 doi:10.20944/preprints202003.0085.v1

Structural elucidation of SARS-CoV-2 vital proteins: computational methods reveal potential drug candidates against Main protease, Nsp12 RNA-dependent RNA polymerase and Nsp13 helicase

Muhammad Usman Mirza[1], Matheus Froeyen[1*]

[1] Department of Pharmaceutical and Pharmacological Sciences, Rega Institute for Medical Research, Medicinal Chemistry, University of Leuven, B-3000 Leuven, Belgium

Corresponding author:

Prof. Matheus Froeyen

Department of Pharmaceutical and Pharmacological Sciences

Rega Institute for Medical Research,

Medicinal Chemistry, University of Leuven, B-3000 Leuven, Belgium

Email: mathy.froeyen@kuleuven.be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

Abstract

The recently emerged severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) caused a major outbreak of coronavirus disease 2019 (COVID-19) and instigated a widespread fear and has threatened global health security. Although phenomenal efforts are in progress to effectively combat this COVID-19 outbreak. Still, no licensed antiviral drugs or vaccines are available, and treatment is limited to supportive care and few repurposed drugs. In this urgency situation, computational drug discovery methods provide both an alternative and a supplement to tiresome high-throughput screening, particularly in the hit-to-lead-optimization stage. Identification of small molecules that specifically target viral replication apparatus has shown the most successful strategy in antiviral drug discovery. The present study deals with the identification of potential compounds that specifically interact with SARS-CoV-2 vital proteins, including main protease (Mpro), Nsp12 RNA-dependent-RNA-polymerase (RdRp) and Nsp13 helicase. A constructive and integrated virtual screening efforts together with molecular dynamics simulations identified potential binding modes and favourable molecular interaction profile of corresponding compounds. Moreover, structurally important binding site residues in conserved motifs located inside the active site are elucidated, which displayed relative importance in ligand binding based on residual energy decomposition analysis. Although the current study lacks experimental validation, the structural information obtained from this computational study paved the way to identify and design specific targeted inhibitors to combat COVID-19 outbreak.

Keywords: SARS-CoV-2, COVID-19, CoV-Mpro, CoV-Nsp12 polymerase, CoV-Nsp13 helicase

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) originated in Wuhan, China (Huang, Wang et al. 2020, Li, Guan et al. 2020), causes respiratory infection and is known to be contagious (Parry 2020). The ongoing 2019–20 Wuhan outbreak of coronavirus disease 2019 (COVID-19) (Benvenuto, Giovanetti et al. 2020, Perlman 2020, Wang, Horby et al. 2020), as of 29 February 2020, has led to 2,838 deaths along with 85,403 confirmed cases of infection of which 79,394 reportedly have been within mainland China (WHO 2020). World Health Organization (WHO) has declared SARS-CoV-2 a global health emergency on 30th January 2020 (WHO 2020).

SARS-CoV-2 belongs to the *Coronaviruses* family (Kasmi, Khataby et al. 2020). Other coronaviruses are known to cause common cold which may lead to severe diseases including the Middle East respiratory syndrome (MERS) (Drosten, Günther et al. 2003, Azhar, Hui et al. 2019) and severe acute respiratory syndrome (SARS) (Peiris, Lai et al. 2003, Yin and Wunderink 2018). The recent SARS-CoV-2 is the seventh known human coronavirus (HCoV) from the same family after 229E, NL63, OC43, HKU1, MERS-CoV, and SARS-CoV (Zhu, Zhang et al. 2020).

Currently, there is no licensed drug or vaccine available for SARS-CoV-2, therefore, treatment is focused on the alleviation of symptoms which may include dry cough, fever and pneumonia (Hui, Madani et al. 2020). In 2003, the SARS outbreak caused more than 800 deaths worldwide

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

(Drosten, Günther et al. 2003, Kuiken, Fouchier et al. 2003, Peiris, Lai et al. 2003). Following the SARS outbreak, a series of helicase and protease (Mpro) inhibitors were reported to prevent viral replication (Bacha, Barrila et al. 2004, Blanchard, Elowe et al. 2004, Jain, Pettersson et al. 2004, Kao, Tsui et al. 2004, Wu, Jan et al. 2004, Tanner, Zheng et al. 2005). For treating coronavirus-related pneumonia, Chinese Center for Disease Control and Prevention (CCDC) is currently testing existing pneumonia treatments for SARS-CoV-2. Existing antivirals, including protease inhibitors (indinavir, saquinavir and lopinavir/ritonavir) as well as RNA polymerase inhibitors including remdesivir (Paules, Marston et al. , Li and De Clercq 2020, Morse, Lalonde et al. 2020) are being tested against SARS-CoV-2. Recently, antiviral efficiency of several FAD-approved drugs including remdesivir (EC50 = 0.77 μM) and chloroquine (EC50 = 1.13 μM) against a clinical isolate of SARS-CoV-2 *in vitro* showed potential inhibition at low-micromolar concentration (Wang, Cao et al. 2020). The efficacy of remdesivir is evident from a recent recovery of US patient infected with SARS-CoV-2 after intravenous treatment (Holshue, DeBolt et al. 2020), while chloroquine is being evaluated in an open-label trial (ChiCTR2000029609). Following this, two phase III trials (NCT04252664 and NCT04257656) were also initiated to evaluate intravenous remdesivir in patients infected with SARS-CoV-2. Others include Nafamostat (EC50 = 22.50 μM), Nitazoxanide (EC50 = 2.12 μM) and Favipiravir (EC50 = 61.88 μM) (Wang, Cao et al. 2020). Randomized trials to evaluate the efficacy of favipiravir plus baloxavir marboxil (ChiCTR2000029544) and favipiravir plus interferon-α (ChiCTR2000029600) are also initiated for patients with SARS-CoV-2. Moreover, the efficacy of interferon beta (Paules, Marston et al.) and previously identified monoclonal antibodies (mAbs) is also under investigation (Chen, Liu et al. 2020) for treating SARS-CoV-2.

CoVs are single-stranded positive-sense RNA (+ssRNA) viruses with 5'-cap and 3'-poly-A tail. The genome size of SARS-CoV-2 is ~30kb which is the largest among all RNA viruses (Chen, Liu et al. 2020, Gralinski and Menachery 2020). Typical CoV genome contains at least six open reading frames (ORFs). The first ORF (ORF1a/b) is about two-thirds of the whole genome length and encodes 16 non-structural proteins (nsp1-16). ORFs near 3' end of the genome encodes at least four main structural proteins including spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins. Most of the non-structural proteins are known to play a vital role in CoV replication. Structural proteins, however, are important for virion assembly as well as for causing CoV infection. Furthermore, particular structural and accessory proteins, such as HE protein, are also encoded by the CoV genome (Chen, Liu et al. 2020). SARS-CoV-2 maintains ~80% nucleotide identity to the original SARS epidemic viruses (Gralinski and Menachery 2020). Crystal structures including, ZC45 and ZXC21, two bat SARS-like CoVs, share ~89% sequence identity with SARS-CoV-2. Phylogenetic analysis has indicated that the novel CoV is a viral recombinant of previously identified bat coronaviruses (Chan, Kok et al. 2020). A recent report has identified bat CoV sequence, RaTG3, with 92% sequence identity with the SARS-CoV-2, which strengthens the aforementioned fact about its origin (Gralinski and Menachery 2020).

To address the current outbreak, the development of wide-spectrum inhibitors against CoV-associated diseases is an attractive strategy. However, it requires identification of a conserved target region within the whole genus Coronavirus (Yang, Xie et al. 2005, De Clercq 2006). On

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

the contrary, all structural proteins including S, E, M, HE, and N proteins among different CoVs have reported considerable variations (Marra, Jones et al. 2003, Rota, Oberste et al. 2003, Woo, Lau et al. 2005) which adds more complexity towards identifying inhibitors against SARS-CoV-2. Subsequently, the Nsp12 RNA-dependent RNA polymerase (Kirchdoerfer and Ward 2019), Nsp13 helicase (Jia, Yan et al. 2019), and main protease (Mpro) or chymotrypsin-like protease (3CLpro) (Bacha, Barrila et al. 2004, Yang, Xie et al. 2005) constitute attractive potential non-structural protein targets for consideration, which are highly conserved among coronaviruses (Xu, Zhao et al. 2020).

However, no structural data was available for SARS-CoV-2 proteins at the start of this study, but high sequence similarity offered to model vital SARS-CoV-2 proteins for rational drug design and downstream modification will be possible for drug leads. The current study focusses on the structural elucidation of a critically important nCoV-Nsp12 polymerase, Nsp13 helicase and Mpro together with the virtual screening strategies, which identified potential hits through molecular dynamics simulation analysis. Yet, the current study warrants experimental elucidation of proposed hits, however, the presented brief structural comparisons and identified potential hits could become the starting point for structure-guided drug discovery.

Methods

Protein modelling

Homology models of critical proteins of SARS-CoV-2, including Mpro, Nsp12 polymerase, Nsp13 helicase were predicted using SWISSMODEL (Schwede, Kopp et al. 2003) as there was no crystal structure available at the time of this study. The templates with the highest identity were selected, and models were generated accordingly. The binding pocket residues were predicted from an analytical tool, COACH meta-server (Yang, Roy et al. 2013) and compared after superimposing various closely related X-ray resolved co-crystallized structures with bound inhibitors. The generated homology models were refined using 20ns molecular dynamics (MD) simulations to remove steric clashes and optimized side-chain geometry.

The AMBER18 simulation package was utilized for unrestrained MD simulations using AMBER ff99SB force field (Maier, Martinez et al. 2015). The stepwise minimization and equilibration protocol were performed accordingly (described in previous studies (Mirza, Rafique et al. 2016, Mirza, Vanmeert et al. 2019)), and solvated system for each model with explicit TIP3 water molecules was submitted to a production run of 20 ns at constant 300K temperature and 1 bar pressure. The protein's backbone conformation was analyzed using CPPTRAJ module (Roe, Cheatham III et al. 2013), and most representative conformation was evaluated through Ramachandhran plots and selected for virtual screening purposes.

Structure-based virtual screening and comparative docking

The small molecules dataset was selected and downloaded in SMILES format from the ZINC database (Irwin, Shoichet et al. 2005) and uploaded accordingly in a Mcule, a fully automated drug discovery platform (Kiss, Sandor et al. 2012). Virtual screening pipeline was established using the following parameter; 1) drug-like filters, and 2) diversity selection was implemented to reduce the size of the library and to maximize the coverage of chemical space at the same

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

time. This procedure excluded the closest analogues based on the Tanimoto similarity coefficient and maximized the diversity of the potential identified active scaffolds from the huge dataset. Prior to docking, careful inspection of binding pocket residues was performed as depicted from COACH and structure comparisons. The docking grid was set up around the predicted binding site residues covering the key residues to reduce the search space for ligand optimization.

An automated in-built Autodock Vina docking engine was utilized to screen the composed docking library, which uses a gradient optimization method in its local optimization process to rank the best poses efficiently (Trott and Olson 2010). Individual compounds of the docking library together with the mcule database were docked iteratively into the respective binding site residues of the representative structures of SARS-CoV-2 Mpro, Nsp12 polymerase, Nsp13 helicase, and ranked based on the Vina empirical scoring function that approximates the ligand binding affinity in kcal/mol. Hence, various docking programs only estimate the real binding affinity, therefore it may be valuable to test different docking programs. The top hits obtained from first round of docking were then evaluated with Glide "extra precision" docking mode (Glide XP) as implemented in Schrödinger's Maestro modelling package (a trial version was utilized for this study). Corresponding proteins were prepared using protein preparation wizard by adding hydrogen atoms and missing chains, while short minimization of only the hydrogen atoms was carried out using the OPLS 3 force field. Selected ligands ranked by AutoDock Vina were prepared using LigPrep accordingly, and the grid box was generated with dimensions covering the predicted active site residues. Top hits were selected after extensive post-docking analysis by comparing the binding poses predicted from AD Vina and Glide.

Molecular dynamics simulations and energy calculations

To gain more understanding of the predicted molecular interactions, the best docked complexes were analysed over a period of 20ns MD simulations using AMBER 18 simulation package. The antechamber module of AMBER18 was used to generate atomic partial charges for the selected hits. We used the same MD simulation protocol as used previously on other small molecule studies (Ikram, Mirza et al. 2019, Mirza, Vanmeert et al. 2019), Briefly, tleap module of AMBER was used to prepare the overall simulation system by generating an octahedral box (10 Å) around the solute with explicit TIP3 water molecules. Charges were neutralized by adding Cl- ions in SARS-CoV-2 Nsp13 helicase, and Na+ ions in Mpro and Nsp12 polymerase. Energy minimization was implemented using steepest descent minimization of 5000 steps followed by a conjugate gradient minimization of 1000 steps. All atoms of the system were energy minimized with gradually reducing restraint force constant on the protein atoms (from 10 to 0 kcal/mol/Å^2) to optimize the solvent position. For equilibration, the same stepwise protocol was used started from, i) 10ps heating of system 10K to 300K with a Langevin thermostat (γ = 1.0 ps-1) keeping the volume constant with 5kcal/mol/Å2 restraint force constant on protein atom position; ii) same as the first step but 20ps without any positional restraints; iii) 20 ps MD at 300K using Langevin thermostat (γ = 0.5 ps-1) keeping the volume constant without any positional restraints; (iv) initial 50ps MD at 300K using Langevin thermostat (γ = 1.0 ps-1) under constant pressure (1 bar) without any positional restraints; (v) additional 400 ps MD at 300K, constant pressure without positional restraints. The prepared

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

solvated complex was subjected to a final 20ns MD simulations at constant temperature (300K) and pressure (1 bar). The root-mean-square-deviation analysed with the CPPTRAJ module. To gain rational insights into the different binding modes and residual contribution, the Molecular Mechanics/Generalized Born Solvent Area (MM/GBSA) method was employed and total binding free energies were calculated. The MM/GBSA approach is well detailed in binding free energy calculations (Hou, Wang et al. 2011) for antiviral inhibitors (Tan, Zu Chen et al. 2006, Srivastava, Sastry et al. 2012).

Results

Structural insights of SARS-CoV-2 proteins

In silico modelling work was initiated by analysing Wuhan seafood market pneumonia virus genome (NCBI genome ID MN908947), which was recently deposited by Wu et al., (Wu and Zhang 2020). For the current work, orf1ab polyprotein sequence (GenBank ID: QHD43415.1) was considered (a total of 7096 amino acids), which is proteolytically processed by virally encoded 3CLpro or Mpro into 16 nsps, including Nsp12 polymerase and Nsp13 helicase (Chen, Liu et al. 2020). In order to locate the potential Nsp12, Nsp13 and Mpro sequence, a multiple sequence alignment was performed with known SARS-CoV proteins (Uniprot ID: P0C6X7). The corresponding regions in SARS-CoV-2 genome were identified based on the sequence identity and similarity score. Sequence alignment revealed that the SARS-CoV-2 Nsp12 polymerase, Nsp13 helicase and Mpro shared high similarity with that of SARS.

Homology modelling was performed using SWISSMODEL and templates were identified against corresponding SARS-CoV-2 proteins. As expected, best templates were identified from SARS-CoV deposited in Protein Data Bank (PDB), having 96.35%, 99.83% and 96.08% identities with SARS-CoV Nsp12 (6NUR), Nsp13 (6JYT) and Mpro (2Z9J). Because of high sequence similarity (>96 %), the homology models revealed a strikingly conserved overall architecture. The superimposition of RdRp, helicase and Mpro with SARS-CoV proteins showed high structural similarity, with RMSD values of 1.35 Å, 1.41 Å and 1.79 Å respectively. Model validation through PROCHECK indicated 96.51% (Mpro) and 97.5% (RdRp) of all residues in Ramachandran favoured regions with 3.27% and 0.28% were rotamer outliers, while helicase showed 84.34% with 4.43% outliers. The models were further refined by 20ns MD simulation. With the availability of the crystal structure of SARS-CoV-2 Mpro (6LU7), homology model was also compared and found less than 0.5 Å RMSD which indicated the reliability of the generated model with only subtle difference. Overall, these models were reliable enough to perform virtual screening. These models revealed the conserved features especially in functional regions due to strikingly similar protein conformation, however, some of the structural features are briefly elucidated together with the identification of inhibitor binding site in SARS-CoV-2 proteins as shown in **Figure 1**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

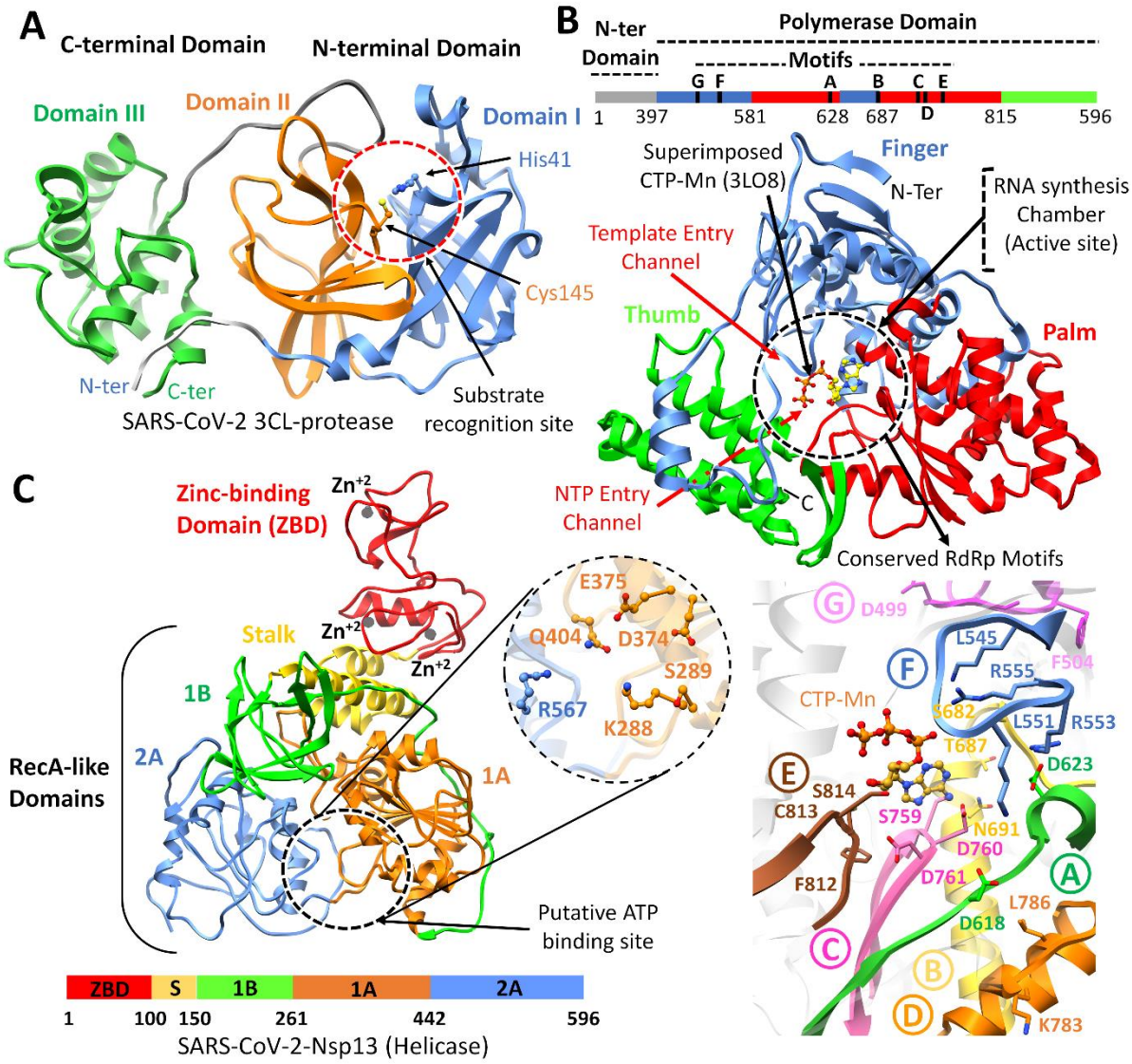

Figure 1: Structural representations of SARS-CoV-2 proteins. **(A)** Overall ribbon representation of SARS-CoV-2 main protease monomer composed of N-terminal domain I (cornflower blue) and domain II (orange), and C-terminal domain III (green). Substrate recognition site is circled and catalytic dyad residues, His41 and Cys145 are highlighted and labelled. **(B)** Linear schematic description of domain architecture of SARS-CoV-2 Nsp12 polymerase followed by its structure composed of thumb (green), palm (red), and fingers (cornflower blue) subdomains. The active site of Nsp12 polymerase is highlighted and arrangement of structurally conserved RdRp motifs in Nsp12 polymerase model coloured green, yellow, hot pink, orange, brown, cornflower blue and magenta for motifs A-F respectively are displayed in bottom right. Superposition of the polio virus elongation complex structure (PDB: 3OL8) CTP (orange) inside the predicted binding site also displayed. **(C)** Overall structure of SARS-CoV-2 Nsp13 helicase composed of ZBD (red), stalk (golden), 1B (green), 1A (orange) and 2A (cornflower blue) domains. Three zinc atoms are shown as dark grey spheres. The binding pocket residues are zoomed and labelled, and linear schematic diagram of the domain organization of SARS-CoV-2 Nsp13 helicase is displayed at the bottom.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

SARS-CoV-2 Mpro: The SARS-CoV-2 Mpro was predicted to contain 306 amino acids (located in the polyprotein, from 3264-3569 a.a). Structural analysis of SARS-CoV-2 Mpro revealed the same location of the active site as reported in SARS-Mpro, which is arranged in the cleft between domains I (8 – 99 a.a) and domain II (100-183 a.a). Both domains contributed one residue to the catalytic dyad composed of His41 and Cys145 **(Figure 1A)**, while these two domains were found connected by a long loop (184-199 a.a) to the helical domain III (200-306 a.a). The hydrogen bond formation between the His41 and Cys145 reported the right conformation of the catalytic dyad, which was found similar in nCoV-Mpro (Bianco, Salvatore et al. 2002). Substrates and inhibitors typically interact with Gln residues at the P1 position, and significant hydrophobic residues Phe/Leu/Met at P2 (Yang, Xie et al. 2005). Superimposition of several co-crystalized SARS-Mpro structures bound with inhibitors (Yang, Xie et al. 2005), N1 (1WOF), I2 (2D2D), N3 (2AMQ), N9 (2AMD) also revealed the similar position of S1, S2, S3 and S4 subsites, especially in the active site close to His41 and Cys145, which is crucial for substrate recognition (Yang, Xie et al. 2005), along with Tyr161 and His163 in the substrate-binding pocket (Chang, Chou et al. 2007). Docking grid was formed around these subsites for virtual screening, and top hits were identified based on the docking scores.

SARS-CoV-2 Nsp12 RdRp: The SARS-CoV-2 nsp12 polymerase was predicted to contain 932 amino acids (located in the polyprotein, from 4393 – 5324 a.a). The nsp12 comprised of N-terminal (1-397 a.a) and a polymerase domain (a.a. 398-919) when compared with SARS-CoV Nsp12 (6NUR). The polymerase domain adopted a structure resembling a cupped "right hand" like other polymerases (Bruenn 2003, Mirza, Vanmeert et al. 2019). The polymerase domain is comprised of a finger (398–581 and 628–687 a.a), a palm (582–627 and 688–815 a.a), and a thumb subdomain (816–919 a.a) **(Figure 1B)**. Like in MERS and SARS-Nsp12, the finger and the thumb subdomains of nCoV-RdRp contact each other, which configured the RdRp active site in the centre for the substrate access through template entry, template-primer exit, and NTP tunnels (Peersen 2017). Alongside, the SARS-CoV-2 nsp12 also revealed seven conserved motifs (A – G) arranged in the polymerase active site chamber, which are involved in a template and nucleotide binding and catalysis (Poch, Sauvaget et al. 1989, Bruenn 2003, Mirza, Vanmeert et al. 2019). The binding site of SARS-CoV-2 nsp12 was further elucidated through the superimposition of elongation complex of Poliovirus bound with CTP-Mn (3LO8) (Gong and Peersen 2010) and crystal Structure of Japanese encephalitis RdRp in complex with ATP (4HDH) (Surana, Satchidanandam et al. 2013). Docking grid was formed around the polymerase active site chamber covering the conserved motifs (A - G), and potential hits were identified.

SARS-CoV-2 Nsp13 helicase: The SARS-CoV-2 nsp13 was predicted to contain 596 amino acids (located in orflab polyprotein, from 5325-5925 a.a). Similar to SARS and MERS-Nsp13, the overall structure of SARS-CoV-2 nsp13 adopted a triangular pyramid shape comprising five domains. Among these, two "RecA-like" domains, 1A (261-441 a.a) and 2A (442-596 a.a), and 1B domain (150-260 a.a) forming the triangular base, while N-terminal Zinc binding domain (ZBD) (1-99 a.a) and stalk domain (100-149 a.a), which connects ZBD and 1B domain, are arranged at the apex of the pyramid **(Figure 1C)**. Small molecules able to inhibit the NTPase activity by interferences with ATP binding posed an ideal strategy to develop

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

inhibitors. The SARS-CoV-2 nsp13 revealed the similar conserved NTPase active site residues including Lys288, Ser289, Asp374, Glu375, Gln404 and Arg567 as present in SARS-Nsp13. All these residues were clustered together in the cleft located at the base between domain 1A and 2B **(Figure 1C)**, while docking grid was established by locating bound ADP of crystalized yeast Upf1 and top hits were identified.

Identification of potential compounds against nCoV proteins

A stepwise structure-based virtual screening pipeline was adopted considering the putative binding sites of SARS-CoV-2 Mpro, Nsp12 polymerase and Nsp13 helicase to identity potential hits as described in previous studies (Mirza, Noor-Ul-Huda Ghori et al. 2015, Mirza and Ikram 2016, Mirza, Vanmeert et al. 2019). The identified potential hits against each target were subjected to *in silico* ADMET predictions, which includes drug-likeness and toxicity potential. This filtering removed substantial hits that exhibited poor ADMET (absorption, distribution, metabolism, elimination, toxicity) due to inhibitory effects on the renal organic cation transporter and CYP450 isozymes. While a large subset of hits was found to contain high-risk chemical groups like epoxides and quinones. The resulted potential hits against each target having the lowest docking energy were analysed for their binding pose inside the binding site. After careful inspection, best hits based on AD Vina predicted binding energies were re-docked with Glide and best docked complexes were further processed through 20ns MD simulations, and MMGBSA method was employed to analyse electrostatic and vdW energy contribution towards total free energy of binding. The MD-simulated complexes that showed protein backbone stability (during the last 5ns in a 20 ns production run) were considered for further molecular interaction analysis. Based on molecular interactions and binding free energy calculations, the subsequent screening resulted in 5, 4 and 3 hits for SARS-CoV-2 Mpro, Nsp12 and Nsp13, respectively (Detailed ADMET profile of these compounds is tabulated in **Table S1,** estimated from Swiss-adme server (Daina, Michielin et al. 2017)**).** The AD Vina, Glide XP-score together with contribution of electrostatic (ΔE_{ele}), van der Waals (ΔE_{vdw}), and solute-solvent energies (ΔG_{np} and ΔG_{p}) towards total binding free energy (ΔG_{tol}) were estimated through Amber-MMGBSA module as tabulated in **Table 1**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

Table 1 Molecular docking predicted binding energies (kcal/mol) and Molecular mechanics generalized born surface area (MM-GBSA) binding free energy calculation of identified hits in complex with SARS-CoV-2 Mpro, Nsp12 polymerase and Nsp13 helicase.

Contributions		AD Vina	XP-Score	ΔE ele	ΔE vdw	ΔE MM	ΔG p	ΔG np	ΔG sol	ΔG tol
Mpro	**cmp3**	-9.2	-7.9	-21.38	-36.26	-57.64	33.04	-4.53	28.51	-29.13
	cmp12	-8.8	-9.87	-26.16	-55.63	-81.79	43.44	-6.87	36.57	-45.22
	cmp14	-8.7	-8.51	-33.56	-44.15	-77.71	38.93	-6.13	32.8	-44.91
	cmp17	-8.7	-8.27	-5.39	-53.87	-59.26	23.26	-5.72	17.54	-41.72
	cmp18	-8.6	-9.2	-36.65	-41.9	-78.55	38.44	-4.44	34	-44.55
Helicase	**cmp1**	-10.9	-9.91	-15.73	-55.49	-71.22	40.73	-6.8	33.93	-37.29
	cmp3a	-10.6	-11.1	-26.89	-51.08	-77.97	44.25	-6.36	37.89	-40.08
	cmp11	-10.2	-10.5	-31.17	-67.39	-98.56	56.76	-7.93	48.83	-49.73
	cmp15	-10.2	-8.4	-27.13	-63.39	-90.52	66.81	-6.85	59.96	-30.56
RdRp	**cmp2**	-8.8	-8.6	-58.59	-48.12	-106.71	71.28	-6.31	64.97	-41.74
	cmp17a	-8.4	-8.9	-49.44	-43.18	-92.62	65.14	-6.89	58.25	-34.37
	cmp21	-8.2	-7.1	-38.06	-47.73	-85.79	66.01	-5.12	60.89	-24.9

Note: ΔG_{tol} is the sum of molecular mechanics energy (ΔE_{MM}) and solvation free energy (ΔG_{sol}). Both ΔE_{MM} and ΔG_{sol} are further divided into internal energy (ΔE_{int}), electrostatic energy (ΔE_{ele}), and van der Waals (ΔE_{vdw}) energy in the gas phase, and polar (ΔG_{p}) and non-polar (ΔG_{np}) contributions to the solvation free energy.

SARS-CoV-2 Mpro hits

Top-ranked SARS-CoV-2 Mpro hits were filtered based on interaction with the catalytic dyad (at least one H-bond with either His41 or Cys145 or strong vdW interactions). A total of 13 compounds were analysed after 20 ns MD simulation based on the stability for at least last 5 ns and MMGBSA calculations were performed for the most stable complexes. Among top hits, total binding free energies (ΔG_{tol}) were very promising for cmp12, cmp14, cmp17 and cmp18 which exhibited ΔG_{tol} = -45.22, -44.91, -41.72, and -44.55 kcal/mol, and showed ligand RMSD values after 20ns less than 1.2 Å (0.91 Å, 1.18 Å, 0.84 Å and 1.03 Å) respectively. All compounds contributed well through H-bonds and exhibited favourable electrostatic energy for cmp3 (ΔG_{elec} = -21.38 kcal/mol), cmp12 (ΔG_{elec} = -26.16 kcal/mol), cmp14 (ΔG_{elec} = -33.56 kcal/mol) and cmp18 (ΔG_{elec} = -36.65 kcal/mol) except cmp17, which displayed weak electrostatic energy (ΔG_{elec} = -5.39 kcal/mol) due to unstable H-bonds over a simulation period **(Table 1)**. Further in-depth molecular interaction analysis after 20ns MD simulations unveiled significant findings with respect to Mpro-subsites **(Figure 2A)** and per-residue decomposition is displayed in **Figure 2B**. Among all compounds, cmp12 and 14 showed some consensus features, where pentacyclic moiety of cmp12 anchored right inside the subsites and formed H-bonds with Ser144 of S1, Cys145 of catalytic dyad, Met165 of S4, and terminal 4-methylaniline established two H-bonds with two neighbouring glutamine residues (Gln189 and Gln192) of S4. Similarly, biphenyl interacted hydrophobically with S2, triazole moiety established one H-bond with Cys145 of catalytic dyad, and pyrazole-pyrimidine moiety formed three H-bonds with Gln189, Thr190, and Gln192 located at S4. The remaining compounds (cmp3, cmp17 and cmp18) showed distinct H-bond patterns, even though the central moieties superimposed well near the catalytic dyad. For instance, the terminal oxadiazol-2-amine moiety of cmp3 mainly interacted through H-bonds with Cys145 of catalytic dyad, Asn142, Ser144 of S1, Glu166 and

 doi:10.20944/preprints202003.0085.v1

His172 of S4, and central carbaldehyde also formed H-bond with Glu166 of S4, Whereas, cmp17 only formed H-bonds with residues of S4 through terminal carbazole-1-carboxamide moiety and cmp18 established H-bonds with residues of S2 and S4 through urea and terminal imidazole moieties. Although the compounds mainly interacted electrostatically, however nonpolar solvation energies also slightly contributed towards ligand binding. The residual decomposition analysis revealed the favourable contribution of catalytic dyad towards total binding free energy, where Cys145 exhibited -1.058, -1.19, -2.12, -1.015, -0.892 kcal/mol and His41 exhibited -0.423, -1.528, -0.656, -1.032, -0.532 kcal/mol with cmp3, 12, 14, 17 and 18 respectively. Other residues located in close vicinity to catalytic dyad also contributed significantly which included, Glu189 (ranging from -3.505 to -1.219 kcal/mol), Glu166 (-2.711 to -0.845 kcal/mol), and Met165 (-3.859 to -0.988 kcal/mol) **(Figure 2B)**. Such favourable residual energy contribution was evident from the H-bonds formed by Glu166, Glu189 and Met165. The 2D chemical structure representation of each compound is displayed in connection with interactions of structural moieties with binding pocket subsites.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

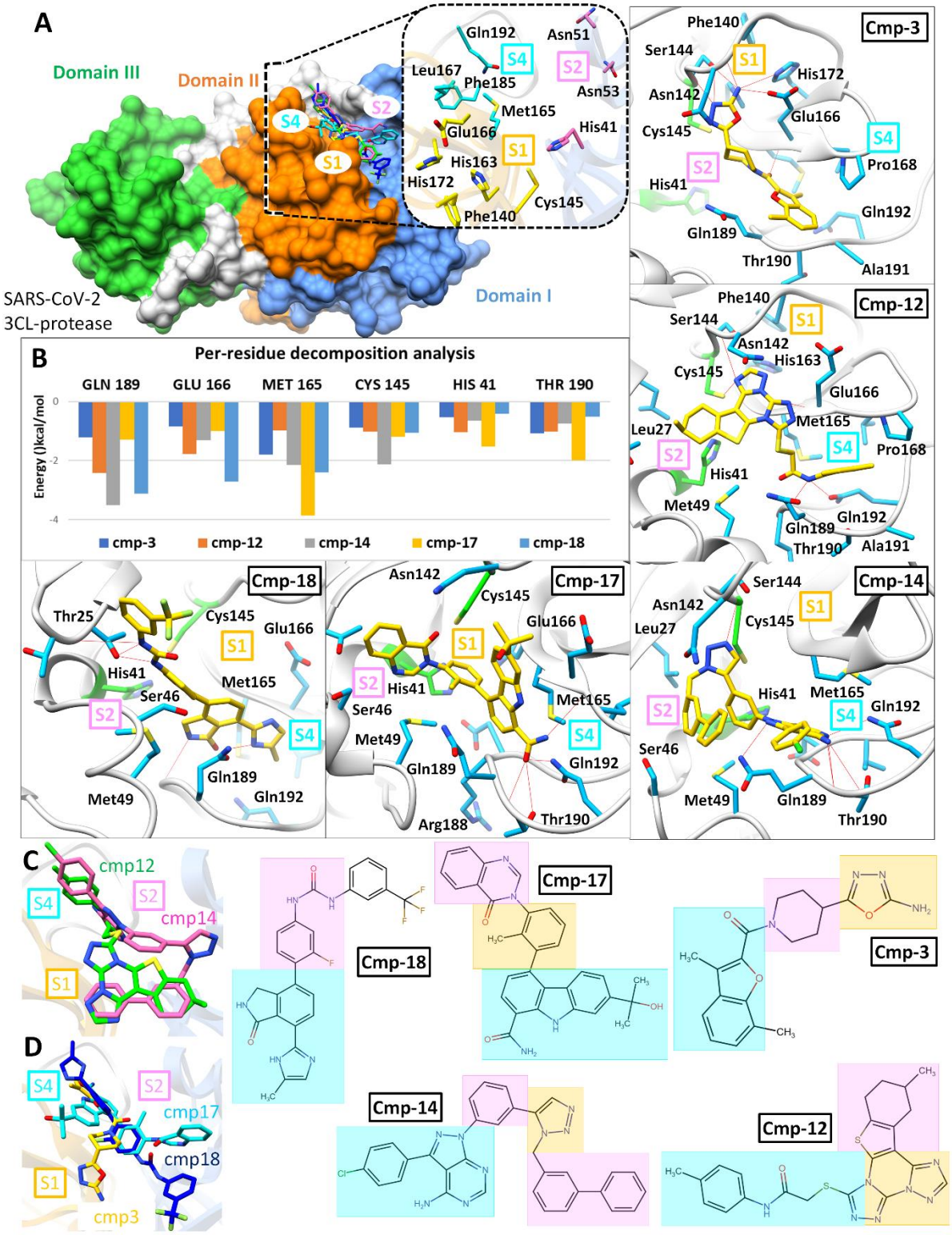

Figure 2: Post-Molecular dynamics (MD) analysis of SARS-CoV-2 Mpro hits. **(A)** Molecular surface representation of Mpro with MD simulated representative conformation of cmp3 (yellow), cmp12 (green), cmp14 (magenta), cmp17 (cyan) and cmp18 (dark blue) inside the substrate binding site zoomed with subsites S1 (orange), S2 (pink) and S4 (cyan) and residues are labelled accordingly. Molecular interactions representations of each complex with interacting residues are highlighted blue sticks and catalytic dyad in green sticks. Terminal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

moieties of cmp12 and cmp14 revealed similar binding modes while cmp3, cmp17 and cmp18 showed different binding mode are displayed in (**C**) and (**D)** followed by chemical structure representations of these compounds in connection with interactions of structural moieties with binding pocket subsites.

SARS-CoV-2 Nsp12 RdRp

For SARS-CoV-2 Nsp12 RdRp, hits were selected based on establishing H-bond with conserved motifs, especially with signature motif C (SDD), which has been experimentally evaluated in altering polymerase activity supported from various mutational studies (Biswas and Nayak 1994, Vázquez, Alonso et al. 2000, Bergeron, Albariño et al. 2010, Zhou, Zheng et al. 2011). A total of 17 hits were analysed through 20ns MD simulations, and complexes with stable RMSD were further processed through MMGBSA binding free energy calculations. Among all hits, cmp2, 17a and 21 showed promising ΔG_{tol} (-41.74, -34.37, -24.9 kcal/mol) which mainly interacted through H-bonds, thus contributed electrostatically **(Table 1)**. A careful molecular inspection after MD revealed the similar conformation of these compounds, which were found extensively interacted with the conserved motifs **(Figure 3)**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

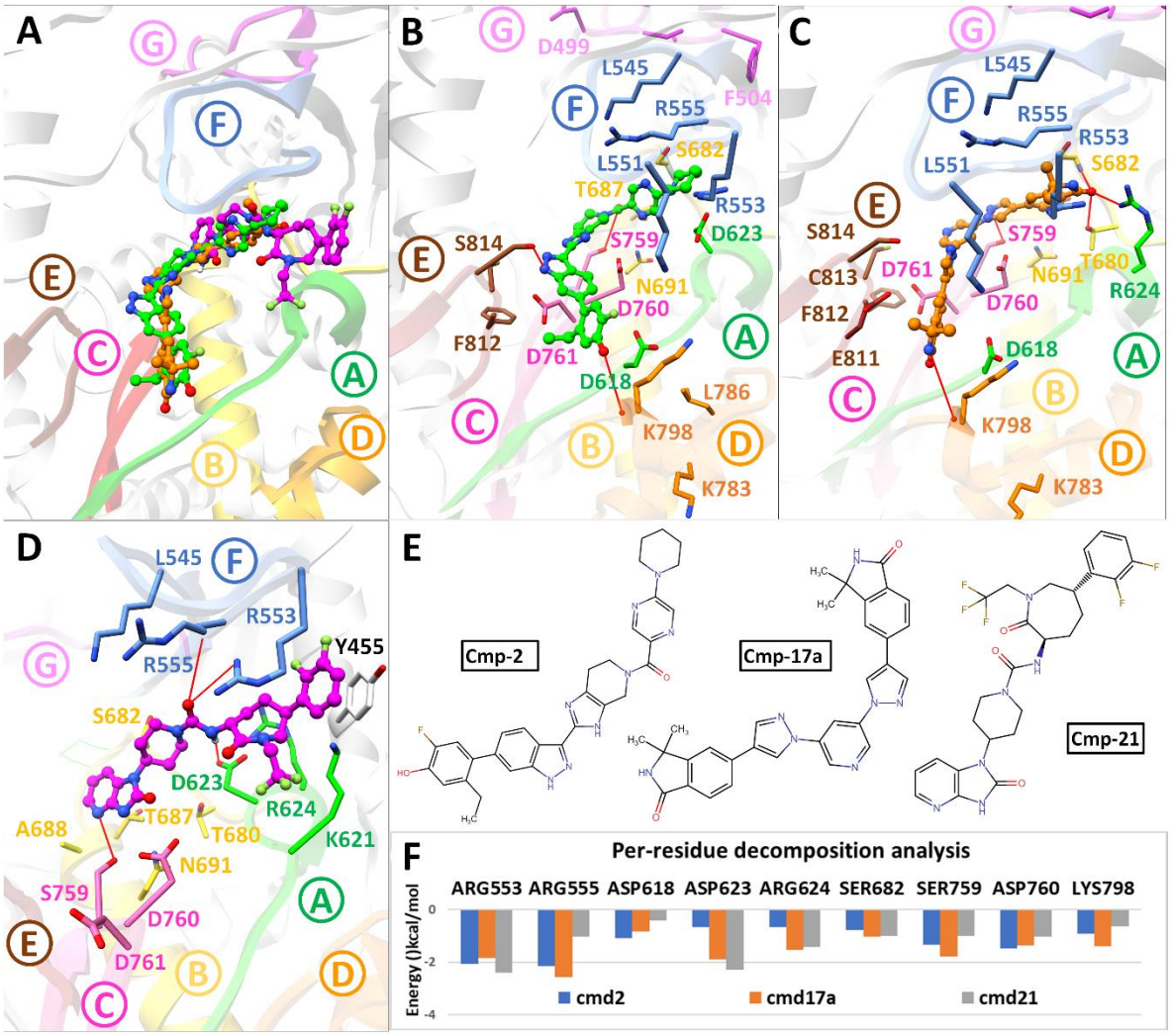

Figure 3: Post-Molecular dynamics (MD) analysis of SARS-CoV-2 Nsp12 RdRp complexes. **(A)** MD simulated conformations of cmp2 (green), cmp17a (green), cmp21 (magenta) inside the predicted binding pocket of Nsp12 polymerase. The arrangement of motifs and colors are same as in Fig. 1B. Molecular interactions of individual compounds are displayed in B and C and D, and residues are labelled accordingly. **(E)** Representations of chemical structures of compounds, and **(F)** Per-residue energy decomposition analysis of potential binding pocket residues

Cmp2 established three H-bonds with the residues belongs to signature motif C (Ser759), motif D (Lys798) and motif E (Ser814), Cmp17a formed five H-bonds with Arg624 (motif A), Thr680 and Ser682 (motif B), Ser756 (motif C) and Lys798 (motif D), whereas cmp21 interacted mainly with the residues of motif A (Asp623), motif F (Arg553 and Arg555), and motif C (Ser759) through four H-bonds. Together with, various residues also contributed through strong hydrophobic interactions, for instance, Tyr455 established strong stalking interaction with terminal difluorobenzene of cmp21. Per-residue binding energy decomposition analysis revealed a decisive role of important residues located in conserved polymerase motifs **(Figure 3F)**. These include the following residues; i) Arg553 and Arg555 of motif F which

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

exhibited most favourable ΔG_{tol} with cmp2 (-2.05 and -2.14 kcal/mol), 17a (-1.84 and -2.57 kcal/mol) and 21 (-2.4 and -1.04 kcal/mol), ii) conserved Aspartate residues of motif A (Asp618 and Asp623) and C (Asp760), iii) conserved Serines of motif B (Ser682) and motif C (Ser759) and, d) Lys798 of motif D. Among these, Ser759 and Asp760 of motif C exhibited significant binding free energy with all compounds (ranging from -0.988 to -1.76 kcal/mol and -1.03 to -1.47 kcal/mol), while most of the residues interacted electrostatically.

SARS-CoV-2 Nsp13 helicase

For SARS-CoV-2 Nsp13 helicase, top hits were selected based on significant interactions with six key residues (Lys288, Ser289, Asp374, Glu375, Gln404 and Arg567) including neighbouring residues involved in NTP hydrolysis (Jia, Yan et al. 2019). A 20 ns MD simulations on top ranked complexes identified 4 best compounds based on considerable interactions with key residues and displayed favourable MMGBSA binding free energy **(Table 1)**. The superimposition of Upf1-ADP showed the overall same binding mode as identified with these 4 compounds **(Figure 4)**.

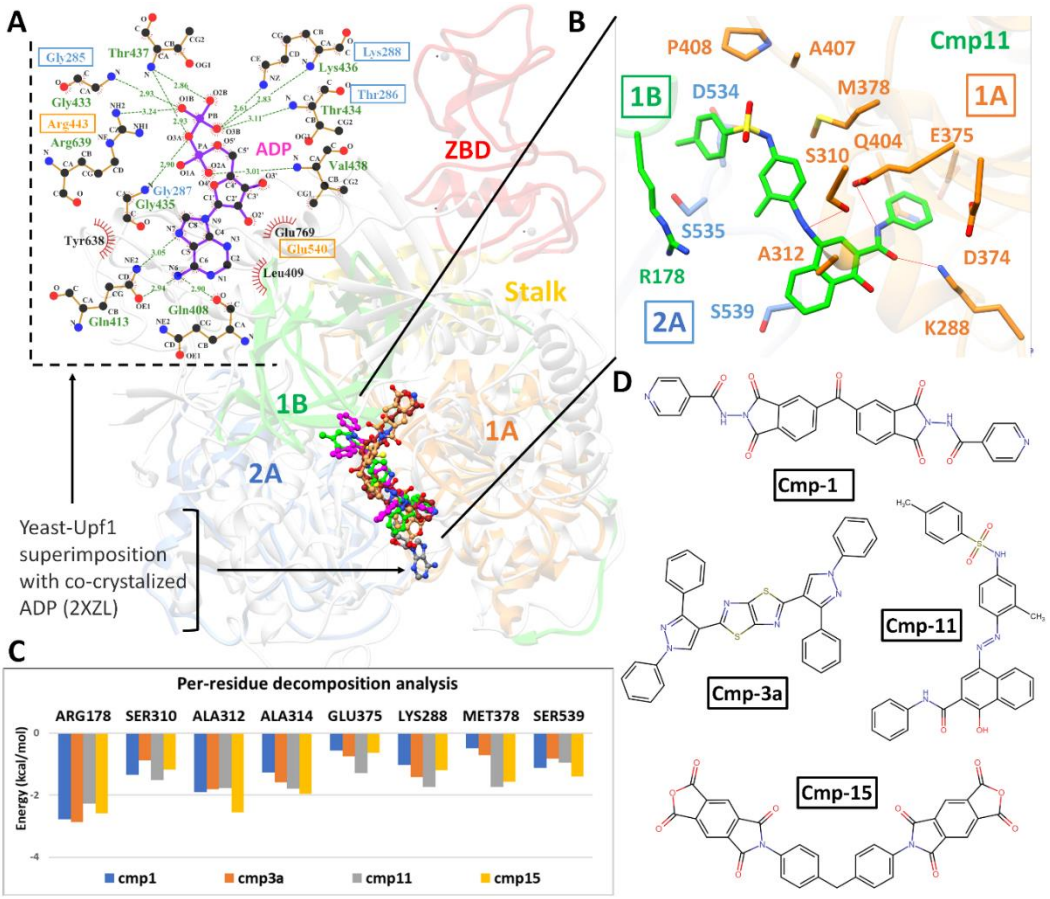

Figure 4: Post-Molecular dynamics (MD) analysis of SARS-CoV-2 Nsp13 helicase complexes. **(A)** The Yeast Upf1-ADP complex (transparent white) is superimposed on Nsp13 helicase to locate the ATP putative binding pocket and 2D-interaction plot is displayed. MD simulated cmp1 (brown), cmp3a (magenta), cmp11 (green) and cmp15 (sandy brown) are displayed inside the pocket. The arrangement of domains is colored same as in Fig. 1C. **(B)** MD simulated conformations of cmp11 (green) inside the predicted binding pocket of Nsp13

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

helicase interacted with key residues of domain 1A, 2A and 1B. **(C)** Per-residue energy decomposition analysis of potential binding pocket residues, and representations of chemical structures of compounds are displayed in **(D).**

The 2D-interaction plot of yeast Upf1-ADP revealed a network of H-bonds established by phosphate groups with Lys436, Thr434, Gly433, Arg639 Gly435 which were also conserved in nCoV-helicase **(Figure 4A)**. MMGBSA method estimated favourable ΔG_{tol} in cmp1 (-37.29 kcal/mol), cmp3a (-40.08 kcal/mol) and cmp11 (-49.73 kcal/mol) except cmp15 (-30.56 kcal/mol) and contributed mainly through vdW energies, while nonpolar solvation energies also slightly contributed **(Table 1)**. In-depth molecular analysis of cmp11 showed most significant MMGBSA total energy which indicated 3 H-bonds established between Ser310, Glu375 and Lys288 and terminal region of cmp11, and interacted mainly with the sidechains of Glu374, Met378, Ala312, Arg178, Ser535. Ser539 and Asp534 **(Figure 4B)**. These residues were also found conserved in Upf1-ADP complex except Met378 and Arg178. Per-residue energy decomposition analysis indicated major contributions from Arg178 of domain 1B, Ala312 and Ala314 of domain 1A which indicated < -1.5 kcal/mol binding energy with all four compounds. Together with, Ser288 and Ser310 also contributed in total free energy of binding **(Figure 4C)**.

Discussion

Human coronavirus (HCoV) causes to respiratory infection, while SARS-CoV and MERS-CoV are highly pathogenic viruses leading to severe respiratory syndrome in humans while other HCoVs including HCoV-OC43, HCoV-229E, HCoV-NL63 and HCoV-HKU1 cause mild upper respiratory disease. After SARS-CoV and MERS-CoV outbreak in 2002 and 2012 respectively, the emergence of SARS-CoV-2 has posed a global health threat (Wang, Horby et al. 2020). Recent literature has estimates of animal-to-human transmission (Riou and Althaus 2020), epidemic size (Oh, Choe et al. 2015) and improved characterization of epidemiological and virological features of the virus (Wu, Hao et al. 2020, Zhou, Yang et al. 2020). SARS-CoV-2 is an RNA virus with an inherent feature of high mutation rate however, it might adapt slow mutation rate in comparison to other RNA viruses due to the presence of genome-encoded exonuclease. However, this aspect leads to the possibility that SARS-CoV-2 might adapt to becoming more virulent alongside becoming more efficiently transmitted via person-to-person contact (Wang, Horby et al. 2020).

There exists no anti-viral therapy for human coronaviruses to date (Liu, Morse et al. 2020). Although, current treatment against COVID-19 includes the use of recombinant IFN with ribavirin which has limited effects (Cinatl, Morgenstern et al. 2003). Following SARS and MERS outbreak, efforts have been made for the development of antivirals specifically targeting CoV proteases, polymerases, MTases as well as entry proteins, however, none of these has shown efficacy during clinical trials (Chan, Chan et al. 2013, Cheng, Cheng et al. 2015, Wang, Sun et al. 2015). Moreover, strategies including the use of inactivated and live-attenuated viruses, vector-based vaccines, subunit vaccines, recombinant proteins and DNA vaccines, have so far been evaluated in animals only (Graham, Donaldson et al. 2013, Wang, Shi et al.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 March 2020 doi:10.20944/preprints202003.0085.v1

2015, de Wit, van Doremalen et al. 2016). Several virological and patient-associated factors pose a significant challenge in the development of novel anti-CoV drugs. One such challenge is the high mutation rate of RNA-based HCoVs. Since, novel CoVs emerge at unpredictable times owing to their rapid mutation rates, therefore, most anti-CoV drugs designed to target the replication apparatus of an existing CoV specifically may not be effective against novel CoV(Li and De Clercq 2020).

The computational methods offer an immediate and scientifically sound basis to design highly specific inhibitor against important viral proteins and give aid in antiviral drug discovery (Jorgensen 2004, Bajorath and Bajorath 2011). A recent structural modelling effort predicted few repurposed drugs against SARS-CoV-2 vital proteins including anti-bacterial, anti-fungal, anti-HIV drugs, and flavonoids from different sources. In the current study, a brief structural elucidation was performed together which revealed strikingly similar features in SARS-CoV-2, when compared with crystal structures of SARS-CoV-Mpro, RdRp and helicase. These viral proteins play a crucial role in the viral life cycle and considered among the most popular strategies for antiviral therapeutics (Sampath and Padmanabhan 2009, Hoenen, Groseth et al. 2019). Structural analyses revealed that the substrate-binding pockets of SARS-CoV-Mpro, active site of RdRp and NTPase binding pocket of helicase were highly conserved (Liu, Morse et al. 2020), which can lead to the concept of "wide-spectrum inhibitors" for targeting CoVs. Furthermore, extensive SBVS procedure identified potential hits against each target, which were assessed through MD simulations. MM-GBSA and per residue energy decomposition divulged the relative importance of amino acid involved in binding supported with the contribution of different components of binding free energy.

Overall, nCoV-Mpro hits were found to interact with catalytic dyad (His41 or Cys145) together with residues located at S4, S2 and S1 subsites of substrate binding pocket and mutational studies of His41 or Cys145 resulted in complete loss of enzymatic activity, which confirmed the role of catalytic dyad (Huang, Wei et al. 2004, Tan, Verschueren et al. 2005). In the substrate-binding pocket, catalytic dyad residues, highly conserved residues of S1, Gln189 and Met165 of S4, displayed relative importance in ligand binding, suggested a possibility for inhibitor design targeting this region in the nCoV-Mpro. In case of nCoV-RdRP, which closely resembles with the polymerases of picornaviruses (Lu, Stratton et al. , Zamyatkin, Parra et al. 2009, Gong and Peersen 2010), the in-depth interaction analysis inside the conserved motifs **(Figure 3)** revealed functionally important aspartate residues of motif A (Asp623) and motif C (Asp760), together with conserved Arginine residues of motif F (Arg553 and Arg555), and Ser759 (motif C) which were found highly interacted with the compounds. Among these, motif A and C are most strikingly conserved aspartate residues that bind divalent metal ions necessary for catalytic activity (Poch, Sauvaget et al. 1989, Gong and Peersen 2010). Moreover, the importance of aspartate residues in RdRps supported from various other mutational studies (Biswas and Nayak 1994, Vázquez, Alonso et al. 2000, Bergeron, Albariño et al. 2010, Zhou, Zheng et al. 2011), suggested a conserved binding feature of cmp2, cmp17a and cmp21 along with side chains of Arg553, Arg555 and Ser759 which line deep in the binding pocket of nCoV-RdRp, thus enhancing it as an anchor for inhibitory molecules.

In nCoV-helicase, the structural information of NTPase active site obtained after superimposing SARS-helicase and Yeast-Upf1-ADP divulged the location of conserved residues. The identified hits positioned similarly with reference to the interaction network reported for co-crystalized ADP **(Figure 4)**. In the active site, the side chains of Ser310, Lys288 and Glu375 actively participated in H-bond, while Arg178 and conserved alanine residues (Ala312 and Ala314) contributed substantially towards total binding free energy. Moreover, the importance of these residues has already been elucidated from mutational studies (Jia, Yan et al. 2019). Together with, the ZBD domain also proven critical for helicase activity or even the life cycle of SARS-CoV (Jia, Yan et al. 2019).

Among various hits identified, cmp2 against nCoV-RdRp was recently reported as Pan-Janus Kinase Inhibitor Clinical Candidate (PF-06263276) for its potential role in the treatment of inflammatory diseases of lungs (Jones, Storer et al. 2017). Hence, the cmp2 could be ideal for reducing the inflammatory activity during pathogenic COVID-19 epidemic (Channappanavar and Perlman 2020), as lung inflammation was observed during the SARS and MERS outbreaks (Guarner 2020).

Conclusively, this comprehensive study presented an integrated computational approach towards the identification of novel inhibitors into the area of anti-nCoV treatment as a starting point, which warrants in vitro testing to evaluate compound efficacy. The in-depth structural elucidation of interacting residues together with the dynamic conformations adopted over a period of 20ns MD simulations of identified compounds offered the way to design wide-spectrum or selective inhibitors against HCoV.

The development of novel, broad-spectrum, pan-CoV antivirals seems to be the ultimate treatment strategy for emerging CoV infections. Future *in vivo* and clinical studies focusing on assessing the effectiveness and safety of promising antiviral drugs, monoclonal and polyclonal neutralising antibody products, and therapeutics directed against immunopathologic host responses may provide cues for treating emerging COVID19 outbreak.

References

Azhar, E. I., D. S. Hui, Z. A. Memish, C. Drosten and A. J. I. D. C. Zumla (2019). "The Middle East Respiratory Syndrome (MERS)." **33**(4): 891-905.

Bacha, U., J. Barrila, A. Velazquez-Campoy, S. A. Leavitt and E. J. B. Freire (2004). "Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro." **43**(17): 4906-4912.

Bajorath, J. and J. Bajorath (2011). Chemoinformatics and computational chemical biology, Springer.

Benvenuto, D., M. Giovanetti, A. Ciccozzi, S. Spoto, S. Angeletti and M. J. b. Ciccozzi (2020). "The 2019-new Coronavirus epidemic: evidence for virus evolution."

Bergeron, É., C. G. Albariño, M. L. Khristova and S. T. J. J. o. v. Nichol (2010). "Crimean-Congo hemorrhagic fever virus-encoded ovarian tumor protease activity is dispensable for virus RNA polymerase function." **84**(1): 216-226.

Bianco, A. C., D. Salvatore, B. Gereben, M. J. Berry and P. R. J. E. r. Larsen (2002). "Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases." **23**(1): 38-89.

Biswas, S. K. and D. P. J. J. o. v. Nayak (1994). "Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1." **68**(3): 1819-1826.

Blanchard, J. E., N. H. Elowe, C. Huitema, P. D. Fortin, J. D. Cechetto, L. D. Eltis, E. D. J. C. Brown and biology (2004). "High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase." **11**(10): 1445-1453.

Bruenn, J. A. J. N. a. r. (2003). "A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases." **31**(7): 1821-1829.
Chan, J. F.-W., K.-H. Kok, Z. Zhu, H. Chu, K. K.-W. To, S. Yuan, K.-Y. J. E. M. Yuen and Infections (2020). "Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan." **9**(1): 221-236.
Chan, J. F., K.-H. Chan, R. Y. Kao, K. K. To, B.-J. Zheng, C. P. Li, P. T. Li, J. Dai, F. K. Mok and H. J. J. o. I. Chen (2013). "Broad-spectrum antivirals for the emerging Middle East respiratory syndrome coronavirus." **67**(6): 606-616.
Chang, H.-P., C.-Y. Chou and G.-G. J. B. j. Chang (2007). "Reversible unfolding of the severe acute respiratory syndrome coronavirus main protease in guanidinium chloride." **92**(4): 1374-1383.
Channappanavar, R. and S. Perlman (2020). Evaluation of Activation and Inflammatory Activity of Myeloid Cells During Pathogenic Human Coronavirus Infection. MERS Coronavirus: Methods and Protocols. R. Vijay. New York, NY, Springer US**:** 195-204.
Chen, Y., Q. Liu and D. J. J. o. M. V. Guo (2020). "Coronaviruses: genome structure, replication, and pathogenesis."
Cheng, K.-W., S.-C. Cheng, W.-Y. Chen, M.-H. Lin, S.-J. Chuang, I.-H. Cheng, C.-Y. Sun and C.-Y. J. A. r. Chou (2015). "Thiopurine analogs and mycophenolic acid synergistically inhibit the papain-like protease of Middle East respiratory syndrome coronavirus." **115**: 9-16.
Cinatl, J., B. Morgenstern, G. Bauer, P. Chandra, H. Rabenau and H. J. T. L. Doerr (2003). "Treatment of SARS with human interferons." **362**(9380): 293-294.
Daina, A., O. Michielin and V. J. S. r. Zoete (2017). "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules." **7**: 42717.
De Clercq, E. J. E. r. o. a.-i. t. (2006). "Potential antivirals and antiviral strategies against SARS coronavirus infections." **4**(2): 291-302.
de Wit, E., N. van Doremalen, D. Falzarano and V. J. J. N. R. M. Munster (2016). "SARS and MERS: recent insights into emerging coronaviruses." **14**(8): 523.
Drosten, C., S. Günther, W. Preiser, S. Van Der Werf, H.-R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova and R. A. J. N. E. j. o. m. Fouchier (2003). "Identification of a novel coronavirus in patients with severe acute respiratory syndrome." **348**(20): 1967-1976.
Gong, P. and O. B. J. P. o. t. N. A. o. S. Peersen (2010). "Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase." **107**(52): 22505-22510.
Graham, R. L., E. F. Donaldson and R. S. J. N. R. M. Baric (2013). "A decade after SARS: strategies for controlling emerging coronaviruses." **11**(12): 836-848.
Gralinski, L. E. and V. D. J. V. Menachery (2020). "Return of the Coronavirus: 2019-nCoV." **12**(2): 135.
Guarner, J. J. A. J. o. C. P. (2020). "Three Emerging Coronaviruses in Two DecadesThe Story of SARS, MERS, and Now COVID-19."
Hoenen, T., A. Groseth and H. J. N. R. M. Feldmann (2019). "Therapeutic strategies to target the Ebola virus life cycle." **17**(10): 593-606.
Holshue, M. L., C. DeBolt, S. Lindquist, K. H. Lofy, J. Wiesman, H. Bruce, C. Spitters, K. Ericson, S. Wilkerson and A. J. N. E. J. o. M. Tural (2020). "First case of 2019 novel coronavirus in the United States."
Hou, T., J. Wang, Y. Li, W. J. J. o. c. i. Wang and modeling (2011). "Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations." **51**(1): 69-82.
Huang, C., Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu and X. J. T. L. Gu (2020). "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China."
Huang, C., P. Wei, K. Fan, Y. Liu and L. J. B. Lai (2004). "3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism." **43**(15): 4568-4574.
Hui, D. S., T. Madani, F. Ntoumi, R. Kock, O. Dar, G. Ippolito, T. Mchugh, Z. Memish, C. Drosten and A. J. I. j. o. i. d. I. o. p. o. t. I. S. f. I. D. Zumla (2020). "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-The latest 2019 novel coronavirus outbreak in Wuhan, China." **91**: 264.

Ikram, N., M. U. Mirza, M. Vanmeert, M. Froeyen, O. M. Salo-Ahen, M. Tahir, A. Qazi and S. J. B. Ahmad (2019). "Inhibition of Oncogenic Kinases: An In Vitro Validated Computational Approach Identified Potential Multi-Target Anticancer Compounds." **9**(4): 124.
Irwin, J. J., B. K. J. J. o. c. i. Shoichet and modeling (2005). "ZINC− a free database of commercially available compounds for virtual screening." **45**(1): 177-182.
Jain, R. P., H. I. Pettersson, J. Zhang, K. D. Aull, P. D. Fortin, C. Huitema, L. D. Eltis, J. C. Parrish, M. N. James and D. S. J. J. o. m. c. Wishart (2004). "Synthesis and evaluation of keto-glutamine analogues as potent inhibitors of severe acute respiratory syndrome 3CLpro." **47**(25): 6113-6116.
Jia, Z., L. Yan, Z. Ren, L. Wu, J. Wang, J. Guo, L. Zheng, Z. Ming, L. Zhang and Z. J. N. a. r. Lou (2019). "Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis." **47**(12): 6538-6550.
Jones, P., R. I. Storer, Y. A. Sabnis, F. M. Wakenhut, G. A. Whitlock, K. S. England, T. Mukaiyama, C. M. Dehnhardt, J. W. Coe and S. W. J. J. o. m. c. Kortum (2017). "Design and synthesis of a pan-Janus kinase inhibitor clinical candidate (PF-06263276) suitable for inhaled and topical delivery for the treatment of inflammatory diseases of the lungs and skin." **60**(2): 767-786.
Jorgensen, W. L. J. S. (2004). "The many roles of computation in drug discovery." **303**(5665): 1813-1818.
Kao, R. Y., W. H. Tsui, T. S. Lee, J. A. Tanner, R. M. Watt, J.-D. Huang, L. Hu, G. Chen, Z. Chen, L. J. C. Zhang and biology (2004). "Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics." **11**(9): 1293-1299.
Kasmi, Y., K. Khataby, A. Souiri and M. M. Ennaji (2020). Coronaviridae: 100,000 Years of Emergence and Reemergence. Emerging and Reemerging Viral Pathogens, Elsevier**:** 127-149.
Kirchdoerfer, R. N. and A. B. J. N. c. Ward (2019). "Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors." **10**(1): 1-9.
Kiss, R., M. Sandor and F. A. J. J. o. c. Szalai (2012). "http://Mcule. com: a public web service for drug discovery." **4**(S1): P17.
Kuiken, T., R. A. Fouchier, M. Schutten, G. F. Rimmelzwaan, G. Van Amerongen, D. van Riel, J. D. Laman, T. de Jong, G. van Doornum and W. J. T. L. Lim (2003). "Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome." **362**(9380): 263-270.
Li, G. and E. De Clercq (2020). Therapeutic options for the 2019 novel coronavirus (2019-nCoV), Nature Publishing Group.
Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung and Z. Feng (2020). "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia."
Liu, W., J. S. Morse, T. Lalonde and S. J. C. Xu (2020). "Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV."
Lu, H., C. W. Stratton and Y. W. J. J. o. M. V. Tang "Outbreak of Pneumonia of Unknown Etiology in Wuhan China: the Mystery and the Miracle."
Maier, J. A., C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. J. J. o. c. t. Simmerling and computation (2015). "ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB." **11**(8): 3696-3713.
Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber and S. Y. J. S. Chan (2003). "The genome sequence of the SARS-associated coronavirus." **300**(5624): 1399-1404.
Mirza, M. U. and N. J. I. j. o. m. s. Ikram (2016). "Integrated computational approach for virtual hit identification against ebola viral proteins VP35 and VP40." **17**(11): 1748.
Mirza, M. U., N. I. Noor-Ul-Huda Ghori, A. R. Adil, S. J. D. d. Manzoor, development and therapy (2015). "Pharmacoinformatics approach for investigation of alternative potential hepatitis C virus nonstructural protein 5B inhibitors." **9**: 1825.

Mirza, M. U., S. Rafique, A. Ali, M. Munir, N. Ikram, A. Manan, O. M. Salo-Ahen and M. J. S. r. Idrees (2016). "Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins." **6**: 37313.
Mirza, M. U., M. Vanmeert, M. Froeyen, A. Ali, S. Rafique and M. J. S. r. Idrees (2019). "In silico structural elucidation of RNA-dependent RNA polymerase towards the identification of potential Crimean-Congo Hemorrhagic Fever Virus inhibitors." **9**(1): 1-18.
Morse, J. S., T. Lalonde, S. Xu and W. Liu (2020). "Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV."
Oh, M.-d., P. G. Choe, H. S. Oh, W. B. Park, S.-M. Lee, J. Park, S. K. Lee, J.-S. Song and N. J. J. J. o. K. M. S. Kim (2015). "Middle East respiratory syndrome coronavirus superspreading event involving 81 persons, Korea 2015." **30**(11): 1701-1705.
Parry, J. (2020). China coronavirus: cases surge as official admits human to human transmission, British Medical Journal Publishing Group.
Paules, C. I., H. D. Marston and A. S. J. J. Fauci "Coronavirus Infections—More Than Just the Common Cold."
Peersen, O. B. J. V. r. (2017). "Picornaviral polymerase structure, function, and fidelity modulation." **234**: 4-20.
Peiris, J., S. Lai, L. Poon, Y. Guan, L. Yam, W. Lim, J. Nicholls, W. Yee, W. Yan and M. J. T. L. Cheung (2003). "Coronavirus as a possible cause of severe acute respiratory syndrome." **361**(9366): 1319-1325.
Perlman, S. (2020). "Another Decade, Another Coronavirus."
Poch, O., I. Sauvaget, M. Delarue and N. J. T. E. j. Tordo (1989). "Identification of four conserved motifs among the RNA-dependent polymerase encoding elements." **8**(12): 3867-3874.
Riou, J. and C. L. J. E. Althaus (2020). "Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020." **25**(4).
Roe, D. R., T. E. J. J. o. c. t. Cheatham III and computation (2013). "PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data." **9**(7): 3084-3095.
Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher and M.-h. J. s. Chen (2003). "Characterization of a novel coronavirus associated with severe acute respiratory syndrome." **300**(5624): 1394-1399.
Sampath, A. and R. J. A. r. Padmanabhan (2009). "Molecular targets for flavivirus drug discovery." **81**(1): 6-15.
Schwede, T., J. Kopp, N. Guex and M. C. J. N. a. r. Peitsch (2003). "SWISS-MODEL: an automated protein homology-modeling server." **31**(13): 3381-3385.
Srivastava, H. K., G. N. J. J. o. c. i. Sastry and modeling (2012). "Molecular dynamics investigation on a series of HIV protease inhibitors: assessing the performance of MM-PBSA and MM-GBSA approaches." **52**(11): 3088-3098.
Surana, P., V. Satchidanandam and D. T. Nair (2013). "RNA-dependent RNA polymerase of Japanese encephalitis virus binds the initiator nucleotide GTP to form a mechanistically important pre-initiation state." Nucleic Acids Research **42**(4): 2758-2773.
Tan, J., K. H. Verschueren, K. Anand, J. Shen, M. Yang, Y. Xu, Z. Rao, J. Bigalke, B. Heisen and J. R. J. J. o. m. b. Mesters (2005). "pH-dependent conformational flexibility of the SARS-CoV main proteinase (Mpro) dimer: molecular dynamics simulations and multiple X-ray structure analyses." **354**(1): 25-40.
Tan, J. J., W. Zu Chen and C. X. J. J. o. M. S. T. Wang (2006). "Investigating interactions between HIV-1 gp41 and inhibitors by molecular dynamics simulation and MM–PBSA/GBSA calculations." **766**(2-3): 77-82.
Tanner, J. A., B.-J. Zheng, J. Zhou, R. M. Watt, J.-Q. Jiang, K.-L. Wong, Y.-P. Lin, L.-Y. Lu, M.-L. He, H.-F. J. C. Kung and biology (2005). "The adamantane-derived bananins are potent inhibitors of the helicase activities and replication of SARS coronavirus." **12**(3): 303-311.
Trott, O. and A. J. J. J. o. c. c. Olson (2010). "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading." **31**(2): 455-461.

Vázquez, A. L., J. M. M. Alonso and F. J. J. o. v. Parra (2000). "Mutation analysis of the GDD sequence motif of a calicivirus RNA-dependent RNA polymerase." **74**(8): 3888-3891.
Wang, C., P. W. Horby, F. G. Hayden and G. F. J. T. L. Gao (2020). "A novel coronavirus outbreak of global health concern."
Wang, L., W. Shi, M. G. Joyce, K. Modjarrad, Y. Zhang, K. Leung, C. R. Lees, T. Zhou, H. M. Yassine and M. J. N. c. Kanekiyo (2015). "Evaluation of candidate vaccine approaches for MERS-CoV." **6**(1): 1-11.
Wang, M., R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong and G. J. C. R. Xiao (2020). "Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro." 1-3.
Wang, Y., Y. Sun, A. Wu, S. Xu, R. Pan, C. Zeng, X. Jin, X. Ge, Z. Shi and T. J. J. o. v. Ahola (2015). "Coronavirus nsp10/nsp16 methyltransferase can be targeted by nsp10-derived peptide in vitro and in vivo to reduce replication and pathogenesis." **89**(16): 8416-8427.
WHO (2020). "Situation report - 40 (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200229-sitrep-40-covid-19.pdf?sfvrsn=7203e653_2)."
Woo, P. C., S. K. Lau, C.-m. Chu, K.-h. Chan, H.-w. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai and W.-k. J. J. o. v. Luk (2005). "Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia." **79**(2): 884-895.
Wu, C.-Y., J.-T. Jan, S.-H. Ma, C.-J. Kuo, H.-F. Juan, Y.-S. E. Cheng, H.-H. Hsu, H.-C. Huang, D. Wu and A. J. P. o. t. N. A. o. S. Brik (2004). "Small molecules targeting severe acute respiratory syndrome human coronavirus." **101**(27): 10012-10017.
Wu, F., Zhao,S., Yu,B., Chen,Y.-M., Wang,W., Hu,Y., Song,Z.-G., Tao,Z.-W., Tian,J.-H., Pei,Y.-Y., Yuan,M.L., Zhang,Y.-L., Dai,F.-H., Liu,Y., Wang,Q.-M., Zheng,J.-J., Xu,L., Holmes,E.C. and and Y.-Z. Zhang (2020). "https://www.ncbi.nlm.nih.gov/nuccore/MN908947."
Wu, P., X. Hao, E. H. Lau, J. Y. Wong, K. S. Leung, J. T. Wu, B. J. Cowling and G. M. J. E. Leung (2020). "Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020." **25**(3).
Xu, J., S. Zhao, T. Teng, A. E. Abdalla, W. Zhu, L. Xie, Y. Wang and X. J. V. Guo (2020). "Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV." **12**(2): 244.
Yang, H., W. Xie, X. Xue, K. Yang, J. Ma, W. Liang, Q. Zhao, Z. Zhou, D. Pei and J. J. P. b. Ziebuhr (2005). "Design of wide-spectrum inhibitors targeting coronavirus main proteases." **3**(10).
Yang, J., A. Roy and Y. J. B. Zhang (2013). "Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment." **29**(20): 2588-2595.
Yin, Y. and R. G. J. R. Wunderink (2018). "MERS, SARS and other coronaviruses as causes of pneumonia." **23**(2): 130-137.
Zamyatkin, D. F., F. Parra, Á. Machín, P. Grochulski and K. K.-S. J. J. o. m. b. Ng (2009). "Binding of 2′-amino-2′-deoxycytidine-5′-triphosphate to norovirus polymerase induces rearrangement of the active site." **390**(1): 10-16.
Zhou, P., X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li and C.-L. J. N. Huang (2020). "A pneumonia outbreak associated with a new coronavirus of probable bat origin." 1-4.
Zhou, Y., H. Zheng, F. Gao, D. Tian and S. J. S. C. L. S. Yuan (2011). "Mutational analysis of the SDD sequence motif of a PRRSV RNA-dependent RNA polymerase." **54**(9): 870-879.
Zhu, N., D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi and R. J. N. E. J. o. M. Lu (2020). "A Novel Coronavirus from Patients with Pneumonia in China, 2019."