

Attacking the SARS-CoV-2 replication machinery with the Pathogen Box's molecules

Cleidy Osorio-Mogollón¹, Gustavo E. Olivos-Ramírez², Kewin Otazu², Manuel E. Chenet-Zuta^{3*}, Georcki Ropón-Palacios^{2*}, Ihosvany Camps^{2*}, Gabriel M. Jimenez-Avalos⁴, Eduardo Apari-Cossio⁵, Natalia E. Torres-Moreira⁶, Reyna G. Cárdenas-Cárdenas⁷

¹*Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brasil*

²*Laboratório de Modelagem Computacional - LaModel, Instituto de Ciências Exatas- ICEx. Universidade Federal de Alfenas - UNIFAL-MG, Alfenas, Minas Gerais, Brasil*

³*Escuela de Posgrado, Universidad San Ignacio de Loyola, Lima, Perú*

⁴*Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Perú*

⁵*Departamento de Bioingeniería, Universidad de Ingeniería y Tecnología, Perú*

⁶*Departamento de Ciencias de la Vida y Agricultura, Universidad de las Fuerzas Armadas ESPE, Ecuador*

⁷*Facultad de Farmacia y Bioquímica, Universidad Nacional de la Amazonía Peruana, Perú*

Abstract

The world is currently facing a pandemic caused by the new 2019 coronavirus disease (COVID-19), caused by SARS-CoV-2. Among the fundamental processes of this virus are viral transcription and replication. They allow the synthesis of genetic material and the consequent multiplication of the virus to infect other cells or organisms. These are performed by a multi-subunit machinery of various nonstructural proteins (nsp); among which the RNA-dependent RNA polymerase (RdRp or nsp12) is the most important, and, at the same time, conserved among coronaviruses. The structure of this protein (PDB ID: 6M71) was used as a target in the application of computational strategies for drug search, like virtual screening and molecular docking. The region considered for virtual screening has three important amino acids for protein catalysis: T680 (located in Motif A), N691 and D623 (located in Motif B), where a grid box was located. In turn, applying the concept of drug repositioning is considered as a quick response in the treatment of sudden outbreaks of diseases. Here, we used the Pathogen Box, a database of chemical compounds analyzed for the treatment against malaria, which were filtered under the criteria of selecting those that do not present

any violation of Lipinski's Rule of Five. At the same time, the Remdesivir, Beclabuvir and Sofosbuvir drug, previously used in *in silico* and clinical studies for inhibition of nsp12, were used as positive controls. The results showed a Top10 potential target inhibitors, with binding energy (ΔG) higher than those of the positive controls, of which TCMDC-134153 and TCMDC-135052, both with $\Delta G = -7.53 \text{ kcal/mol}$, present interactions with the three important residues of the nsp12 catalytic site. These proposed ligands would be used for subsequent validation by molecular dynamics, where they can be considered as drugs for the development of effective treatments against this new pandemic.

Keywords: SARS-CoV-2, nsp12, RNA polymerase, molecular docking, drug repurposing

1. Introduction

According to the World Health Organization (WHO), a pandemic is defined as the worldwide spread of a new disease, the most frequent being those caused by RNA viruses [1, 2, 3]. Among these viruses, there is the family of coronaviruses (CoV) that, in the last 20 years, have been responsible for serious infections of the respiratory tract [4] such as: 1. The CoV causing the syndrome severe acute respiratory (SARS-CoV) [5]; 2. The CoV causing Middle East respiratory syndrome (MERS-CoV) [6]; and currently, 3. The new CoV called SARS-CoV-2 (due to its similarity to SARS-CoV), causing of coronavirus disease 2019 (COVID-19). The latter has become a global public health emergency, since, to date, it has affected more than 7,570,800 individuals and has caused more than 422,900 deaths worldwide in 188 countries [3].

Among the fundamental processes of SARS-CoV-2 are viral transcription and replication, which allow the synthesis of genetic material and the consequent multiplication of the virus to infect other cells or organisms [7]. These are performed by a multi-subunit machinery of various nonstructural proteins (nsp); among which the RNA-dependent RNA polymerase

*Corresponding authors

Email addresses: manchenet@gmail.com (Manuel E. Chenet-Zuta³), georopon@gmail.com (Georcki Ropón-Palacios²), icamps@unifal-mg.edu.br (Ihosvany Camps²)

(RdRp or nsp12) is the most important [8]. It should be noted that this protein is highly conserved in the coronavirus family [9].

The structure of the SARS-CoV-2 nsp12 has 932 amino acids (residues S1-Q932) and is made up of a nidovirus-associated nucleotidyltransferase (NiRAN) domain, an interface region and a “right-hand” domain [8]. The NiRAN domain consists of 190 amino acids (residues D60-R249). Following this, the interface region (residues A250-R365) is found, which in turn is the junction of the cofactors nsp7 (residues S1-Q83) and nsp8 (residues A1-Q198) [10]. Finally, the “right hand” domain consists of 554 amino acids (residues S367-F920), with polymerase activity, made up of three conserved subdomains of the family of viral polymerases: a finger subdomain (residues L366-A581 and K621-G679), a palm subdomain (residues T582-P620 and T680-Q815) and a thumb subdomain (residues H816-E920) [11].

The “right hand” domain has an active site, formed by the preserved polymerase motifs AG in the palm subdomain. The structure of motif A consists of a β chain and a helix or loop in the subdomain of the fingers and together with motif B, plays a discriminating role between NTP and NTP deoxy [12]. Motif B is a loop, linking a β chain from the finger subdomain to the N-terminal helix emerging from the palm. This union has a preserved glycine that is essential for the function of the polymerase. Another of its functions is the union of the RNA template and the discrimination of the substrate [13]. In these two motifs, highly conserved residues exist through the CoV family, D623 located in motif A, T680 and N691 in motif B [14].

The nsp12 protein forms a complex for RNA replication, with the nsp7 and nsp8 cofactors [8, 14, 15, 16, 17]. The nsp8 cofactor acts as a primase RNA, producing oligonucleotides for subsequent extension by nsp12 [12, 18]. The nsp7-nsp8 complex is involved in the binding role of nucleic acids [19]. Various studies have shown that the nsp12 protein has minimal activity on its own; however, when bound to nsp7 and nps8 proteins, it increases its nucleotide polymerization activity [15, 20]. That is why this complex is considered as the minimum structure for these processes and as a potential pharmacological target [10, 14, 21].

The Pathogen Box [22], is an open database of small molecules, frequently used for treat-

ment against malaria [23]. However, in recent years, it has been used in the search for drugs for the treatment of other diseases such as trypanosomiasis, cryptosporidiosis, toxoplasmosis, tuberculosis, filariasis, schistosomiasis, dengue and trichuriasis [24, 25], giardiasis [26], candidiasis [27], among others.

In this context, the objective of the present study is to search for inhibitors of the nsp12 protein, whose structure has been previously crystallized [8]. For this, different computational biology tools were used to predict interactions between drugs and proteins. The principle of the study was based on the discovery of new uses for existing drugs, also known as drug repositioning; being a novel alternative to *de novo* discovery and development, which is considered as a series of long, complicated and expensive processes [28, 29, 30, 31].

2. Material and methods

2.1. Drug database

To identify potential SARS-CoV-2 RNA polymerase inhibitors, a database of 20,000 chemical compounds from The Pathogen Box was used. These compounds were filtered, under the criterion of selecting those that do not present any violation of Lipinski's Rule of Five [32]. Next, each chemical compound was prepared for virtual screening, for this reason, they were transformed from SMILE files to SDF, PDB and PDBQT, using OpenBabel v.2.4.0 [33], which polar hydrogens and a protonation at pH 7.4 were added, following the methodology described in Ref. [34]. Subsequently, to achieve the optimization of the geometry of the structures of these compounds, these were minimized using the force field MMFF94 [35]. This procedure used an *in-house* Python script that automatically executes each step.

2.2. Receptor preparation

The crystal structure of the SARS-CoV-2 RNA polymerase enzyme was obtained in PDB (PDB ID: 6M71, 2.90 Å resolution) [8] format from the Protein Data Bank platform [36]. Subsequently, this structure was repaired with the addition of missing atoms, modification

of non-standard residues and elimination of heteroatoms, using the CHARMM-GUI software [37], with which two disulfide bridges were added (C301 -C306 and C487-C645). In addition to this, the protein structure was converted to PDBQT format, with the previous addition of Gasteiger-Marsili charges and polar hydrogens, using the MGLTools v.1.5.7 [38], to be afterward used in virtual screening and molecular docking tests.

2.3. Virtual screening

To search for inhibitors, the compound database was used in the virtual screening tests, using Autodock-Vina v1.1.2 software [39]. To perform this, the grid-box was built considering three amino acids important for receptor catalysis (T680, N691 and D623), which are highly conserved among the coronavirus species [14]. This procedure was performed considering the following search parameters: exhaustiveness of 20, spacing of 1.00 Å and grid size of 50 x , 50 y and 50 z . In the same way, another *in-house* Python script was used to automatically run virtual screening and select the Top1000 compounds with the best binding energy from the obtained results.

2.4. Molecular docking

Using the Top1000 compounds from the result of the virtual screening, a molecular docking was done to validate the results and carry out a new screening. For this procedure, a new grid-box was prepared on the previously mentioned amino acids considering a spacing of 0.375 Å. The Autodock GPU software [38], which implements the algorithm from [40], was used in this procedure. Molecular docking parameters were as follows: population size 350, number of evaluations 2,500,000, number of generations 27,000, mutation ratio 0.02, crossover ratio 0.8, and run 100. In each trial, selection of the best pose was performed based on the 2D score (2DS), described by Blanco (2019), which normalizes two proper variables for each pose using the following arithmetic combination $2DS = N_{pop} - NMBE$, where N_{pop} is the population of the cluster where the pose of the ligand was classified, and $NMBE$ is the average coupling energy. Next, each pose was extracted with the PyMol v.2.0 [41] and converted to PDB format using a MGLTools Python script. This procedure was performed in the same way using an *in-house* Python script, to automate the procedures.

2.5. Data analysis

The electrostatic potential of the nsp12 protein was calculated with the APBS module [42] employing the PyMol v.2.0 [41]. The resulting data from virtual screening was handled using a Python and C *in-house* scripts. The first one was used to select the top of the best ligands based on the binding energy and the second one, to convert the binding energy to the dissociation constant, with equation $\Delta G = RT\ln(k_d)$ (R is the gas constant and T is the temperature, taken equal to 300 K). The generated poses that showed the best interaction energies (ΔG) were analyzed with PLIP [43] to determine the type and number of interactions between drugs and the receptor.

3. Results

The structure of the SARS-CoV-2 RNA polymerase complex, linked to its cofactors, has been represented three-dimensionally in figure 1A. The replication machinery of this virus has the nsp12 protein as its catalytic center, while the cofactors (an nsp8 protein and an nsp7-nsp8 heterodimer), linked to the center, allow an increase in polymerase activity. Furthermore, the calculation of the surface of the electrostatic potential indicates that the nucleus of the structure is mainly composed of positively charged amino acids, unlike the periphery, where neutral and negative charged amino acids are observed (figure 1B).

The structure of nsp12, presents in its central region the active site (figure 2A), which is made up of seven motifs (figure 2B), which gives the machinery polymerase activity. Two of these motifs, A and B, have three highly conserved amino acids (D623, T680 and N691) among the coronavirus family (figure 2C).

The results of the virtual screening were classified according to the free binding energy (ΔG) of each compound. The Top1000 compounds with the best affinity were considered for subsequent validation by molecular docking. This last technique allowed obtaining the Top10 of compounds that have a high affinity with the active site (see table 1), with values ≤ -7.29 kcal/mol. We note that nine of the best compounds interact by forming hydrogen bonds with at least one of the conserved residues. Whereas, the TCMDC-135052 and

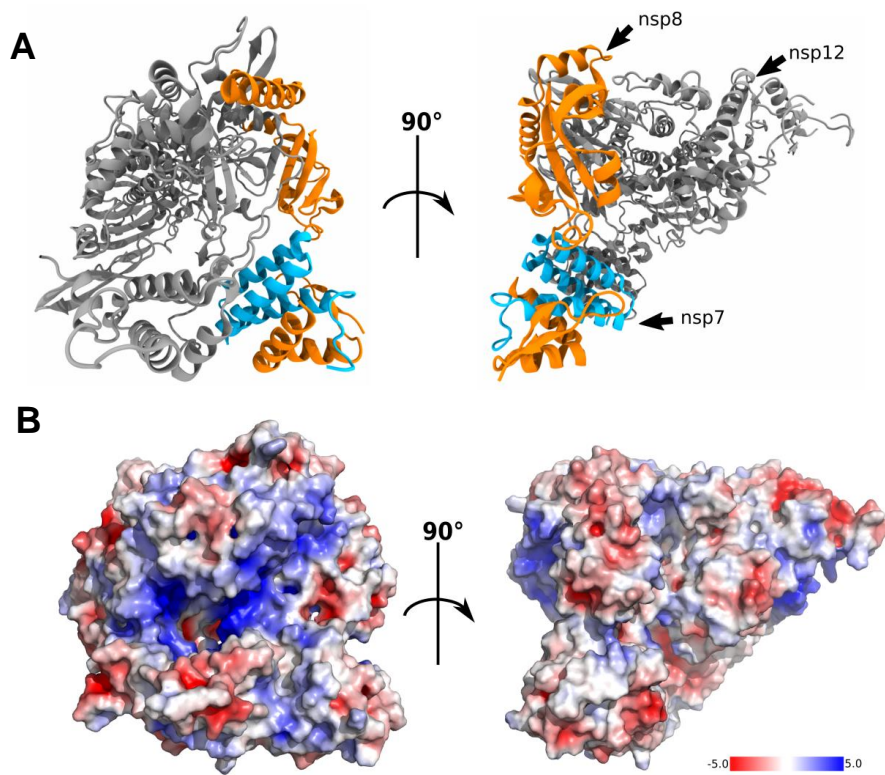


Figure 1: Frontal and lateral view of the nsp12-nsp7-nsp8 complex. A) Structure of nsp12 and cofactors nsp7 and nsp8, these are in gray, cyan and orange, respectively. B) Molecular electrostatic potential surface. Color coding: Blue: positive potential; Red: negative potential.

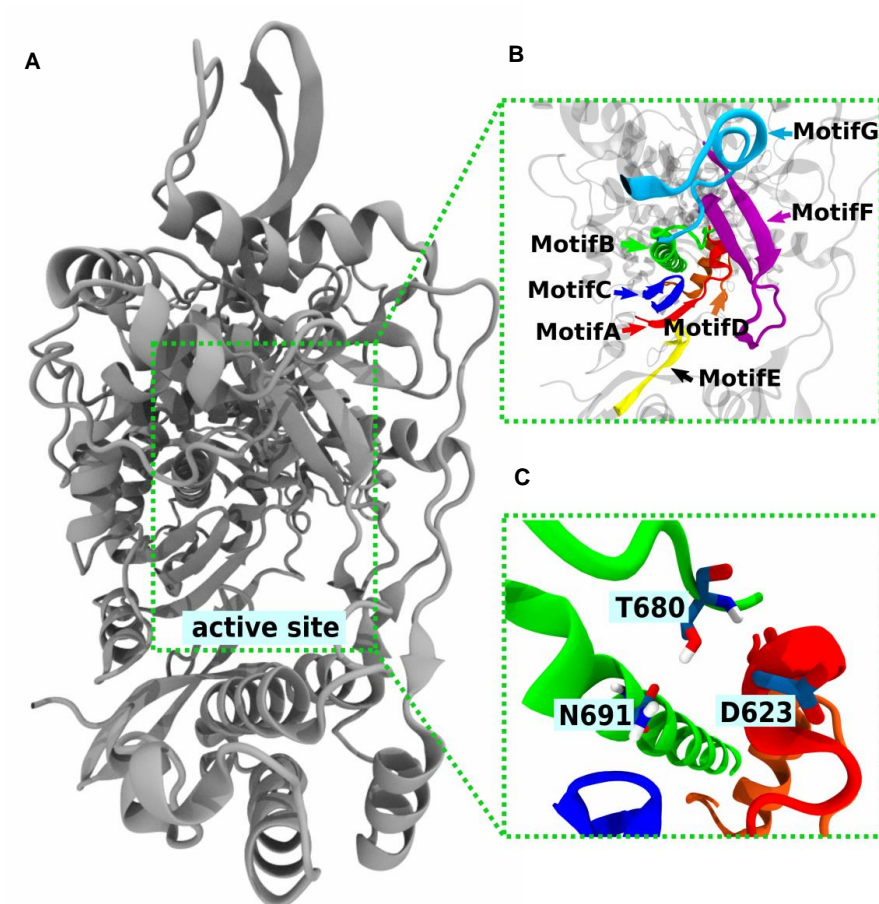


Figure 2: nsp12 protein A) The nsp12 active site is showed in gray. B) Zoom of active site shows motifs A-G. C) The zoom of motifs A and B shows conserved amino acids D623, T680 and N691. These residues are represented in stick and are their atoms hydrogen, oxygen and nitrogen are showed in white, red and blue respectively.

TCMDC-125802 compounds, which have the best binding energy (-7.53 kcal/mol), have at least one interaction with the three conserved amino acids of nsp12.

In the analysis using the PLIP software, we determined that the compounds evaluated, interact with key residues of the active site of nsp12, forming hydrogen bonds, followed by hydrophobic interactions, salt bridge interactions, aromatic and halogen bonds, where the amino acids that highlights are D623, followed by D760 and N691, which are present in most complexes. On the other hand, salt bridges were observed, mostly linked to D760 and at least one halogen bond with the Thr657 residue and the TCDMDC-124626 ligand (see table 1 and figure 3).

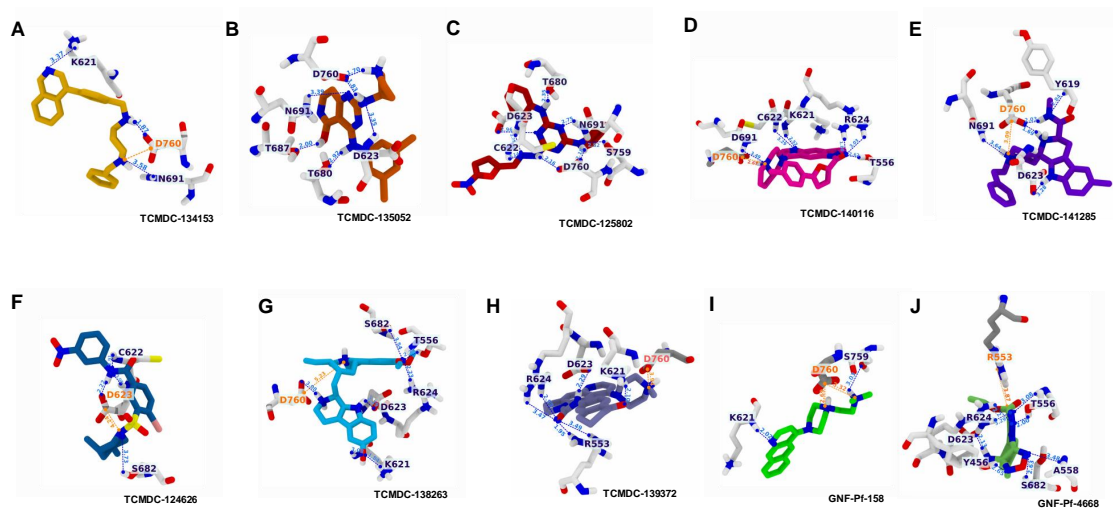


Figure 3: Molecular recognition in nsp12. Here we show in sticks, the Top10 ligands (table 1) after validation by molecular docking. Each ligand are in different colors and receptor residues are in white. The hydrogen, nitrogen and oxygen atoms are in white, blue and red, respectively.

Table 1: Ligands interaction after validation by molecular docking with SARS-CoV-2 RNA polymerase.

Ligand	ΔG	k_d	Interactions												
			Number	H-Bond Residue	Distance	Number	Aromatic Residue	Number	Hydrophobic Residue	Number	Salt bridges Residue	Distance	Number	Residue	Distance
TCMDC-134153	-7.53	28.86	3	K621(NH-N)	3.37	1	R553	2	Y455, D623	1	D760	2.79			
				N691(NH-N)	3.58										
				D760(O-HN)	1.87										
TCMDC-135052	-7.53	28.86	6	D623(NH-N)	3.35										
				T680(O-HN)	2.92										
				T687(OH-O)	2.00										
				N691(NH-N)	3.39										
				D760(O-HN)	1.83										
				D760(O-HN)	1.70										
TCMDC-125802	-7.51	29.86	8	C622(NH-N)	3.44			3	K621, T687, A688						
				D623(O-HN)	2.96										
				D623(NH-N)	2.75										
				T680(NH-O)	2.35										
				N691(NH-N)	2.75										
				S759(O-HN)	2.52										
				D760(NH-O)	2.18										
				D760(NH-O)	2.16										
TCMDC-140116	-7.51	29.86	6	T556(OH-O)	2.62					1	D760	2.67			
				K621(NH-N)	2.52										
				C622(NH-N)	3.29										
				R624(NH-O)	1.86										
				R624(NH-O)	3.03										
				N691(NH-N)	3.49										
TCMDC-141285	-7.47	31.95	5	Y619(NH-N)	2.84			1	D623	1	D760	3.09			
				D623(O-HN)	3.28										
				N691(NH-N)	3.64										
				D760(O-HN)	2.03										
				D760(O-HN)	1.69										
TCMDC-124626	-7.41	35.37	4	C622(NH-N)	2.91			1	D623	1	D623	4.24	1	T687	3.53
				D623(NH-N)	3.08										
				D623(O-HN)	2.22										
				S682(OH-N)	3.73										
TCMDC-138263	-7.38	37.21	7	T556(OH-O)	2.28					1	D760	5.23			
				K621(NH-O)	3.08										
				K621(NH-O)	1.84										

Continued on next page

Table 1 – continued from previous page

Ligand	ΔG	k_d	Interactions												
			H-Bond			Aromatic		Hydrophobic		Salt bridges			Halogen bonds		
			Number	Residue	Distance	Number	Residue	Number	Residue	Number	Residue	Distance	Number	Residue	Distance
				D623(O-HN)	2.29										
				R624(NH-O)	2.23										
				S682(OH-O)	3.54										
				D760(O-HN)	3.08										
TCMDC-139372	-7.38	37.21	6	R553(NH-O)	1.95	1	R553	4	Y455, R553, K621, D623	1	D760	3.66			
				R553(NH-O)	3.49										
				K621(NH-O)	2.10										
				D623(O-HN)	2.29										
				R624(NH-O)	2.93										
				R624(NH-O)	3.47										
GNF-Pf-158	-7.32	41.19	2	K621(NH-N)	2.02	1	R553	2	Y455, K621	2	D760	3.32			
				S759(O-HN)	3.03						D760	2.69			
GNF-Pf-4668	-7.29	43.34	8	Y456(OH-N)	2.65			2	D623, T687	1	R553	3.87			
				T556(OH-N)	2.00										
				T556(NH-N)	3.08										
				A558(NH-N)	3.48										
				D623(O-HN)	2.13										
				R624(NH-N)	3.29										
				R624(NH-N)	2.20										
				S682(NH-N)	2.63										

Note: ΔG is in units of kcal/mol. k_d is in units of nM and distances are in Å.

One of the two best compounds, TCMDC-135052, interacts with residue D760 by forming two hydrogen bonds with two of its amino groups (primary and secondary) at distances of 1.70 Å, 1.83 Å, respectively. Likewise, the formation of a hydrogen bonds with residues D623, T680, T687, N691 was observed through their secondary amino, hydroxyl, tertiary amino and secondary amino groups at distances of 2.92 Å, 2.00 Å, 3.39 Å and 3.35 Å respectively (see table 1 and figure 3).

On the other hand, compound TCMDC-125802 interacts with residues D623 and D760 forming two hydrogen bonds with two tertiary amino groups at distances of 2.96 Å and 3.75 Å; and with secondary amino groups at distances of 2.16 Å and 2.18 Å. Likewise, the formation of a hydrogen bonds with residues C622, T680, N691 and S759 was observed through tertiary and secondary amino groups, at distances of 3.14 Å, 2.35 Å, 2.75 Å and 2.52 Å (see table 1 and figure 3).

Regarding positive controls, only the presence of hydrogen bonds interactions and hydrophobic interactions were observed, where the compound Sofosbuvir interacts with the three residues of the core of the active site (D623, T680 and N691) (table 2 and figure 4). On the other hand, when evaluating the binding-free energies of the compounds taken as positive controls, it was found that Remdesivir, Beclabuvir and Sofosbuvir have values of -3.84 , -3.40 and -2.41 kcal/mol respectively, with Remdesivir being the compound with the best interaction between positive controls (table 2).

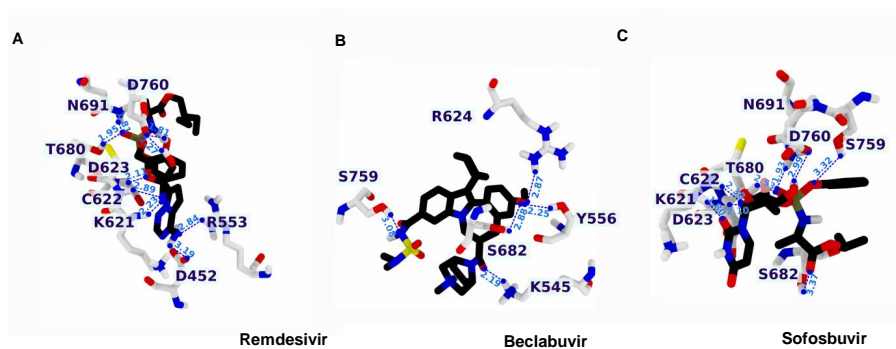


Figure 4: Molecular recognition in nsp12 with positive controls. Here we show in sticks, the validation by molecular docking. Positive controls are in black and their receptor residues are in white. The hydrogen, nitrogen and oxygen atoms are in white, blue and red, respectively.

Table 2: Positive controls interaction after validation by molecular docking with SARS-CoV-2 RNA polymerase.

Ligand	ΔG	Interactions				
		Number	H-Bond Residue	Distance	Number	Hydrophobic Residue
Remdesivir	-3.84	9	D452(O-HN)	3.19	2	D593, Y657
			R553(NH-N)	2.84		
			K621(NH-N)	2.23		
			C622(NH-N)	2.89		
			D623(NH-O)	2.11		
			T680(OH-O)	1.95		
			N691(NH-O)	1.81		
			D760(O-HO)	1.81		
			D760(O-HO)	2.27		
Beclabuvir	-3.40	5	K545(NH-O)	2.19	1	D593
			Y556(OH-O)	2.25		
			R624(NH-O)	2.87		
			S682(OH-O)	2.88		
			S759(OH-N)	3.08		
Sofosbuvir	-2.41	8	K621(NH-O)	1.80	3	D593, Y657, A658
			C622(NH-O)	2.30		
			D623(NH-O)	2.10		
			T680(OH-O)	2.53		
			S682(OH-O)	3.37		
			N691(NH-O)	1.93		
			S759(OH-O)	3.32		
			D760(O-HO)	2.99		

Note: ΔG is in units of kcal/mol. k_d is in units of nM and distances are in Å.

4. Discussion

The absence of an efficient treatment for COVID-19 represents a public health problem, since this disease continues to cause thousands of deaths worldwide [3, 44]. In this context, the methodologies applied in the present study were aimed at conducting a search and repositioning of drugs, through the use of tools in computational biology. This is a key to reducing the experimental testing periods, and to improve the specificity of the active ingredients, since they resemble the behavior that the drug would have in the body [45].

In the molecular docking analysis, we identified compounds that interact with conserved amino acids from the nsp12 active site (D623, T680, and N691) [14]. In a preliminary study, it was suggested to consider two of these residues (D623 and N691), for the search for new nsp12 inhibitors [46]. Interestingly, compounds TCMDC-135052 and TCMDC-125802 form

hydrogen bonds with these residues, suggesting a high potential to inhibit the active site of this protein. Furthermore, these compounds may have similar activity in nsp12 homologous proteins.

It is important to note that the D623 residue is highly conserved in the replication machinery of nidovirals [47]. Structural and thermodynamic evidence indicates that this residue has a discrimination function between rNTPs and dNTPs [48], in addition, it participates in the recognition of sugar from rNTPs [49]. In this sense, compounds that interact with this residue (table 1) can compromise RNA replication activity. Similarly, we observed that the compound Remdesivir (an FDA approved drug), used as a docking control, interacts with D623 (table 2), as reported in other studies [8, 49].

On the other hand, the T680 residue of the nsp12 protein has a structural function that helps the coupling of the NTPs that enter the active site of the replication machinery [49]. Hydrogen bonds between compounds TCMDC-135052 and TCMDC-125802 and this residue were observed at distances of 2.92 Å and 2.35 Å, respectively. Likewise, the Remdesivir and Sofosbuvir compounds form hydrogen bonds with T680, however, the interaction with Sofosbuvir occurs at a greater distance from the TCMDC-125802 compound (2.53 Å), which suggests a greater probability of hydrogen bond formation between this residue and the compound.

In the case of the amino acid N691, it is involved in the recognition of the 2'-OH of the incoming nucleotides in the replication machine [49, 50]. In our results, the first five compounds in our Top10 form hydrogen bonds with this residue, mainly mediated by amino groups (table 1). Similarly, other drugs such as Galidesivir and Penciclovir inhibit the activity through this residue [10]. However, they present functional groups that differ from our results. Consequently, the functional groups of our compounds could allow a better interaction with nsp12, according to the energies obtained by the molecular docking (table 1).

As it is well known, it is important to consider specific functional groups in rational drugs design [51]. Based on this, the greater presence of amino groups in compound-target interactions suggests potential inhibitory compounds. It is also supported by the functional groups observed in Sofosbuvir, which was used as a control in this work. Therefore, these

features can be exploited in structure-based drug design using a pharmacophore model.

Finally, the compounds with the best binding energy need to be evaluated using molecular dynamics to determine the stability of the complexes, mediated by the main residues that confer catalytic activity on nsp12. Furthermore, experimental studies of these compounds will validate their effectiveness in the treatment against SARS-CoV-2.

5. Conclusion

In the present work, it was proposed that the chemical compounds TCMDC-135052 and TCMDC-125802, possess a high inhibition potential of nsp12 of SARS-CoV-2, a protein important in the replication machinery of this virus. These compounds need to be evaluated using molecular dynamics, in order to determine the stability of the complexes, mediated by the main conserved residues that confer catalytic activity to nsp12. Likewise, experimental studies will validate the effectiveness in the treatment against SARS-CoV-2.

6. Acknowledgments

Part of the results presented here were developed with the help of CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho em São Paulo) grant UNICAMP/FINEP-MCT and CENAPAD-UFC (Centro Nacional de Processamento de Alto Desempenho, at Universidade Federal do Ceará). We are grateful to Prof. João Marcelo Pereira Alves from Department of Parasitology of Institute of Biomedical Sciences of USP for his computational support in the virtual screening simulations.

References

- [1] M. E. J. Woolhouse, K. Adair, L. Brierley, RNA viruses: a case study of the biology of emerging infectious diseases, *Microbiol. Spectr.* 1 (2014) 1 (jan 2014). doi:10.1128/microbiolspec.0H-0001-2012.
- [2] R. Carrasco-Hernandez, R. Jácome, Y. L. Vidal, S. P. de León, Are RNA viruses candidate agents for the next global pandemic? a review, *ILAR Journal* 58 (3) (2017) 343–358 (2017). doi:10.1093/ilar/ilx026.
- [3] Corona virus disease (COVID-19) Situation Report-105, Tech. rep., World Health Organization (2020). URL https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200504-covid-19-sitrep-105.pdf?sfvrsn=4cdda8af_2
- [4] A. R. Fehr, S. Perlman, Coronaviruses: An overview of their replication and pathogenesis, in: *Coronaviruses*, Springer New York, 2015, pp. 1–23 (2015). doi:10.1007/978-1-4939-2438-7_1.
- [5] V. C. C. Cheng, S. K. P. Lau, P. C. Y. Woo, K. Y. Yuen, Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection, *Clin. Microbiol. Rev.* 20 (4) (2007) 660–694 (oct 2007). doi:10.1128/cmr.00023-07.
- [6] J. F. W. Chan, S. K. P. Lau, K. K. W. To, V. C. C. Cheng, P. C. Y. Woo, K.-Y. Yuen, Middle east respiratory syndrome coronavirus: Another zoonotic betacoronavirus causing SARS-like disease, *Clin. Microbiol. Rev.* 28 (2) (2015) 465–522 (mar 2015). doi:10.1128/cmr.00102-14.
- [7] K. A. White, L. Enjuanes, B. Berkhout, RNA virus replication, transcription and recombination, *RNA Biology* 8 (2) (2011) 182–183 (mar 2011). doi:10.4161/rna.8.2.15663.
- [8] Y. Gao, L. Yan, Y. Huang, F. Liu, Y. Zhao, L. Cao, T. Wang, Q. Sun, Z. Ming, L. Zhang, J. Ge, L. Zheng, Y. Zhang, H. Wang, Y. Zhu, C. Zhu, T. Hu, T. Hua, B. Zhang, X. Yang, J. Li, H. Yang, Z. Liu, W. Xu, L. W. Guddat, Q. Wang, Z. Lou, Z. Rao, Structure of the RNA-dependent RNA polymerase from COVID-19 virus, *Science* 368 (6492) (2020) 779–782 (apr 2020). doi:10.1126/science.abb7498.
- [9] S. K. Lal (Ed.), *Molecular biology of the SARS-coronavirus*, Springer Berlin Heidelberg, 2010 (2010). doi:10.1007/978-3-642-03683-5.
- [10] W.-F. Zhang, P. Stephen, J.-F. Thériault, R. Wang, S.-X. Lin, Novel coronavirus polymerase and nucleotidyl-transferase structures: Potential to target new outbreaks, *J. Phys. Chem. Lett.* 11 (11) (2020) 4430–4435 (may 2020). doi:10.1021/acs.jpcllett.0c00571.
- [11] S. Venkataraman, B. Prasad, R. Selvarajan, RNA dependent RNA polymerases: insights from structure, function and evolution, *Viruses* 10 (2) (2018) 76 (feb 2018). doi:10.3390/v10020076.
- [12] A. J. W. te Velhuis, Common and unique features of viral RNA-dependent polymerases, *Cell. Mol. Life Sci.* 71 (22) (2014) 4403–4420 (aug 2014). doi:10.1007/s00018-014-1695-z.
- [13] D. Garriga, C. Ferrer-Orta, J. Querol-Audí, B. Oliva, N. Verdaguer, Role of motif B loop in allosteric regulation of RNA-dependent RNA polymerization activity, *J. Mol. Biol.* 425 (13) (2013) 2279–2287

- (jul 2013). doi:10.1016/j.jmb.2013.03.034.
- [14] R. N. Kirchdoerfer, A. B. Ward, Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors, *Nat. Commun.* 10 (1) (2019) 2342 (may 2019). doi:10.1038/s41467-019-10280-3.
- [15] L. Subissi, C. C. Posthuma, A. Collet, J. C. Zevenhoven-Dobbe, A. E. Gorbalenya, E. Decroly, E. J. Snijder, B. Canard, I. Imbert, One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities, *PNAS* 111 (37) (2014) E3900–E3909 (sep 2014). doi:10.1073/pnas.1323705111.
- [16] Q. Wang, J. Wu, H. Wang, Y. Gao, Q. Liu, A. Mu, W. Ji, L. Yan, Y. Zhu, C. Zhu, X. Fang, X. Yang, Y. Huang, H. Gao, F. Liu, J. Ge, Q. Sun, X. Yang, W. Xu, Z. Liu, H. Yang, Z. Lou, B. Jiang, L. W. Guddat, P. Gong, Z. Rao, Structural basis for RNA replication by the SARS-CoV-2 polymerase, *Cell* (2020) In press (may 2020). doi:10.1016/j.cell.2020.05.034.
- [17] W. Yin, C. Mao, X. Luan, D.-D. Shen, Q. Shen, H. Su, X. Wang, F. Zhou, W. Zhao, M. Gao, S. Chang, Y.-C. Xie, G. Tian, H.-W. Jiang, S.-C. Tao, J. Shen, Y. Jiang, H. Jiang, Y. Xu, S. Zhang, Y. Zhang, H. E. Xu, Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir, *Science* (2020) eabc1560 (may 2020). doi:10.1126/science.abc1560.
- [18] I. Imbert, J.-C. Guillemot, J.-M. Bourhis, C. Bussetta, B. Coutard, M.-P. Egloff, F. Ferron, A. E. Gorbalenya, B. Canard, A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus, *The EMBO Journal* 25 (20) (2006) 4933–4942 (oct 2006). doi:10.1038/sj.emboj.7601368.
- [19] Y. Zhai, F. Sun, X. Li, H. Pang, X. Xu, M. Bartlam, Z. Rao, Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer, *Nat. Struct. Mol. Biol.* 12 (11) (2005) 980–986 (oct 2005). doi:10.1038/nsmb999.
- [20] D.-G. Ahn, J.-K. Choi, D. R. Taylor, J.-W. Oh, Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates, *Archives of Virology* 157 (11) (2012) 2095–2104 (jul 2012). doi:10.1007/s00705-012-1404-x.
- [21] Q. Peng, R. Peng, B. Yuan, J. Zhao, M. Wang, X. Wang, Q. Wang, Y. Sun, Z. Fan, J. Qi, G. F. Gao, Y. Shi, Structural and biochemical characterization of the nsp12-nsp7-nsp8 core polymerase complex from SARS-CoV-2, *Cell Reports* 31 (11) (2020) 107774 (jun 2020). doi:10.1016/j.celrep.2020.107774.
- [22] W. C. V. Voorhis, J. H. Adams, R. Adelfio, V. Ahyong, M. H. Akabas, P. Alano, A. Alday, Y. A. Resto, A. Alsibae, A. Alzualde, K. T. Andrews, S. V. Avery, V. M. Avery, L. Ayong, M. Baker, S. Baker, C. B. Mamoun, S. Bhatia, Q. Bickle, L. Bounaadja, T. Bowling, J. Bosch, L. E. Boucher, F. F. Boyom, J. Brea, M. Brennan, A. Burton, C. R. Caffrey, G. Camarda, M. Carrasquilla, D. Carter, M. B. Cassera, K. C.-C. Cheng, W. Chindaudomsate, A. Chubb, B. L. Colon, D. D. Colón-López, Y. Corbett, G. J. Crowther, N. Cowan, S. D’Alessandro, N. L. Dang, M. Delves, J. L. DeRisi, A. Y. Du, S. Duffy, S. A.

- E.-S. El-Sayed, M. T. Ferdig, J. A. F. Robledo, D. A. Fidock, I. Florent, P. V. T. Fokou, A. Galstian, F. J. Gamo, S. Gokool, B. Gold, T. Golub, G. M. Goldgof, R. Guha, W. A. Guiguemde, N. Gural, R. K. Guy, M. A. E. Hansen, K. K. Hanson, A. Hemphill, R. H. van Huijsduijnen, T. Horii, P. Horrocks, T. B. Hughes, C. Huston, I. Igarashi, K. Ingram-Sieber, M. A. Itoe, A. Jadhav, A. N. Jensen, L. T. Jensen, R. H. Y. Jiang, A. Kaiser, J. Keiser, T. Ketas, S. Kicka, S. Kim, K. Kirk, V. P. Kumar, D. E. Kyle, M. J. Lafuente, S. Landfear, N. Lee, S. Lee, A. M. Lehane, F. Li, D. Little, L. Liu, M. Llinás, M. I. Loza, A. Lubar, L. Lucantoni, I. Lucet, L. Maes, D. Mancama, N. R. Mansour, S. March, S. McGowan, I. M. Vera, S. Meister, L. Mercer, J. Mestres, A. N. Mfopa, R. N. Misra, S. Moon, J. P. Moore, F. M. R. da Costa, J. Müller, A. Muriana, S. N. Hewitt, B. Nare, C. Nathan, N. Narraido, S. Nawaratna, K. K. Ojo, D. Ortiz, G. Panic, G. Papadatos, S. Parapini, K. Patra, N. Pham, S. Prats, D. M. Plouffe, S.-A. Poulsen, A. Pradhan, C. Quevedo, R. J. Quinn, C. A. Rice, M. A. Rizk, A. Ruecker, R. S. Onge, R. S. Ferreira, J. Samra, N. G. Robinett, U. Schlecht, M. Schmitt, F. S. Villela, F. Silvestrini, R. Sinden, D. A. Smith, T. Soldati, A. Spitzmüller, S. M. Stamm, D. J. Sullivan, W. Sullivan, S. Suresh, B. M. Suzuki, Y. Suzuki, S. J. Swamidass, D. Taramelli, L. R. Y. Tchokouaha, A. Theron, D. Thomas, K. F. Tonissen, S. Townson, A. K. Tripathi, V. Trofimov, K. O. Udenze, I. Ullah, C. Vallieres, E. Vigil, J. M. Vinetz, P. V. Vinh, H. Vu, N. aki Watanabe, K. Weatherby, P. M. White, A. F. Wilks, E. A. Winzeler, E. Wojcik, M. Wree, W. Wu, N. Yokoyama, P. H. A. Zollo, N. Abla, B. Blasco, J. Burrows, B. Laleu, D. Leroy, T. Spangenberg, T. Wells, P. A. Willis, Open source drug discovery with the malaria box compound collection for neglected diseases and beyond, *PLOS Pathogens* 12 (7) (2016) e1005763 (jul 2016). doi:10.1371/journal.ppat.1005763.
- [23] S. Duffy, M. L. Sykes, A. J. Jones, T. B. Shelper, M. Simpson, R. Lang, S.-A. Poulsen, B. E. Sleebs, V. M. Avery, Screening the medicines for malaria venture pathogen box across multiple pathogens reclassifies starting points for open-source drug discovery, *Antimicrob. Agents Chemother.* 61 (9) (2017) e00379 (jul 2017). doi:10.1128/AAC.00379-17.
- [24] J. Spalenka, S. Escotte-Binet, A. Bakiri, J. Hubert, J.-H. Renault, F. Velard, S. Duchateau, D. Aubert, A. Huguenin, I. Villena, Discovery of new inhibitors of toxoplasma gondii via the pathogen box, *Antimicrob. Agents Chemother.* 62 (2) (2017) e01640 (nov 2017). doi:10.1128/AAC.01640-17.
- [25] C. G. L. Veale, Unpacking the Pathogen Box—an open source tool for fighting neglected tropical disease, *ChemMedChem* 14 (4) (2019) 386–453 (jan 2019). doi:10.1002/cmdc.201800755.
- [26] K. M. Hennessey, I. C. Rogiers, H.-W. Shih, M. A. Hulverson, R. Choi, M. C. McCloskey, G. R. Whitman, L. K. Barrett, E. A. Merritt, A. R. Paredez, K. K. Ojo, Screening of the pathogen box for inhibitors with dual efficacy against *giardia lamblia* and *cryptosporidium parvum*, *PLOS Neglected Tropical Diseases* 12 (8) (2018) e0006673 (aug 2018). doi:10.1371/journal.pntd.0006673.
- [27] T. Vila, J. L. Lopez-Ribot, Screening the pathogen box for identification of candida albicans biofilm in-

- hibitors, *Antimicrob. Agents Chemother.* 61 (1) (2016) e02006 (oct 2016). doi:10.1128/AAC.02006-16.
- [28] R. A. Hodos, B. A. Kidd, K. Shameer, B. P. Readhead, J. T. Dudley, In *silicomethods for drug repurposing and pharmacology*, Wiley Interdiscip. Rev. Syst. Biol. Med. 8 (3) (2016) 186–210 (apr 2016). doi:10.1002/wsbm.1337.
- [29] K. Park, A review of computational drug repurposing, *Transl. Clin. Pharmacol.* 27 (2) (2019) 59 (2019). doi:10.12793/tcp.2019.27.2.59.
- [30] M. García-Serradilla, C. Risco, B. Pacheco, Drug repurposing for new, efficient, broad spectrum antivirals, *Virus Res.* 264 (2019) 22–31 (apr 2019). doi:10.1016/j.virusres.2019.02.011.
- [31] A. Pandey, A. N. Nikam, A. B. Shreya, S. P. Mutalik, D. Gopalan, S. Kulkarni, B. S. Padya, G. Fernandes, S. Mutalik, R. Prassl, Potential therapeutic targets for combating SARS-CoV-2: Drug repurposing, clinical trials and recent advancements, *Life Sci.* 256 (2020) 117883 (sep 2020). doi:10.1016/j.lfs.2020.117883.
- [32] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Delivery Rev.* 23 (1-3) (1997) 3–25 (jan 1997). doi:10.1016/s0169-409x(96)00423-1.
- [33] N. M. OBoyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, Open babel: An open chemical toolbox, *J. Cheminf.* 3 (2011) 33 (oct 2011). doi:10.1186/1758-2946-3-33.
- [34] G. Ropón-Palacios, M. E. Chenet-Zuta, G. E. Olivos-Ramirez, K. Otazu, J. Acurio-Saavedra, I. Camps, Potential novel inhibitors against emerging zoonotic pathogen nipah virus: a virtual screening and molecular dynamics approach, *J. Biomol. Struct. Dyn.* (2019) 1–10 (aug 2019). doi:10.1080/07391102.2019.1655480.
- [35] T. A. Halgren, Merck molecular force field. i. basis, form, scope, parameterization, and performance of MMFF94, *J. Comput. Chem.* 17 (5-6) (1996) 490–519 (apr 1996). doi:10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.
- [36] H. M. Berman, The protein data bank, *Nucleic Acids Res.* 28 (1) (2000) 235–242 (jan 2000). doi:10.1093/nar/28.1.235.
- [37] S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A web-based graphical user interface for CHARMM, *J. Comput. Chem.* 29 (11) (2008) 1859–1865 (mar 2008). doi:10.1002/jcc.20945.
- [38] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (16) (2009) 2785–2791 (dec 2009). doi:10.1002/jcc.21256.
- [39] O. Trott, A. J. Olson, AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2010) 455–461 (2010). doi:10.1002/jcc.21334.

- [40] F. J. Solis, R. J.-B. Wets, Minimization by random search techniques, *Math. Oper. Res.* 6 (1) (1981) 19–30 (feb 1981). doi:10.1287/moor.6.1.19.
- [41] W. L. DeLano, Pymol: An open-source molecular graphics tool, software, *CCP4 Newsletter on protein crystallography* (2002).
- [42] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, *PNAS* 98 (18) (2001) 10037–10041 (aug 2001). doi:10.1073/pnas.181342398.
- [43] S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme, M. Schroeder, PLIP: fully automated protein–ligand interaction profiler, *Nucleic Acids Res.* 43 (W1) (2015) W443–W447 (apr 2015). doi:10.1093/nar/gkv315.
- [44] D. S. Jones, History in a crisis — lessons for covid-19, *N. Engl. J. Med.* 382 (18) (2020) 1681–1683 (apr 2020). doi:10.1056/NEJMp2004361.
- [45] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, *Acta Pharmaceutica Sinica B* 10 (5) (2020) 766–788 (may 2020). doi:10.1016/j.apsb.2020.02.008.
- [46] J. Ziebuhr, Molecular biology of severe acute respiratory syndrome coronavirus, *Curr. Opin. Microbiol.* 7 (4) (2004) 412–419 (aug 2004). doi:10.1016/j.mib.2004.06.007.
- [47] C. C. Posthuma, A. J. te Velthuis, E. J. Snijder, Nidovirus RNA polymerases: Complex enzymes handling exceptional RNA genomes, *Virus Res.* 234 (2017) 58–73 (apr 2017). doi:10.1016/j.virusres.2017.01.023.
- [48] R. Jácome, A. Becerra, S. P. de León, A. Lazcano, Structural analysis of monomeric RNA-dependent polymerases: Evolutionary and therapeutic implications, *PLOS ONE* 10 (9) (2015) e0139001 (sep 2015). doi:10.1371/journal.pone.0139001.
- [49] C. J. Gordon, E. P. Tchesnokov, E. Woolner, J. K. Perry, J. Y. Feng, D. P. Porter, M. Götte, Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency, *J. Biol. Chem.* 295 (20) (2020) 6785–6797 (apr 2020). doi:10.1074/jbc.RA120.013679.
- [50] H. S. Hillen, G. Kokic, L. Farnung, C. Dienemann, D. Tegunov, P. Cramer, Structure of replicating SARS-CoV-2 polymerase, *Nature* (2020) 1476–4687 (may 2020). doi:10.1038/s41586-020-2368-8.
- [51] Z. He, J. Zhang, X.-H. Shi, L.-L. Hu, X. Kong, Y.-D. Cai, K.-C. Chou, Predicting drug-target interaction networks based on functional groups and biological features, *PLoS ONE* 5 (3) (2010) e9603 (mar 2010). doi:10.1371/journal.pone.0009603.