



Analysis of non-structural proteins, NSPs of SARS-CoV-2 as targets for computational drug designing

Resal Raj*

Dr. Z.H (P.G) College, Agra Road, Etah, Uttar Pradesh, India

ARTICLE INFO

Keywords:

Non-structural proteins
Drugs
Docking
Pharmacophore
Polar surface
Binding pockets

ABSTRACT

Background: The ORF1ab of Severe Acute Respiratory Syndrome, SARS Corona Virus, SARS-CoV-2 genome is processed into 15 non-structural proteins, NSPs by proteases and each NSP has a specific role in the life cycle and pathogenicity of the virus. This research analyzes possible drugs for these proteins as targets in computational drug designing using already available experimental drugs from the drug bank database.

Methods: Out of 471 proteins and 8820 drugs download from Protein Data Bank, PDB and Drug Bank database respectively, 16 proteins similar to NSP 1–15 and 31 drugs as per the “Rule of three” were selected for docking. Out of 88 docking results using PyRx, 18 proteins/chains with three promising drugs, DB01977, DB07132 and DB07535 were analyzed using PyMOL for final results.

Results: NSPs 3, 5, 11, 14 and 15 were identified as targets for the drugs, DB01977, BD07132 and DB07535. Drugs, DB01977 and DB07535 bind in the same binding pockets of NSP 5 and NSP 15. Drug, DB07132 binds with more number of residues when compared with the other two drugs and this indicates that the strength of protein-drug association is more by this drug with the NSPs than other drugs. Binding pockets of NSPs for these three drugs are very close with many sharing residues in common suggesting of similarity of pharmacophore of these drugs with the target binding pockets.

Conclusion: The binding pockets of NSPs are well matched with the pharmacophore of drugs and with polar surface of drugs less than or equal to 100 Å², drugs, DB01977, DB07132 and DB07535 bind individually and effectively with NSPs 3, 5, 11, 14 and 15 of ORF1ab of SARS-CoV-2 genome to bring changes in the activity of SARS-CoV-2 which may be useful for biological and clinical considerations.

1. Introduction

There are two ORFs, ORF1a and ORF1b in the genomic RNA of Severe Acute Respiratory Syndrome, SARS-CoV-2 encoding for various non-structural proteins, NSPs at 5' terminal and few structural proteins such as envelop protein, membrane proteins etc. at the 3' terminal of the genome. The translated polypeptides of ORF 1 ab are processed into approximately 1-15 NSPs [1]. NSP 1 is used by the virus to evade the host immune system, inhibition of host gene expression [2,3] and hence, it is a target protein for vaccine development. NSP 2 is dispensable for viral replication and its function is not well clear [4]. NSPs 2 and 3 interact to form proteases that cleave ORF1ab [5]. The structure of NSP 3 shows the presence of RNA-binding domains [6], SARS Unique Domain, SUD [7] which in turn has three sub domains, N-terminal, Middle and C-terminal domains [8] and papain like protease, PL-PRO domain to achieve full activity of the protein [9]. NSP 3 and NSP 4

interact with other cofactors to induce membrane rearrangement for the mechanism of viral replication and the loss of NSP 3- NSP 4 complex eliminates viral replication [10]. NSP 5 is a cysteine like protease, 3CL-PRO which processes 11 cleavage sites between NSP 4 and 16 during replication and also has a conserved 3-domain structure and catalytic residues [11–13]. NSP 6 generates autophagosomes from the endoplasmic reticulum and is involved in autophagy [14,15]. Moreover, different NSPs have different roles in virus life cycle. For example, NSP 12 in complex with NSP7 and NSP 8 forms viral replicase machinery [16–19], NSP 9 in complex with NSP 8 is involved in RNA replication and virulence [20–22], NSP10 - NSP16 complex is essential for capping viral mRNA transcripts for efficient translation and to evade immune surveillance [23], NSP 14 in complex with its activator NSP 10 is involved in exonuclease activity [24,25], NSP13 is for RNA TPase activity [26], NSP14 is for exoribonuclease activity [27,28], NSP 11 and NSP 15 are involved in endoribonuclease activity [29–31] etc. table R1.

* Department of Computational Biology & Bioinformatics, Dr. Z.H (P.G) College, Agra Road, Etah, Uttar Pradesh, India.

E-mail address: resalraj@yahoo.com.

<https://doi.org/10.1016/j.bbrep.2020.100847>

Received 29 June 2020; Received in revised form 12 October 2020; Accepted 27 October 2020

Available online 11 December 2020

2405-5808/© 2020 The Author.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

It is found that the adaptive evolution in ORF1a contribute to host shifts or immune evasion due to selection pressure and the positive selection drives the evolution of NSPs, shift and evade the immune system [32, 33]. Although most of the NSPs are orthologous, coronavirus NSP 2 is homologous to the bacterial DNA Topoisomerase I and IV needed for (–) strand RNA synthesis which suggests that NSP 2 is considered as target for drug and vaccine development [34].

Table R1

Review table, role of NSPs of SARS-CoV-2.

S. No.	Protein	Function	Reference
1	NSP1	It is involved in host-range restriction in countering innate host antiviral response and in suppressing induction of apoptosis during early stages of infection to promote viral growth.	[2,3]
2	NSP2	Involved in disruption of intracellular host signaling during SARS-CoV infections.	[35]
3	NSP3	It is proposed to facilitate translation of the mRNA transcripts and to suppress host protein synthesis.	[11–15]
4	NSP4	Essential role is replication and the assembly of the replicative structures.	[36]
5	NSP5	Protease activity	[37]
6	NSP6	Generates autophagosomes from the endoplasmic reticulum and is involved in autophagy	[14,15]
7	NSP7	Primer-Independent RNA polymerase Activity	[38]
8	NSP8	Primase activity	[39]
9	NSP9	In complex with NSP 8, involved in RNA replication and virulence of virus.	[20–22]
10	NSP10	It is a cofactor for both the 2'O-methyltransferase activity of NSP16, and the N7-guanine-methyltransferase/exoribonuclease activities of NSP14	[40–42]
11	NSP11	Essential for replication	[43]
12	NSP12	RNA polymerase/Replicase activity	[44]
13	NSP13	Helicase and RNA TPase activity	[26]
14	NSP14	Methyl transferase and Exoribonuclease activity	[27,28],
15	NSP15	Uridylate-specific Endoribonuclease activity	[29–31]

Since the progress of development of vaccine/drug has been highlighted by many research and review articles, this effort is unique to the additionally existing/ongoing drug discoveries because of the rule of three [45] which emphasizes on fragment based drug designing. Knowing the pandemic nature of the disease and unavailability of the vaccine/drug, the docking was undertaken from the existing drugs from drug bank which were in an experimental stage and selected the drugs based on the rule of three [45]. The drugs obtained were having the polar surface area less than or equal to 100 Å² instead, using the rule of five [46], the selected drugs would be of higher molecular weight, polar surface area etc. which were already under drug discovery.

2. Materials and methods

1. Download of proteins from Protein Data Bank, PDB and Screening of the downloaded proteins for docking.

The respective sequences and the associated predicted structures of NSPs 1–15 of ORF1ab of SARS-CoV-2 were downloaded from the website, <https://zhanglab.cmb.med.umich.edu/COVID-19/>, Zhang Lab, University of Michigan. The downloaded predicted protein structure lacked the chains, however the respective similar proteins from Protein Data Bank, PDB had chains, A, B, C. etc. When we tried to make Protein Data Bank, Partial Charge (Q), & Atom Type (T)) format, pdbqt file using software, Python based virtual screening tool, PyRx, Open Babel etc. for docking, some proteins produced error messages due to structural problems, low template modelling score, TM-score and therefore, the sequences of respective NSPs were used to search for proteins from PDB. The sequence of each NSP was used to search for protein 3D structures

from PDB under search and sequence search tabs. The generated protein-IDs for each NSP were used to download the structures of every protein using download multiple data files tab under download tab of PDB. Nearly, 471 proteins were downloaded for NSPs 1–15 except for NSP 2 and 6, no matching protein structure was available. Moreover, NSPs 2 and 6 were not used in docking because their predicted structure created error message and also using them might give discrepancies in docking score as others were downloaded proteins from PDB.

Each protein was analyzed with respect to its role in disease development, number of chains, the bound ligands, etc. The downloaded protein sequences were compared with the sequences of respective NSPs using multiple sequence alignment software, BaseByBase to get sequence similarity percentage. Python based molecular visualization software PyMOL was used to compare predicted protein structure by Zhang Lab, University of Michigan and the downloaded proteins from PDB to get structural similarity with the help of its “super” command. The proteins, predicted and downloaded were superimposed in PyMOL for structural similarity score using “super” command of PyMOL and a root mean square deviation, RMSD score was obtained for each protein. This was done to know how much the downloaded protein structure had deviated from the predicted protein structure and for example if RMSD score was zero, the downloaded protein and predicted protein, both were considered to have cent percent structural similarity, no deviation of structure or no structural difference.

2. Download of drugs from Drug Bank and screening of the downloaded drug for docking.

Drugs were available for download when logged in to the drug bank database account and 8820 drug compounds were downloaded after logging in to the drug bank database account. Out of 8820 drug compounds downloaded from drug bank database, 1414 drug compounds were selected based on the Lipinski's rule of 5 [46] and these compounds had high polar surface area, molecular weight, volume etc. and therefore, finally, 31 drug compounds were selected based on the rule of three [45] using DataWarrior software. The selected drugs had very low molecular weight, volume, surface area etc. and were used for molecular docking using docking software, PyRx.

3 Molecular Docking and analysis

The selected proteins were docked with the selected drugs using PyRx with automatically generated grid parameters of PyRx and the proteins were docked separately, chain wise and the entire protein to know the binding residues involved in each case for better analysis. The docked products were analyzed in terms of type of binding residues, bond lengths and different binding pockets of proteins for each drug with the help of PyMOL.

4 Analysis of target proteins of the drugs and their Molecular Docking using PyRx

The structures and sequences of target proteins of the drugs, DB01977, BD07132 and DB07535 were downloaded from PDB using the link given by Drug Bank database, analyzed and compared with structures and sequences of NSPs. Moreover, the target protein-drug complex was compared with the NSP-drug complex in terms of binding sites, bond lengths and residues.

3. Rules and software used in the research

Lipinski's rule of 5 is useful to differentiate between drug like and non-drug like molecules and its success or failure is due to drug likeness. According to the rule, the drug should have molecular mass less than 500 Da, high lipophilicity, LogP less than 5, less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors, molar refractivity

Table 1
Sequence and structural similarities of proteins.

Predicted Non-Structural Proteins (NSPs)_ID	Type of Protein	PDB Description for proteins					Sequence Similarity (%)	Structural Similarity RMSD
		PDB_ID	Chains	Seq. Start	Seq.End	Ligand(s)		
QHD43415_3	Papain-like proteinase	6wu	A,B,C,D tetramer, Vir250, G-J	-1	324	ACE, DPP,GVE, ZN, UB4, MG	89.17	0.51
QHD43415_5	Proteinase 3CL-PRO.	4HI3	A,B dimer	0	314	Nil	95.75	0.527
QHD43415_5	Proteinase 3CL-PRO.	5C5O	A,B dimer	1	306	SDJ	95.75	0.41
QHD43415_5	Proteinase 3CL-PRO.	7BRP	A,B dimer	0	306	HU5	100	0.509
QHD43415_5	Proteinase 3CL-PRO.	7BRR	A,B dimer	0	306	K36	100	0.563
QHD43415_10	viral transcription	2FYG	A	5	132	GOL, ZN	94.53	0.406
QHD43415_11	RNA-directed RNA polymerase (RdRp)	6NUR	A-D	(-1), 1, 0	953, 198, 83	ZN	96.24	0.338
QHD43415_11	RNA-directed RNA polymerase (RdRp)	6NUS	A,B	(-1), 1	953, 198	ZN	96.24	0.49
QHD43415_14	Uridylate-specific endoribonuclease (NendoU)	2H85	A	-1	345	Nil	88.12	0.339
QHD43415_14	Uridylate-specific endoribonuclease (NendoU)	6VWW	A,B dimer	-23	347	ACY, GOL, CL, MG	100	0.478
QHD43415_14	Uridylate-specific endoribonuclease (NendoU)	6W01	A,B dimer	-23	347	CIT, EDO, PEG	100	0.505
QHD43415_14	Uridylate-specific endoribonuclease (NendoU)	6WLC	A,B dimer	-2	347	ACT, EDO, FMT, TRS,U5P, SO4	100	0.526
QHD43415_14	Uridylate-specific endoribonuclease (NendoU)	6WXC	A,B dimer	-2	347	EDO, FMT, CMU, PO4	100	0.509
QHD43415_15	2'-O-methyltransferase (2'-O-MT)	6W4H	A,B	6796, 4252	7096, 4393	ACT, BDF, SAM, SO3, ZN	100	0.341
QHD43415_15	2'-O-methyltransferase (2'-O-MT)	6W75	A,C dimer, B,D dimer	6796, 4252	7096, 4393	FMT,SAM,NA, ZN	100	0.385
QHD43415_15	2'-O-methyltransferase (2'-O-MT)	6WJT	A,C dimer, B,D dimer	6796, 4252	7096, 4393	FMT,SAH,NA, ZN	100	0.4

Table 2
Binding energies of drugs using PyRx docking software.

NSPs	Drug_ID	DB00150	DB01977	DB02441	DB03225	DB03314	DB07132	DB07535	DB08136	DB08466	DB12291
NSP 3	Binding energy in Kcal/mol of 6wu	-6.4	-8.4	-7.3	-6.3	-6.7	-8.3	-8.2	-7.6	-7.1	-7.5
NSP 5	Binding energy in Kcal/mol of 6wu	-7.3	-9.2	-6.6	-7.3	-7.2	-9	-8.2	-7.7	-7.2	-8.4
	Binding energy in Kcal/mol of 6wu	-7.1	-8.2	-6.9	-7	-6.7	-7.5	-8.6	-7.7	-7.2	-7.4
NSP 10	Binding energy in Kcal/mol of 6wu	-6.1	-7.1	-6	-6.1	-7	-7.2	-6.9	-6.5	-6.3	
NSP 11	Binding energy in Kcal/mol of 6wu	-6	-7.7	-7	-6.1	-6.3	-7.3	-7.5	-7.2	-6.2	-6.9
	Binding energy in Kcal/mol of 6wu	-6.8	-8	-6.7	-6.8	-6.6	-7.3	-7.5	-7.3	-7	-7.2
NSP 14	Binding energy in Kcal/mol of 6wu	-7	-8.8	-6.9	-7	-7.2	-8	-8.6	-7.9	-7.2	-7.8
	Binding energy in Kcal/mol of 6wu	-7.2	-8.1	-7.2	-7.6	-7.6	-7.1	-8	-7.1	-7	-8.3
	Binding energy in Kcal/mol of 6wu	-7.1	-8.1	-7.3	-7.1	-7.3	-7.8	-8.2	-7.8	-7	-8.2
	Binding energy in Kcal/mol of 6wu	-7	-7.8	-6.9	-7.1	-7.1	-7.8	-7.3	-7.2	-6.9	-7.9
	Binding energy in Kcal/mol of 6wu	-7.1	-7.8	-6.9	-7.1	-7.2	-8.2	-7.6	-7.3	-7.2	-7.7
	Binding energy in Kcal/mol of 6wu	-7.4	-8.1	-6.9	-7.3	-7.5	-7.9	-7.8	-7.2	-7.5	-7.8
	Binding energy in Kcal/mol of 6wu	-7.2	-7.9	-7	-7.2	-7.4	-8.4	-7.9	-7.4	-7.4	-7.8
	Binding energy in Kcal/mol of 6wu	-7.1	-7.9	-6.9	-7.1	-7.3	-8	-7.8	-7.3	-7.3	-7.8
	Binding energy in Kcal/mol of 6wu	-7.1	-7.9	-7	-7.1	-7.3	-8.1	-7.8	-7.2	-7.3	-7.8
NSP 15	Binding energy in Kcal/mol of 6wu	-6.7	-8.2	-6.6	-6.7	-7.1	-8	-8.3	-7.6	-6.9	-7.6
	Binding energy in Kcal/mol of 6wu	-7	-8.3	-6.6	-7	-6.9	-7.4	-8.2	-7.5	-7.1	-7.5
	Binding energy in Kcal/mol of 6wu	-6.9	-7.5	-6.6	-6.9	-7	-8.5	-7.9	-7	-6.9	-6.8

Note: The binding energy given above are for the poses with RMSD = 0.

between 40 and 130 etc. **Rule of three** explains about the fragment-based screening for lead finding strategy and the physicochemical properties of fragments obey the "rule of three". According to the rule, the drug should have lower molecular mass, <300 Da, <3 hydrogen bond donors, <3 hydrogen bond acceptors, <3 rotatable bonds or in short all in threes. **PyMOL** is an open source molecular visualization software created by Warren Lyford DeLano and commercialized by Schrödinger. PyMOL can produce high quality 3D images for structural biology. PyMOL can be obtained from <https://www.schrodinger.com/pymol>. **PyRx** is virtual screening software for computational drug designing to screen libraries of compounds against drug targets. PyRx includes docking wizard with easy-to-use user interface which makes it a valuable tool for Computer-Aided Drug Design. PyRx also includes chemical spreadsheet-like functionality and powerful visualization

engine that are essential for Rational Drug Design. PyRx is available from <http://sourceforge.net/projects/pyrx/files>. For every docking, a docking parameter file is created which tells AutoDock which docking algorithm to use and how many runs to do and it usually has the file extension, ". dpf", docking parameter file. Moreover, the parameter file has details of receptors, ligands, exhaustiveness, grid details etc. and if there are no grid parameter, default grid parameters based on binding pockets were taken for docking by PyRx. Calculation of binding score or energy or affinity is completely based on the pharmacophores of ligand and receptors and mathematical expressions of the same can be found from the articles [47,48] even though explanation of mathematical derivation is beyond the scope of this paper. BaseByBase is a whole genome pairwise and multiple alignment editor. The program highlights differences between pairs of alignments and allows the user to easily

Table 3
Binding pockets and residues with bond length.

Drug ID		DB01977					DB07132 *					DB07535				
NSP	Bonded residues and bond length of 6wu	T257 = 2.7	Y305 = 2.8	E252 = 2.4			T291 = 2,3.1	K217 = 3.1	K306 = 2.3	T259 = 2.9	T259 = 2.3	S278 = 2.3,2.5	Q250 = 3.0			
NSP	Bonded residues and bond length of 7brr, Chains A&B	K5 (B) = 2.6	V125 = 2.7	K5 (A) = 2.4			L282 = 2.6	F3 = 2.2,2.5	K5 = 2.5,3.0	R4 = 2.1,2.6	L282 = 2.3					
NSP	Bonded residues and bond length of 7brr_A	K102 = 2.5	N151 = 2.2	D295 = 2.2, 2.6	T111 = 2.3	T292 = 1.6	T111 = 2.4	D295 = 2.7	Q110 = 2.7		N151 = 2.5,2.4	D295 = 2.4,2.5	T292 = 2.6	Q110 = 2.4	T111 = 2.6	
NSP	Bonded residues and bond length of 2fyg	T111 = 2.7	V116 = 3.1				T111 = 2.2	D91 = 2.2			D91 = 3.5					
NSP	Bonded residues and bond length of 6nur	E665 = 2.1,2.4	T556 = 3.0	R624 = 3.3			Y619 = 2.4	C622 = 3.0	R624 = 3.1	A554 = 2.7	D452 = 2.5,3.3	N459 = 1.9,2.5				
NSP	Bonded residues and bond length of 6nus	Y346 = 2.2					P323 = 2.2	N459 = 1.9				N628 = 2.8	P677 = 3.6	H347 = 2.5		
NSP	Bonded residues and bond length of 2h85	W86 = 2.7	P67 = 2.8	S161 = 2.4			D272 = 2.3	S273 = 2.2	T274 = 2.7	D199 = 2.2	S197 = 2.8	E68 = 2.2,5				
	Bonded residues and bond length of 6vww_A	K71 = 3.2	D268 = 2.3				M272 = 2.5	Y279 = 2.5	ACY, GOL		T275 = 2.4	S274 = 2.7				
	Bonded residues and bond length of 6vww_B	L201 = 3.1	K90 = 2.5				D297 = 3.2	S274 = 2.7	GOL		D268 = 2.5,2.2	M272 = 3.5				
	Bonded residues and bond length of 6w01_A	K71 = 3.1	D268 = 2.7				S274 = 3.3	D273 = 3.2	K71 = 3.2	Y279 = 1.9	D268 = 2.8	T275 = 3.1	D297 = 3	Y279 = 2.3	EDO	
	Bonded residues and bond length of 6w01_B	L266 = 2.4					Y279 = 2.1	D273 = 2.8			D268 = 2.2,2.4					
	Bonded residues and bond length of 6wlc_A	K71 = 3.3					E69 = 2.6	K71 = 3	D273 = 3.4	Y279 = 2	T275 = 2.1	V295 = 2.2,5	Y279 = 3.1			
	Bonded residues and bond length of 6wlc_B	K71 = 3.2	D268 = 2.1				Y279 = 2	D273 = 2.7	S274 = 3.3	T275 = 3.3	K71 = 2.8	D268 = 2.1,2.4				
	Bonded residues and bond length of 6wxc_A	K71 = 3.1					T275 = 3.3	S274 = 3.3	D273 = 2.6	K71 = 2.8	Y279 = 2	D268 = 2.2,3				
	Bonded residues and bond length of 6wxc_B	L266 = 2.6					T273 = 3.2	S274 = 3.0	D273 = 2.5	Y279 = 2.3	K71 = 2.8	D268 = 2.3,2.3				
NSP	Bonded residues and bond length of 6w75	S7090 A = 2.6, 2.2, 2.5	S7090C = 2.7				T6908 A = 3.4	S7090 = 2.8,3.4	T6908C = 2.1		S7090 A = 2.4,2.8	S7090C = 2.3	S6907 = 2.8			
	Bonded residues and bond length of 6wjt_D	T4364 = 3.1	T4368 = 2.4				D4344 = 3.3,3.3	L4365 = 2.6,2.1	N4367 = 2.7		V4369 = 3.0					

Table 4
Selected drugs and their detailed targets proteins.

Drug ID	Target Protein ID with Ligand code	Active sties with bond length	PyRx Binding energy and bonded residues	General Function of Target Protein ^a	Specific action of Target Protein ^a	Type of Target Protein ^a
DB01977	lowe, 675	D205 = 2.8,2.9 S206 = 2.8 Q208 = 2.9 G234 = 2.9	-7.5 D205 = 2.3 S206 = 2.6	Serine-type endopeptidase activity	Specifically cleaves the zymogen plasminogen to form the active enzyme plasmin.	Urokinase-type plasminogen activator
DB07132 ^a	2pe1, 517	T222 = 3.0 D223 = 3.3 K111 = 3.0	-9.1 T222 = 2.6 E90 = 2.2 E209 = 2.2	Protein serine/threonine kinase activity	Serine/threonine kinase which acts as a master kinase, phosphorylating and activating a subgroup of the AGC family of protein kinases. Its targets include: protein kinase B (PKB/AKT1, PKB/AKT2, PKB	3-phosphoinositide-dependent protein kinase 1
DB07535	2va5, C8C	D32 = 2.6,3.0 D228 = 3.2	-9.1 T222 = 2.5 E90 = 2.3 E209 = 2.2,2.7	Peptidase activity	Responsible for the proteolytic processing of the amyloid precursor protein (APP). Cleaves at the N-terminus of the A-beta peptide sequence, between residues 671 and 672 of APP, leads to the genera	Beta-secretase 1

^a Details taken from drug bank.

Table 5
Difference in binding energies at different chains of same protein with different binding residues.

Protein	Binding Energy (Kcal/mol)	Binding residues and bond length				
6wu	-8.4	T257 = 2.7	Y305 = 2.8	E252 = 2.4		
6wu_A	-5.8	Q174 = 2.7	E203 = 2.5	M206 = 2.3		
6wu_B	-7.1	A153 = 2.7	Y154 = 2.4	D76 = 2.3		
6wu_C	-7.1	Y154 = 2.3	T74 = 2.7	A153 = 2.4	N156 = 2.6	
6wu_D	-6	Y137 = 2.1	K126 = 3.1	L125 = 3.1		

navigate large alignments of similar sequences. It gives percentage of identity also and available at <https://4virology.net/virology-ca-tool/s/base-by-base/>. **DataWarrior** combines dynamic graphical views and interactive row filtering with chemical intelligence. Scatter plots, box plots, bar charts and pie charts not only visualize numerical or category data, but also show trends of multiple scaffolds or compound substitution patterns. This software was used in this research to select drugs according to the rule of three and available at <http://www.openmolecules.org/datawarrior/download.html>. **Molinspiration** offers broad range of cheminformatics software tools supporting molecule manipulation and processing, including SMILES and SD file conversion, normalization of molecules, etc. and available at <https://www.molinspiration.com/cgi-bin/properties>.

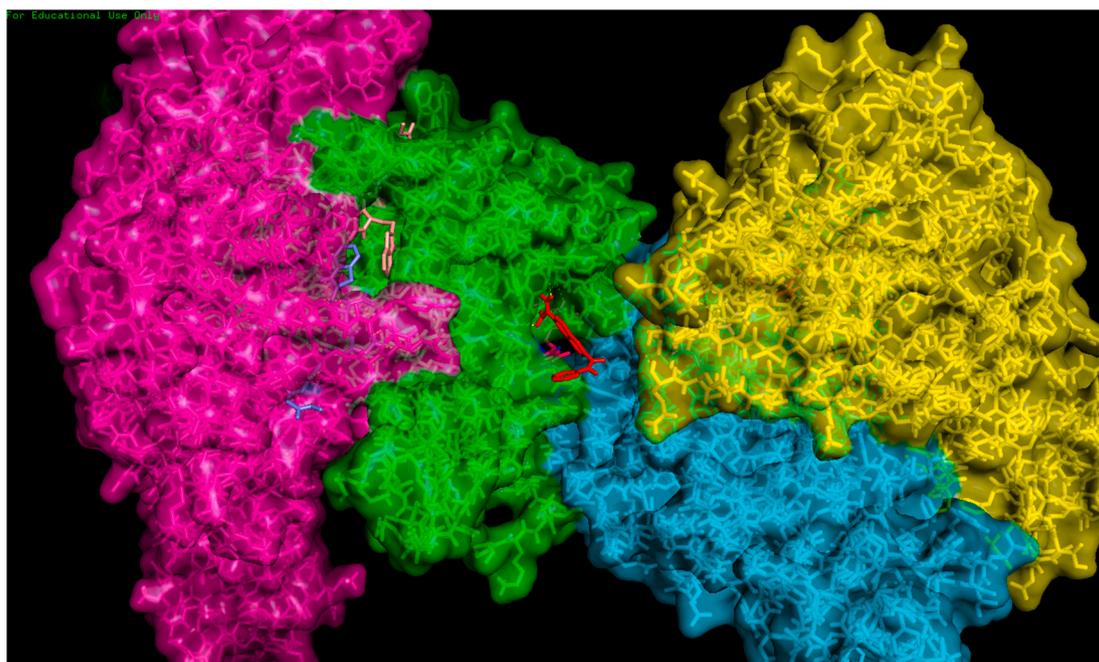


Fig. 1. Drug DB01977 binds with the binding pocket between the three chains, A, B and C of 6wu.pdb. Chain A is green, chain B is blue, chain C is yellow and chain D is magenta. The drug DB01977 is red with binding bonds with chain A (green). This shows that drug binds with the binding pockets situated between chains than binding pockets of individual chains for better affinity. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

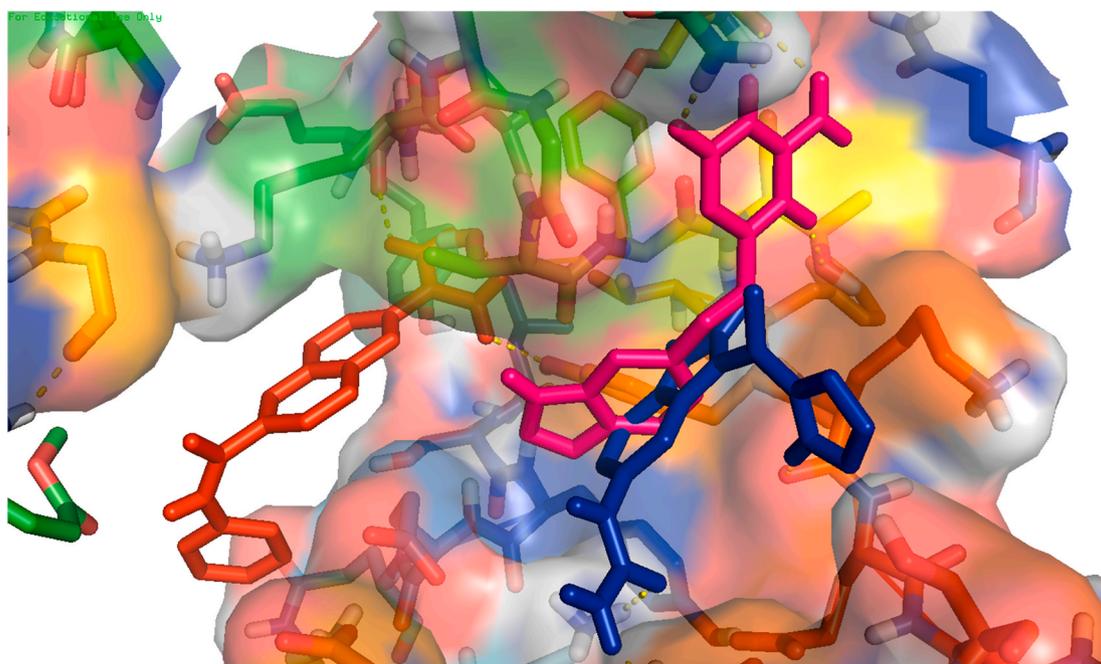


Fig. 2. Protein 6WUU.pdb showing classing/overlapping poses of drugs DB07132, blue and DB0753, magenta in the close by binding pockets. The binding residues and bond lengths are already calculated. It also shows non overlapping pose of drugs DB01977, red and DB07132 or DB01977 and DB07535. Blocking the function of NSP by the drugs in both cases, overlapping and non-overlapping needs to be proved in wet lab. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 6

Comparison of pharmacophores of drugs.

Drug	Remdesivir	Ritonavir	Chloroquine	Darunavir	Lopinavir	Azithromycin	Elbasvir	DB01977	DB07132	DB07535
(Molinspiration) milogP	2.82	7.51	5	4.32	5.69	2.73	8.85	2.37	1.71	1.9
Total polar surface area (A ²)	203.57	145.8	28.16	140.4	120	180.09	188.8	78.97	100.88	87.57
Number of atoms	42	50	22	38	46	52	65	22	21	19
Molecular weight	602.59	721	319.88	547.7	628.8	749	882	289.3	282.3	254.3
Number of bond donors	14	11	3	10	9	14	16	4	6	5
Number of bond acceptors	5	4	1	4	4	5	4	4	4	4
Number of violations	2	3	1	1	2	2	3	0	0	0
Number of rotatable bonds	14	18	8	12	15	7	13	3	3	3
Volume(A ³)	523.04	663.1	313.12	490.5	608	736.4	799.4	264.4	247.9	229.1
Mutagenic	No	No	Yes	No	No	No	No	No	No	No
Tumorigenic	Yes	No	No	No	No	No	No	No	No	No
Irritant	Yes	No	Yes	No	Yes	No	No	No	No	No
Reproductive effect	Yes	No	No	No	No	No	No	No	No	No
clogP	0.3	4.72	4.01	2.24	4.85	1.66	7.04	2.69	0.19	1.07
Solubility	-4.99	-6.07	-4.06	-3.96	-6.13	-3.09	-8.33	0.4	-2.54	-3.58
Druglikeness	-30.39	-8.93	7.39	-12.8	7.64	13.85	-7.68	-0.08	4.41	0.49
Drug-Score	0.05	0.13	0.25	0.29	0.17	0.48	0.06	0.59	0.92	0.71

Table 7

Bioactivity table for drugs.

Properties	DB01977	DB07132	DB07535
GPCR ligand	0.31	0.12	0.25
Ion channel modulator	0.08	-0.1	-0.13
Kinase inhibitor	0.19	0.58	0.43
Nuclear receptor ligand	-0.27	-0.36	-0.22
Protease inhibitor	0.59	-0.33	-0.27
Enzyme inhibitor	0.11	0.2	0.52

4. Results

1 NSP 3, 5, 11, 14 and 15 as targets for drugs DB01977, DB07132 and DB07535

There are 471 proteins downloaded for molecular docking, [Supplementary Table S1](#). The chains, the details of sequences, the ligands, the

sequence similarity percentage and the RMSD score for structural similarity for each downloaded protein from PDB were analyzed, [Supplementary Table S2](#). The downloaded drugs from the drug bank database were processed to only 31 drugs in accordance with the rule of three [45], [Supplementary Table S3](#).

Using [Table S2](#), proteins were selected based on the RMSD score around 0.5, cut off for selection of highly similar proteins and the selected proteins had above 88% sequence similarity. This was very much sufficient and 16 such highly similar proteins based on sequence and structural similarity were selected for docking, [Table 1](#). Docking was done between the highly similar 16 proteins downloaded from the PDB and the selected 31 drugs using PyRx. The docking score, binding energy for each protein with all chains and individual chain was recorded from the first pose with RMSD = 0 out of the nine poses of the PyRx, [Supplementary Table S4](#). With the help of [Table S4](#), 18 proteins/chains and 10 drugs were finally selected from the docking results, cut off binding energy less than or equal to -6 kcal/mol for further analysis, [Table 2](#).

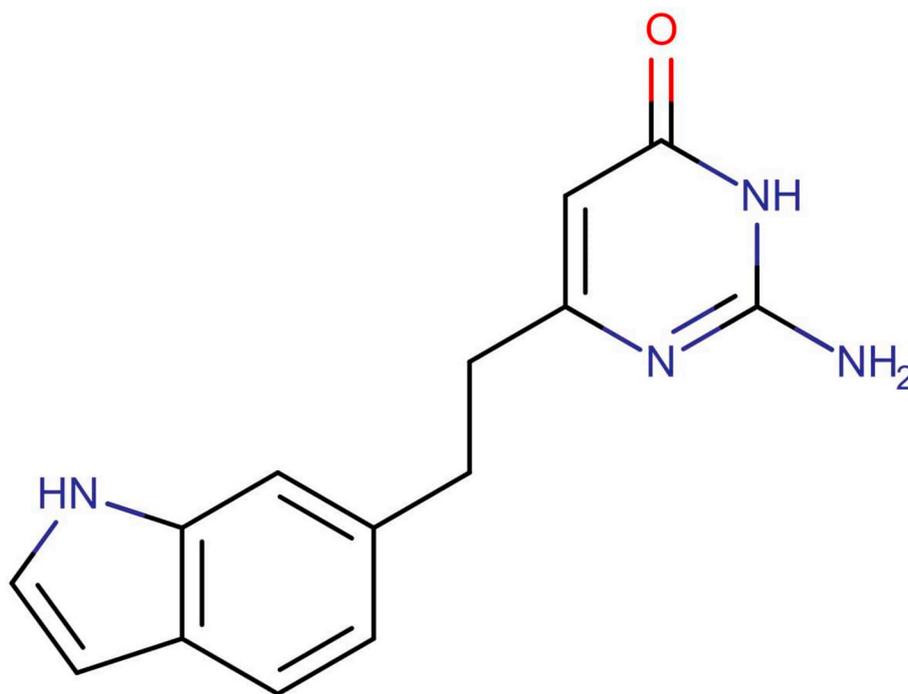


Fig. 3a. Chemical structure of DB07535, with smiles, NC1=NC(CCC2=CC3=C(C=CN3)C=C2) = CC(=O)N1.

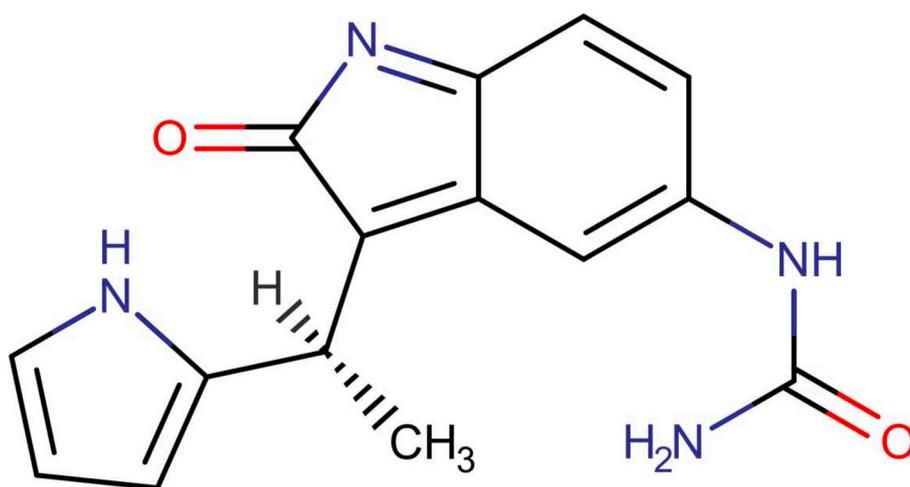


Fig. 3b. Chemical structure of DB07132, with smiles, [H][C@](C)(C1=CC=CN1)C1=C2C=C(NC(N)=O)C=CC2=NC1=O

NSPs 3, 5, 11, 14 and 15 show better binding association, low (less than -8KCal/mol) binding energy with the docked drugs, DB01977, DB07132 and DB07535 and hence, selected as target NSPs in this research also, Table 2. The highly scored NSP-drug complex, lower (more negative) binding energy was further analyzed with respect to binding pockets, binding residues, bond length etc. using PyMOL, Table 3.

2. Target proteins of DB01977, DB07132 and DB07535 and their binding association with NSPs

The target proteins of these drugs, 1owe-675 (675 is DB01977) complex, 2pe1-517 (517 is DB07132) complex and 2va5-C8C (C8C is DB07535) complex were first analyzed, secondly docked using PyRx and analyzed using PyMOL and finally compared with NSPs to establish the difference, Tables 3 and 4. It was found that the binding pockets were different due to different resolutions of the target proteins however, the binding residues are similar due to the pharmacophore of the drugs.

Therefore, we found that these drugs show equal affinity with the binding pockets of the targets and NSPs, low binding energy that is more negative value as per the protein-drug complex, Table 4.

3. Drugs, DB01977 and DB07535 and NSPs 5 and 15 have same/similar binding pockets.

Drugs DB01977 and DB07535 binds with same binding pockets of NSPs 5 and 15, Table 3 however, NSP 11 and NSP 14 have completely different binding pockets for these drugs. Moreover, binding pockets for both drugs have few sharing residues which suggests that binding pockets of these two drugs are situated very close for each drug, Table 3.

4. Association of binding pockets of NSPs for the drug, DB07132 and their pharmacophore.

DB07132 has highest polarizability, polar surface area, PKA and

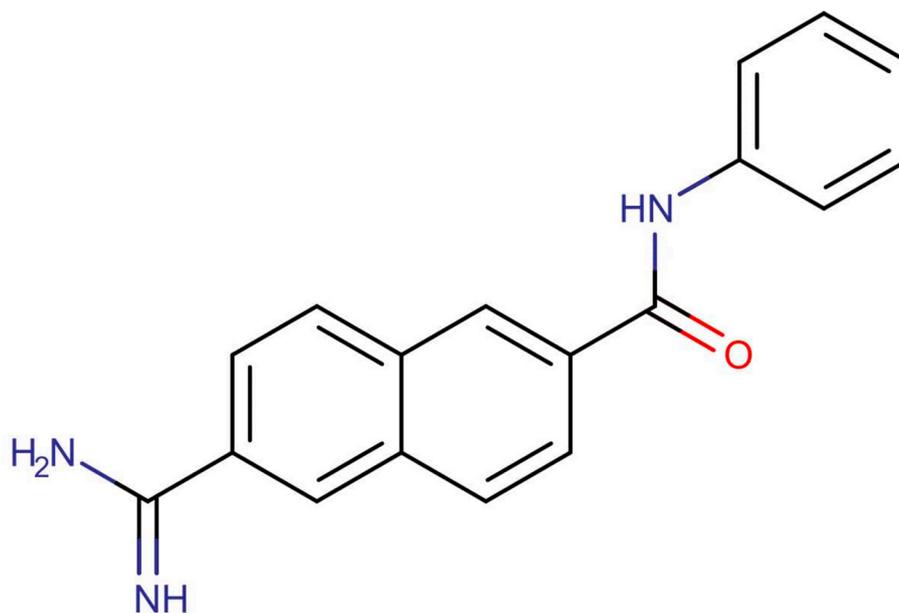


Fig. 3c. Chemical structure of DB01977 with smiles, NC(=N)C1=CC=C2C=C(C(=CC2=C1)C(=O)NC1=CC=CC=C1).

strongest acidity among the selected drugs keeping apart the constant parameters, number of rotatable bonds, hydrogen donors and acceptors to 3, [Supplementary Table S3](#). The drug binds with the NSPs with the help of many binding residues with lower (less than -8 KCal/mol) binding energy, [Table 3](#) and the number of residues involved in binding by DB07132 is more compared with other two drugs, [Table 3](#) suggesting that there are matching pharmacophores between NSPs and the drug.

5. Discussion

Computational receptor based drug designing depends on the integrity of receptors selected for drug designing [49] and here the receptors are highly similar to NSPs of SARS-CoV-2 in sequence and structure. Although many research and reviews state about the ideal sequence similarity from 30% to 40% [50], the selected proteins here are above 88% sequence similarity due to selected RMSD around 0.5 and this cut off is to minimize the structural deviation from the predicted structure. Moreover, this was to get enough number of proteins for each NSP, [Table 1](#) and as aimed, we had downloaded highly similar proteins from PDB for NSPs 3, 5, 10, 11, 14 and 15. The NSPs have different roles in virus life cycle such as PL PRO, 3CL PRO, RNA dependent RNA polymerase, RdRp, endonucleases etc. This screening and selection make it worth docking to reduce the differences in the dry and wet lab experiments and the false positive error. Like in this research, these NSPs are already on the targets for drug designing [51,52].

Actually, 16 proteins with 10 drugs have been docked to get 88 docking results, [Table S4](#) and this is because of docking with individual chains of a protein and the entire protein to analyze the binding affinity at different pockets of each NSP. For the difference in binding pockets with different chains of proteins, a special analysis was done with respect to binding residues and bond length in each chain of the same protein, [Table 5](#). This establishes the fact that the drugs mostly bind with binding pockets situated between the chains rather than the binding pockets of individual chain, [Fig. 1](#). Like in this research, the proteins, NSP 3, Papain-Like Protease, PL-PRO and NSP 5, 3C-like protease, 3CL-PRO are targets for drug discovery in other studies [53] also. This drug discovery is because of their important functions, ORF1ab is processed by proteases, PL-PRO and 3CL-PRO into replication complexes of positive-stranded RNA viruses and stored in double membrane vehicles in the cytoplasm which are essential for viral RNA replication [54,55].

Binding depends on many parameters of receptors and ligand/drugs

[56]. Drugs DB01977 and DB07535 have same binding pockets of NSPs 5 and 15 which highlights the matching pharmacophore of drug and the binding pockets. Additionally, we analyzed different poses of three drugs with the NSPs for their very close pockets and sharing binding residues. It was found that the poses of two drugs, DB07535 and DB07132 overlap in their binding pockets such that their effect may or may not be exhibited completely in a bound state. This is the case with the clinicians administering different drugs at a time such that drugs may have same binding pockets with clashing their conformation to exhibit different effect or no effect of each drug, [Fig. 2](#). However, the binding poses of the drugs, DB01977 and DB07535/DB07132 are different in protein, 6wu0.pdb, [Fig. 2](#) which may be having the effect of each drug.

Moreover, when a drug is administered, as per this research, each of the three drugs may bind considerably with all NSPs 3, 5, 11, 14 and 15 or may be along with other NSPs at a time to bring biological and clinical effects on the virus life cycle. Furthermore, we compared the pharmacophore of some well-known (see [Table 6](#)) drugs with these three drugs using Molinspiration online software, [Table 6](#) and analyzed bioactivity of these drugs, [Table 7](#) using the same software along with their structures, [Fig. 3 \(a\), \(b\) & \(c\)](#). The well-known drugs have high volume and molecular weight which may or may not bind easily with variable binding pockets of NSPs to bring biological effect. The score in each drug in [Table 7](#) shows that these drugs are positive potential drugs as protease inhibitor, kinase inhibitor etc.

6. Conclusion

It is clear that these three drugs, preferably drug DB07132 bind effectively with NSP-3, NSP-5, NSP-11, NSP-14 and NSP-15 with shorter bond length to bring biological effect in SARS-CoV-2 in turn to humans. In conclusion, binding pockets of proteins are well matched with the pharmacophore of drugs and with polar surface of drugs less than or equal to 100 \AA^2 , drugs, DB01977, DB07132 and DB07535 bind individually and effectively with NSPs 3, 5, 11, 14 and 15 of ORF1ab of SARS-CoV-2 genome to bring changes in the activity of SARS-CoV-2 which may be useful for biological and clinical considerations.

List of abbreviations

Not applicable, all abbreviations are expanded in the text itself.

Consent for publication

Not Applicable.

Funding

None.

Authors' contributions

One author contributed for the article.

Declaration of competing interest

None.

Acknowledgements

Authors are thankful to the college for helping in completing the experiments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbrep.2020.100847>.

References

- [1] S. Hussain, et al., Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus, *J. Virol.* 79 (9) (2005) 5288–5295.
- [2] A.R. Jauregui, et al., Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling, *PLoS One* 8 (4) (2013), e62416.
- [3] K. Narayanan, et al., Coronavirus nonstructural protein 1: common and distinct functions in the regulation of host and viral gene expression, *Virus Res.* 202 (2015) 89–100.
- [4] R.L. Graham, et al., The nsp2 proteins of mouse hepatitis virus and SARS coronavirus are dispensable for viral replication, *Adv. Exp. Med. Biol.* 581 (2006) 67–72, https://doi.org/10.1007/978-0-387-33012-9_10.
- [5] E.J. Snijder, et al., Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage, *J. Mol. Biol.* 331 (5) (2003) 991–1004.
- [6] B.W. Neuman, et al., Proteomics analysis unravels the functional repertoire of coronavirus nonstructural protein 3, *J. Virol.* 82 (11) (2008) 5279–5294.
- [7] J. Tan, et al., The SARS-unique domain (SUD) of SARS coronavirus contains two macrodomains that bind G-quadruplexes, *PLoS Pathog.* 5 (5) (2009), e1000428.
- [8] M.A. Johnson, et al., SARS coronavirus unique domain: three-domain molecular architecture in solution and RNA binding, *J. Mol. Biol.* 400 (4) (2010) 724–742.
- [9] P. Serrano, et al., Nuclear magnetic resonance structure of the N-terminal domain of nonstructural protein 3 from the severe acute respiratory syndrome coronavirus, *J. Virol.* 81 (21) (2007) 12049–12060.
- [10] Y. Sakai, et al., Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication, *Virology* 510 (2017) 165–174.
- [11] J.H. Lu, et al., Sequence analysis and structural prediction of the severe acute respiratory syndrome coronavirus nsp5, *Acta Biochim. Biophys. Sin.* 37 (7) (2005) 473–479.
- [12] K. Anand, et al., Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs, *Science* 300 (5626) (2003) 1763–1767.
- [13] C.C. Stobart, et al., Chimeric exchange of coronavirus nsp5 proteases (3CLpro) identifies common and divergent regulatory determinants of protease activity, *J. Virol.* 87 (23) (2013) 12611–12618.
- [14] E.M. Cottam, M.C. Whelband, T. Wileman, Coronavirus NSP6 restricts autophagosome expansion, *Autophagy* 10 (8) (2014) 1426–1441.
- [15] D. Benvenuto, et al., Evolutionary analysis of SARS-CoV-2: how mutation of Non-Structural Protein 6 (NSP6) could affect viral autophagy, *J. Infect.* 81 (1) (2020) e24–e27.
- [16] A.J.W. te Velthuis, S.H.E. van den Worm, E.J. Snijder, The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension, *Nucleic Acids Res.* 40 (4) (2012) 1737–1747.
- [17] 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.
- [18] B. Krichel, et al., Processing of the SARS-CoV ppla/ab nsp7-10 region, *Biochem. J.* 477 (5) (2020) 1009–1019.
- [19] P. Kumar, et al., The nonstructural protein 8 (nsp8) of the SARS coronavirus interacts with its ORF6 accessory protein, *Virology* 366 (2) (2007 Sep 30) 293–303, <https://doi.org/10.1016/j.virol.2007.04.029>. Epub 2007 May 25.
- [20] G. Sutton, et al., The nsp9 replicase protein of SARS-coronavirus, structure and functional insights, *Structure* 12 (2) (2004 Feb) 341–353, <https://doi.org/10.1016/j.str.2004.01.016>.
- [21] Z.J. Miknis, et al., Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth, *J. Virol.* 83 (7) (2009) 3007.
- [22] D.R. Littler, et al., Crystal Structure of the SARS-CoV-2 Non-structural Protein 9, *Nsp9*, *bioRxiv*, 2020, 2020.03.28.013920.
- [23] M. Rosas-Lemus, et al., The Crystal Structure of Nsp10-Nsp16 Heterodimer from SARS-CoV-2 in Complex with S-Adenosylmethionine, *bioRxiv*, 2020, 2020.04.17.047498.
- [24] Y. Ma, et al., Structural basis and functional analysis of the SARS coronavirus nsp14â€nsp10 complex, *Proc. Natl. Acad. Sci.* (2015), 201508686.
- [25] M. Bouvet, et al., Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes, *J. Biol. Chem.* 289 (37) (2014) 25783–25796.
- [26] K.A. Ivanov, et al., Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase, *J. Virol.* 78 (11) (2004) 5619–5632.
- [27] P. Chen, et al., Biochemical characterization of exoribonuclease encoded by SARS coronavirus, *J. Biochem. Mol. Biol.* 40 (5) (2007) 649–655.
- [28] E. Minskaia, et al., Discovery of an RNA virus 3'→5' exoribonuclease that is critically involved in coronavirus RNA synthesis, *Proc. Natl. Acad. Sci. U. S. A.* 103 (13) (2006) 5108–5113.
- [29] K.A. Ivanov, et al., Major genetic marker of nidoviruses encodes a replicative endoribonuclease, *Proc. Natl. Acad. Sci. U. S. A.* 101 (34) (2004) 12694–12699.
- [30] X. Deng, S.C. Baker, An "Old" protein with a new story: coronavirus endoribonuclease is important for evading host antiviral defenses, *Virology* 517 (2018) 157–163.
- [31] M. Zhang, et al., Structural biology of the arterivirus nsp11 endoribonucleases, *J. Virol.* 91 (1) (2017).
- [32] D. Forni, et al., Extensive positive selection drives the evolution of nonstructural proteins in lineage C betacoronaviruses, *J. Virol.* 90 (7) (2016) 3627–3639.
- [33] R. Cagliani, et al., Computational inference of selection underlying the evolution of the novel coronavirus, severe acute respiratory syndrome coronavirus 2, *J. Virol.* 94 (12) (2020) e00411–e00420.
- [34] A.K. Chakraborty, Coronavirus Nsp2 protein homologies to the bacterial DNA Topoisomerase I and IV suggest Nsp2 protein is a unique RNA Topoisomerase with novel target for drug and vaccine development, *Viol. Mycol.* 9 (185) (2020) 1.
- [35] C.T. Cornillez-Ty, et al., Severe acute respiratory syndrome coronavirus nonstructural protein 2 interacts with a host protein complex involved in mitochondrial biogenesis and intracellular signaling, *J. Virol.* 83 (19) (2009) 10314.
- [36] M. Oostra, et al., Localization and membrane topology of coronavirus nonstructural protein 4: involvement of the early secretory pathway in replication, *J. Virol.* 81 (22) (2007) 12323–12336.
- [37] A.M. Mielech, et al., Nidovirus papain-like proteases: multifunctional enzymes with protease, deubiquitinating and deISGylating activities, *Virus Res.* 194 (2014) 184–190.
- [38] Y. Xiao, et al., Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity, *J. Virol.* 86 (8) (2012) 4444.
- [39] Y.W. Tan, et al., Coronavirus infectious bronchitis virus non-structural proteins 8 and 12 form stable complex independent of the non-translated regions of viral RNA and other viral proteins, *Virology* 513 (2018) 75–84.
- [40] M. Bouvet, et al., RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex, *Proc. Natl. Acad. Sci. U. S. A.* 109 (24) (2012) 9372–9377.
- [41] E. Decroly, et al., Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity, *J. Virol.* 82 (16) (2008) 8071–8084.
- [42] D.J. Deming, et al., Processing of open reading frame 1a replicase proteins nsp7 to nsp10 in murine hepatitis virus strain A59 replication, *J. Virol.* 81 (19) (2007) 10280–10291.
- [43] S.G. Fang, et al., Proteolytic processing of polyproteins 1a and 1ab between non-structural proteins 10 and 11/12 of Coronavirus infectious bronchitis virus is dispensable for viral replication in cultured cells, *Virology* 379 (2) (2008) 175–180.
- [44] A.J.W. te Velthuis, et al., The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent, *Nucleic Acids Res.* 38 (1) (2010) 203–214.
- [45] M. Congreve, et al., A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* 8 (19) (2003) 876–877.
- [46] C.A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution, *Drug Discov. Today Technol.* 1 (4) (2004) 337–341.
- [47] Z. Courmia, B. Allen, W. Sherman, Relative binding free energy calculations in drug discovery: recent advances and practical considerations, *J. Chem. Inf. Model.* 57 (12) (2017) 2911–2937.
- [48] Y. Fukunishi, et al., Prediction of Protein–compound binding energies from known activity data: docking-score-based method and its applications, *Mole. Info.* 37 (6–7) (2018), 1700120.
- [49] V. Cherezov, et al., High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, *Science* (New York, N.Y.) 318 (5854) (2007) 1258–1265.
- [50] Z. Xiang, Advances in homology protein structure modeling, *Curr. Protein Pept. Sci.* 7 (2006) 217–227.
- [51] J.H. Lu, et al., Sequence analysis and structural prediction of the severe acute respiratory syndrome coronavirus nsp5, *Acta Biochim. Biophys. Sin.* 37 (7) (2005 Jul) 473–479, <https://doi.org/10.1111/j.1745-7270.2005.00066.x>.

- [52] Y. Ma, et al., Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex, *Proc. Natl. Acad. Sci. U. S. A.* 112 (30) (2015 Jul 28) 9436–9441, <https://doi.org/10.1073/pnas.1508686112>. Epub 2015 Jul 9.
- [53] Y. Chen, et al., X-ray structural and functional studies of the three tandemly linked domains of non-structural protein 3 (nsp3) from murine hepatitis virus reveal conserved functions, *J. Biol. Chem.* 290 (42) (2015) 25293–25306.
- [54] R. Gosert, et al., RNA replication of mouse hepatitis virus takes place at double-membrane vesicles, *J. Virol.* 76 (8) (2002) 3697–3708.
- [55] J. Ziebuhr, The coronavirus replicase: insights into a sophisticated enzyme machinery, *Adv. Exp. Med. Biol.* 581 (2006) 3–11.
- [56] T. Pantzar, A. Poso, Binding affinity via docking: fact and fiction, *Molecules* 23 (8) (2018) 1899.