides that were invented to make functional genes may have better inhibitory activity than that of the above drugs. Among these nucleotides, dB, dJ, dP, dK, dS, dV, dZ are found to make tight interactions with the RdRp protein. It is further found that sequences containing

up to three artificial nucleotides can also inhibit the RdRp viral protein. Among different residues, Lys545, Arg553, Arg555, Thr556, Tyr619, Lys621, Cys622, Asp623, Arg624, Thr680, Ser682, Thr687, Asn691, and Asp760 are found to play a vital role in stabilizing different inhibitors. Hence, the design of new inhibitors should be done in such a way that they can make interactions with these residues.

## Acknowledgment

NRJ is thankful to the Science and Engineering Research Board (SERB) of the Department of Science and Technology (DST, NewDelhi) for financial support.

## References

1. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng Clinical characteristics of 138 hospitalized patients with 2019 coronavirus-infected pneumonia in Wuhan, China, *JAMA*, **2020**, 323, 1061-1069.
2. A.R. Sahin, A. Erdogan, P.M. Agaolglu, Y. Dineri, A. Y. Cakirci, M. E. Senel; R. A. Okyay, A. M. Tasdogan, 2019 Novel coronavirus (COVID-19) outbreak: a review of the current literature. *Eur. J. Med. Oncol.* **2020**, 4, 1-7.
3. C. Liu, Q. Zhou, Y. Li, L. V. Garner, S.P. Watkins, L. J. Carter, J. Smoot, A. C. Gregg, A. D. Daniels, S. Jerve, D. Albaiu, Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Cent. Sci.* **2020**, 6, 315-331.
4. U.J. Buchholz, A. Bukreyev, L. Yang, E.W. Lamirande, B.R. Murphey, K. Subbarao, P.L. Collins, Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 9804-9809.
5. T. P. Sheahan, A.C. Sims, S.R. Leist, A. Schafer, J. Won, A.J. Brown, S.A. Montgomery, A. Hogg, D. Babusis, M. O. Clarke, J.E. Spahn, L. Bauer, S. Sellers, D. Porter, J.Y. Feng, T. Cihlar, R. Jordan, M.R. Denison, R.S. Baric, Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. *Nat. Commun.* **2020**, DOI: 10.1038/s41467-019-13940-6.
6. R. U. Kadam, I.A. Wilson, Structural basis of influenza virus fusion inhibition by the antiviral drug Arbidol. *Proc. Natl. Acad. Sci. U.S. A.* **2017**, 114, 206-214.
7. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). <https://www.nature.com/articles/d41573-020-00016-0>.

8. The efficacy of Lopinavir Plus Ritonavir and Arbidol against novel coronavirus Infection (ELACOI). <https://clinicaltrials.gov/ct2/show/NCT04252885>.
9. S. Pant, M. Singh, V. Ravichandran, U.S.N. Murty, H. Srivastava, Peptide-like and small molecule inhibitors against COVID-19. *Biomol. Struct. Dyn.* **2020**, DOI: 10.1080/07391102.2020.1757510.
10. C. Cava, G. Bertoli, I. Castiglioni, In silico discovery of candidate drugs against COVID-19, *Viruses*, **2020**, 12, 404.
11. Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, F. Cheng, Network-based repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery*, **2020**, 6, 14.
12. R. Wu, L. Wang, H.C.D. Kuo, A. Shannar, R. Peter, P. J. Chou, S. Li, R. Hudlikar, X. Liu, Z. Liu, G. J. Poiani, L. Amorosa, L. Brunetti, A.N. Kong, *Curr. Pharmacol. Rep.* **2020**, DOI:10.1007/s40495-020-00216-7
13. M. Ceccarelli, M. Berretta, E.V. Rullo, G. Nunnari, B. Cacopardo, Differences and similarities between severe acute respiratory syndrome (SARS)-coronaVirus (CoV) and SARS-CoV-2. Would a rose by another name smell as sweet? *Eur. Rev. Med. Pharmacol. Sci.* **2020**, 24, 2781-2783.
14. S.J.R. Da Silva, C.T.A. Da Silva, R.P.G. Mendes, L. Pena, Role of non-structural proteins in the pathogenesis of SARS-CoV-2. *J. Med. Virol.* **2020**, DOI: 10.1002/jmv.25858.
15. E.J. Snijder, E. Decroly, J. Ziebuhr, The non-structural proteins directing coronavirus RNA synthesis and processing. *Adv. Virus Res.* **2016**, 96, 59-126.
16. W. Yin, *et al.* Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*, **2020**, DOI: 10.1126/science.abc1560.
17. Y. Gao, *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, **2020**, DOI:10.1126/science.abb7498.
18. C.-C. Lu, M.-Y. Chen, Y.-L. Chang, Potential therapeutic agents against COVID-19: What we know so far. *J. Chin. Med. Assoc.* **2020**, DOI:10.1097/JCMA.0000000000000318.
19. T. P. Sheahan, A. C. Sims, S. Zhou, R. L. Graham, A. J. Pruijssers, M. L. Agostini, S. R. Leist, A. Schäfer, K. H. Dinno 3rd, L. J. Stevens, J. D. Chappell, X. Lu, T. M. Hughes, A. S. George, C. S. Hill, S. A. Montgomery, A. J. Brown, G. R. Bluemling, M. G. Natchus, M. Saindane, A. A. Kolykhalov, G. Painter, J. Harcourt, A. Tamin, N. J. Thornburg, R. Swanstrom, M. R. Denison, R. S. Baric, An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci. Transl. Med.* **2020**, 12, eabb5883.
20. D.S. Wishart C. Knox, A.C. Guo AC, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, Drugbank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **2006** 34 (Database issue), D668-D672.
21. S.A. Benner, Z. Yang, F. Chen, Synthetic biology, tinkering biology, and artificial biology. what are we learning? *Comptes Rendus Chim.* **2011**, 14, 372-387.
22. K.K. Merritt, K.M. Bradley, D. Hutter, M.F. Matsuura, D.J. Rowold, S.A. Benner, Autonomous assembly of synthetic oligonucleotides built from an expanded DNA alphabet. total synthesis of a gene encoding kanamycin resistance. *Beilstein J. Org. Chem.* **2014**, 10, 2348-2360.
23. E. Biondi, S. A. Benner, Artificially expanded genetic information systems for new aptamer technologies. *Biomedicines* **2018**, 6, 53.

24. L. Zhang, *et al.* Evolution of functional six-nucleotide DNA. *J. Am. Chem. Soc.* **2015**, 137, 6734-6737.
25. J.J. Voegel, S.A. Benner, Nonstandard hydrogen bonding in duplex oligonucleotides. the base pair between an acceptor-donor-donor pyrimidine analog and a donor-acceptor-acceptor purine analog. *J. Am. Chem. Soc.* **1994**, 116, 6929-6930.
26. N.R. Jena, P. Das, B. Behera, P.C. Mishra, Analogues of P and Z as efficient artificially expanded genetic information system. *J. Phys. Chem. B* **2018**, 122, 8134-8146.
27. B. Behera, P. Das, N.R. Jena, Accurate base pair energies of artificially expanded genetic information systems (AEGIS): clues for their mutagenic characteristics. *J. Phys. Chem. B*, **2019**, 123, 6728-6739.
28. N.R. Jena, Electron and hole interactions with P, Z, and P:Z and the formation of mutagenic products by proton transfer reactions, *Phys. Chem. Chem. Phys.* **2020**, 22, 919-931.
29. M.M. Georgiadis, I. Singh, W.F. Kellett, S. Hoshika, S.A. Benner, N.G. Richards, Structural basis for a six nucleotide genetic alphabet. *J. Am. Chem. Soc.* **2015**, 137, 6947-6955.
30. GaussView, Version 5, R. Dennington, T. Keith, J. Millam, J. Inc. *Semichem*, KS Shawnee Mission 2009.
31. Gaussian09, Revision A.1, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, V. Nakai, T. Vreven, J. Jr. J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O., Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian, Inc., Wallingford CT, 2009.
32. G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.*, **1997**, 267, 727-748.
33. J. W. M. Nissink, C. Murray, M. Hartshorn, M. L. Verdonk, J. C. Cole and R. Taylor, A new test set for validating predictions of protein-ligand interaction. *Proteins*, **2002**, 49, 457-471.
34. M. J. Hartshorn, M. L. Verdonk, G. Chessari, S. C. Brewerton, W. T. M. Mooij, P. N. Mortenson, C. W. Murray, Diverse, High-Quality Test Set for the Validation of Protein-Ligand Docking Performance. *J. Med. Chem.*, **2007**, 50, 726-741.
35. T.C. Appleby, J.K. Perry, E. Murakami, O. Barauskas, J. Feng, A. Chao, D. Fox, D.R. Wetmore, M.E. McGrath, A.S. Ray, M.J. Sofia, S. Swaminathan, T.E. Edwards, Structural basis for RNA replication by the hepatitis C virus polymerase. *Science*, **2015**, 347, 771-775.
36. F.G. Hayden, N. Shindo, Influenza virus polymerase inhibitors in clinical development. *Curr. Opin. Infect. Dis.* **2019**, 32, 176-186.

37. B. Kurriya, M.D. Cohen, E. Keystone. Baricitinib in rheumatoid arthritis: evidence-to-date and clinical potential. *Theor. Adv. Musculoskelet. Dis.* **2017**, 9, 37-44.
38. K. Pyrc, B.J. Bosh, B. Berkhout, M.F. Jebbink, R. Dijkman, P. Rottier, L. van der Hoek. Inhibition of human coronavirus NL63 infection at the early stages of the replication cycle. *Antimicrob. Agents Chemother.* **2006**, 50, 2000-2008.
39. T.P. Sheahan et al. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci. Transl. Med.* **2020**, 12, eabb5883.
40. S.L. Barriere, J. Mills. Ceforanide: antibacterial activity, pharmacology, and clinical efficacy. *Pharmacotherapy*, **1982**, 2, 322-327.