Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Molecular Structure 1231 (2021) 129953

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

Probing CAS database as prospective antiviral agents against SARS-CoV-2 main protease

Komal Zia[a], Salman Ali Khan[a], Sajda Ashraf[a], Mohammad Nur-e-Alam[b], Sarfaraz Ahmed[b], Zaheer Ul-Haq[a,*]

[a] Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan

[b] Department of Pharmacognosy, College of Pharmacy, King Saud University, P.O. Box. 2457, Riyadh 11451, Kingdom of Saudi Arabia

ARTICLE INFO

Article history:
Received 19 September 2020
Revised 11 January 2021
Accepted 12 January 2021
Available online 19 January 2021

Keywords:
SARS-CoV-2
Consensus scoring
COVID-19
CAS database
MD simulation

ABSTRACT

The pandemic of COVID-19 has an unprecedented impact on global health and economy. The novel SARS-CoV-2 is recognized as the etiological agent of current outbreak. Because of its contagious human-to-human transmission, it is an utmost global health emergency at present. To mitigate this threat many scientists and researchers are racing to develop antiviral therapy against the virus. Unfortunately, to date no vaccine or antiviral therapeutic is approved thus there is an urgent need to discover antiviral agent to help the individual who are at high risk. Virus main protease or chymotrypsin-like protease plays a pivotal role in virus replication and transcription; thus, it is considered as an attractive drug target to combat the COVID-19. In this study, multistep structure based virtual screening of CAS antiviral database is performed for the identification of potent and effective small molecule inhibitors against chymotrypsin-like protease of SARS-CoV-2. Consensus scoring strategy combine with flexible docking is used to extract potential hits. As a result of extensive virtual screening, 4 hits were shortlisted for MD simulation to study their stability and dynamic behavior. Insight binding modes demonstrated that the selected hits stabilized inside the binding pocket of the target protein and exhibit complementarity with the active site residues. Our study provides compounds for further in vitro and in vivo studies against SARS-CoV-2.

1. Introduction

The world is witnessing the deadliest pandemic of current times in the form of COVID19, first identified in Wuhan, China dates back to December 2019 [1]. SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) is identified as the etiological agent of COVID-19 [2].The mysterious COVID-19 emerged in china, growing exponentially, crossing the territorial barriers of modern world and now penetrated in approximately ~196 countries [3]. Due to the high infectivity of SARS-CoV-2 and the capacity of spreading abruptly on an international scale, the WHO on 11 March 2020 declared it as a pandemic situation [4].

Over the last 2 decades, the world witnesses the two beta-CoVs pandemics, including the SARS-CoV responsible for severe acute respiratory syndrome (SARS) in 2003 and MERS-CoV responsible for Middle Eastern respiratory syndrome (MERS) in 2012 [5]. Currently emerged another novel Coronavirus called SARS-CoV-2, responsible for the ongoing outbreak. The zoonotic Coronaviruses are naturally circulating in different animal species like bats, birds, cats, pigs, mice and cause several respiratory and intestinal infections in humans. Bat was the common primary host of MERS-CoV, and SARS-CoV. However, the direct zoonotic source of SARS-CoV-2 remains enigmatic [6]. Although the fatality rate of SARS-CoV-2 is lower than MERS and SARS but its production rate is higher [7]. However, the exact fatality and production rate for COVID-19 is still under debate. The spreading pandemic of COVID-19 has unusual impact on people and economy of the world due to isolation, social distancing, quarantine and community containment, which is implemented in many countries to reduce the risk of the infection [8,9,10].

Coronaviruses are a large family of lipid-enveloped viruses containing largest known single-stranded and positive polarity multipartite RNA genomes ranging from 27 kb to 31.5 kb in size [11]. The taxonomist of *Coronaviridae* Study Group (CSG) place the SARS-CoV-2 in class of *Betacoronavirus* within the *Coronaviridae* family [12]. Typically, its genome contains 10 open reading frames (ORFs). The two thirds of the whole SARS-CoV-2 genome harbors a large open reading frame ORF1 at the 5-terminal. The ORF1 acts as messenger RNA (mRNA) and directly translated into ORF1a and ORF1b

* Corresponding author.
E-mail address: zaheer.qasmi@iccs.edu (Z. Ul-Haq).

https://doi.org/10.1016/j.molstruc.2021.129953
0022-2860/

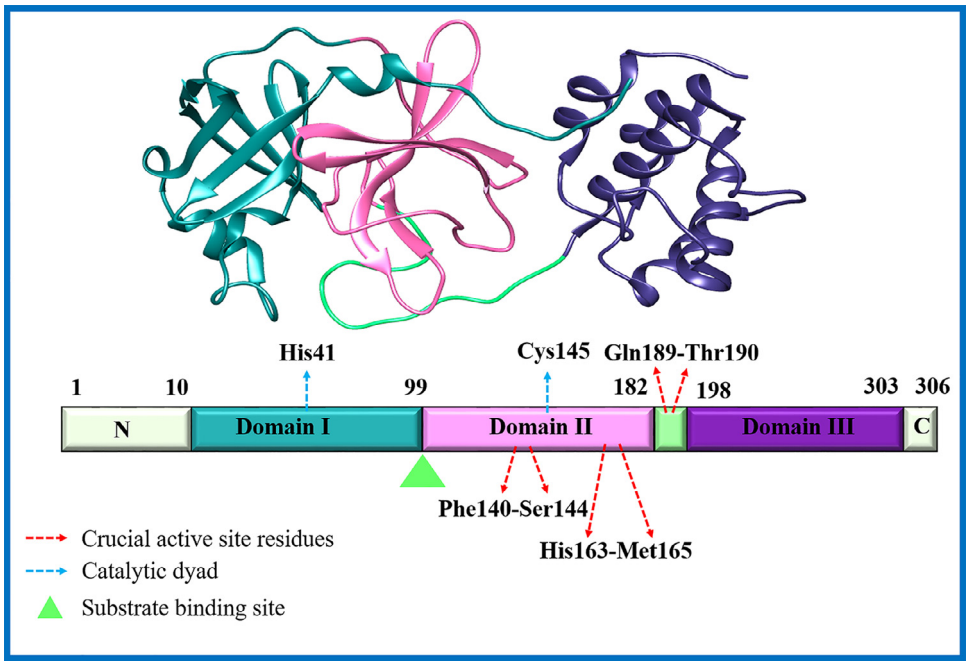

Fig. 1. Structural details of 3CLPro of SARS-CoV-2. Domain organization with role of amino acid residues and 3D structure of main protease. Inter domain boundaries are labelled with amino acid residues. Domain-I shown in cyan color, Domain-II shown in hot pink color while Domain-III shown in dark purple color.

by the host cell ribosomal frameshifting machinery. The remaining ORFs (2–10) at 3′-terminal of SARS-CoV-2 genome encode eight accessory proteins (3a, 3b, p6, 7a, 7b, 8b, 9b and 10) and four important structural proteins [spike (S), nucleocapsid (N), envelope (E) and membrane (M) protein] [13].

During SARS-CoV-2 infection in the host targeted cells, two overlapping "polyproteins," pp1a (486 kDa) and pp1ab (790 kDa) translated from the genomic ORF1ab region of virus, and cleaved into different non-structural proteins [14]. 3C-like protease (3CL^{pro}, also referred as the main protease, M^{pro}) and PL2^{pro} protease are also present among these polyproteins. PL2^{pro} cleaves at three sites while 3CLpro process the large viral polypeptide at 11 sites producing functional nonstructural proteins (nsp1–16) including a helicase, a single-stranded RNA-binding protein, an RNA-dependent RNA polymerase, an endoribonuclease, an exoribonuclease and a 2′-O-ribose-methyltransferase which are important for the coronavirus replication and packaging [15]. Therefore M-pro is considered as a high-profile drug target in anti-SARS-CoV-2 drug discovery [16,17].

The main protease of SARS-CoV-2 exhibit high structural and sequence homology to 3CLpro of SARS-CoV (96% sequence identity) and their sequence alignment indicate only 12 mutated/variant residues (T35V, A46S, S65N, L86V, R88K, S94A, H134F, K180N, L202V, A267S, T285A, I286L) with no significant role in its enzymatic process. The 306 amino acid long Main protease (Mpro) in monomeric form consists of three domains namely N-terminal Domain-I, *N*-terminal Domain-II and C-terminal Domain-III (Fig. 1). The catalytic dyad (His41 and Cys145) located in a cleft between the β-barrel folds of domains I (residues 10−99) and domain-II (100−182). The thiol group of Cys145 acts as nucleophile in the proteolytic process. A long loop region is connected the catalytic domains to the C-terminal domain-III (residues 198−303) composed of five antiparallel α-helical [18].

The M^{pro} exist catalytically active in dimeric state and arranged perpendicular to each another. The dimeric form has a unique substrate binding preference for glutamine residue at the P1 site (Leu-Gln ↓ Ser, Ala, Gly), that distinguish it from other closely related proteases in host, suggesting that it is feasible to achieve high selectivity by targeting viral M^{pro}. The dimerization of two M^{pro} protomer, facilitate the formation of substrate-binding cleft in which conformationally distinct solvent-exposed His41 and Cys145 in the cleft act independently [19].

Scientists around the world are racing to discover the drug or vaccine against COVID-19. In this regard extensive theoretical and experimental work has been done [20–27]. Many antiviral drugs are also in clinical trials against SARS-CoV-2, yet no vaccine or antiviral agent against this deadly virus is discovered. Herein, we perform a structure based virtual screening of CAS COVID-19 database with potential antiviral properties to identify prototypic inhibitors against the SARS-CoV-2 M^{pro} target. The approach involved score-based virtual screening, molecular dynamics simulations and binding free energy calculations. Based on consensus scoring strategy followed by flexible docking and post docking analysis (i.e. protein-ligand interactions fingerprinting), four compounds SKS-01, SKS-02, SKS-03 and SKS-04 were selected for further validation through MD simulation.

2. Methodology

2.1. Database preparation

The Chemical Abstract Services released an open source CAS COVID-19 database, containing ~50,000 potential antiviral candidates. The dataset was downloaded (https://www.cas.org/covid-19-antiviral-compounds-dataset) and the redundant compounds were checked by their CAS number and eliminated. Afterwards, the database was converted from 2D to 3D followed by geometry optimization using WASH module implemented in MOE. Subsequently, hydrogens were added followed by energy minimization using MMFF94 force field and single low energy conformation was generated with original state of chirality and ionization states for each compound. Considering the larger size of binding pocket of 3CLPro and high molecular weight of effetive antiviral drugs against coronaviruses main protease (e.g. Saquinavir,

Darunavir, Lopinavir etc.), compounds with molecular weight upto 715 g/mol were selected. The final database consisting of 35,000 compounds, was saved in mdb format for screening.

2.2. Receptor preparation

Structure-based virtual screening was carried out using the x-ray crystal structure of chymotrypsin-like protease (PDB-ID 6LU7, resolution 2.16 Å) [18] of SARS-CoV-2 bound to peptide like inhibitor (N3). Co-crystalized artifacts such as missing atoms and loops were modeled by Structure Preparation module implemented in MOE 2019.01. All the crystallized water molecules were deleted, and the crystal structure was subjected to protonation and energy minimization using Amber99 force field in MOE. A docking grid of 6 Å was generated, centered on the protein active site by selecting the atoms of co-crystalized ligand, N3.

2.3. Structure-based virtual screening

Prepared CAS antiviral database containing 35,000 compounds were docked into the define grid center on the co-crystalized ligand. During docking the receptor kept flexible and Triangular Matcher was used as placement method, which is the most efficient method for well-defined binding sites. For comprehensive extraction of active compounds, comparative assessment of different scoring functions implemented in MOE was used to evaluate the performance of scoring function. Approximately 34,000 compounds were docked, poses for each compound were calculated by using five different scoring functions: London dG, GBVI/WSA dG, ASE, Affinity dG and Alpha HB. Compounds with different scoring function were sort out from highest to lowest value and top 10% of the data in all scoring functions were selected for further analysis.

Protein-ligand interaction fingerprint (PLIF) module implemented in MOE was used to analyzed all interactions between protein and the selected compounds [28]. The top 78 compounds were selected and subjected to protein ligand interaction profile (PLIP) for their binding interactions. As a result of this analysis compounds were shortlisted on the basic of interactions with hotspot residues in addition to catalytic dyad; Phe140, Leu141, Asn142, Gly143, Ser144, Met165 and Gln189. Finally, 4 hits, containing tetrahydrofuran and purine ring as chemical scaffold, obtained from structure-based virtual screening were subjected to molecular dynamics simulation. The schematic representation of stepwise screening strategy shown in Fig. 2.

2.4. Molecular dynamics simulation

All the simulations were carried out using GROMACS version 2018 [29] on GPU. The topological parameters of protein structure were processed by GROMOS 57a7 force field [30] while Automated Topological Builder was used for the parameterization of all the ligands. All the complex structures were solvated using single point charge water model by placing into a cubic water box and by keeping at a distance of 1.0 nm from the edge of the water box with periodic boundary condition. Subsequently, the systems were neutralized by the addition of counter ions. Prepared systems were minimized to remove the bad contacts and to correct the geometry by using steepest descent algorithm with maximum force $<$ 1000 kJ mol$-$1 nm$-$1. The electrostatic interactions were treated by Particle Mesh Ewald algorithm with a cut-off of 1.2 nm while LINCS algorithm was used to constrained the bond length. To obtain further convergence, the minimized systems were equilibrated for 500 ps in an NVT ensemble followed by NPT ensemble. During equilibration, pressure and volume were maintained by using isobaric-isothermal algorithm. Finally, the production run of 100 ns was carried out and coordinates along with their energies were saved after every 10 ps in output trajectory file. The convergence for all the simulated systems were analyzed by RMSD RMSF, RoG, and hydrogen bond contacts using GROMACS and VMD [31]. Moreover, the Coulombic potentials within R-coulomb (Coul-SR) and the short-range Lennard-Jones energies (LJ-SR) were calculated using the g_energy program available in GROMACS, extracting the information from .edr files produced during the MD simulations.

3. Results and discussion

3.1. Structure-based virtual screening

The recent coronavirus pandemic demands the immediate discovery of therapeutic agent to cure and protect the people who are suffering from the COVID-19. In order to combat the spreading pandemic and to accelerate the discovery of antiviral agent against SARS-CoV-2, structure-based virtual screening was carried out employing the molecular docking approach against 3CLPro. For virtual screening, an open access CAS database is utilized which consists of nearly 50,000 compounds with potential or known antiviral activity.

The crystal structure of 3CLPro in complex with a peptide inhibitor N3 (PDB ID 6LU7) recently reported by Liu et al. was utilized for score-based screening. The structure of 3CLPro consist of three domains in which substrate binding site or active site is located at the cleft of domain I and II where N3 is bound. Like other coronaviruses, the active site of 3CLPro of SARS-CoV-2 consists of conserved catalytic dyad (His41 & Cys145) with other crucial residues of five subsites. This substrate binding pocket served as receptor grid while position of N3 served as center of the grid during the docking based screening. During docking simulation, induced fit method was employed thus providing flexibility of the residues of the active site.

It has been reported that combining different scoring functions or consensus scoring leads to a higher probability of enrichment of true positives. This study mainly focused on empirical scoring functions implemented in MOE to provide a real feature of combination of scoring functions to extract potential hit compounds. In this regard, the prepared database consists of 35,000 compounds was docked into the define grid using different scoring function including London dG, GBVI/WSA dG, ASE, Affinity dG and Alpha HB with Triangular Matcher as placement method (Table 1). Total 34,372 compounds were correctly docked into the defined pocket, and the docked poses of each candidate were sort out from highest to lowest docking score. Afterwards, the compounds were shortlisted by employing consensus scoring strategy [32,33] which led the selection of top 10% of data containing compounds from each scoring function and the top-scored ~3400 compounds were considered for further screening. Thus, the outcome of the virtual screening was inspected in terms of (i) docking accuracy i.e., rmsd to known solutions, (ii) scoring accuracy i.e., prediction of the absolute binding free energy), (iii) comparisons of different scoring, (iv) discrimination of true positive from random compounds, and (v) enrichment factors and hit rates among the top scorers.

Further the compounds were subjected to PLIF to fingerprint the interactions between hit compounds and the active site residues of the 3CLPro. During the interactions fingerprinting, His41 and Cys145 (catalytic dyad) were mainly considered as crucial residue and their interaction analysis led to the selection of 78 compounds. Subsequently, the obtained compounds form PLIF ranking were subjected to PLIP (Protein-Ligand Interaction Profile) and visual inspection at molecular level. As a result of PLIP analysis, 4 compounds (containing tetrahydrofuran and purine ring as chemical scaffold) were shortlisted on the basis of their interactions with the crucial residues of the active site in addition to catalytic dyad (His41, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145,

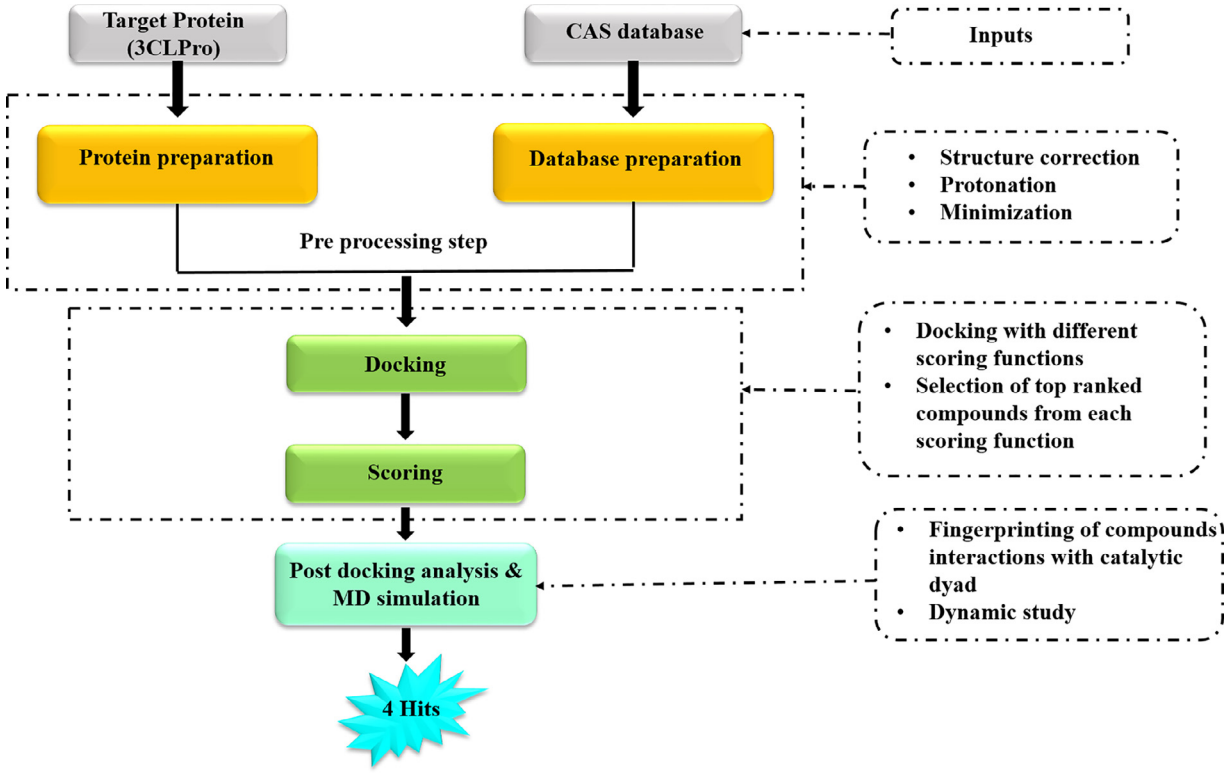

Fig. 2. Schematic representation of structure-based virtual screening of CAS antiviral database.

Table 1
Combination of scoring function utilized for docking.

Placement Method	Refinement Method	Scoring Function	Rescoring Function
Triangular Matcher	Induced Fit	London dG	GBVI/WSA dG
Triangular Matcher	Induced Fit	London dG	ASE
Triangular Matcher	Induced Fit	ASE	Affinity dG
Triangular Matcher	Induced Fit	Alpha HB	London dG
Triangular Matcher	Induced Fit	Affinity dG	Alpha HB
Triangular Matcher	Induced Fit	London dG	Affinity dG
Triangular Matcher	Induced Fit	London dG	Alpha HB

Met165 and Gln189). These hits compounds were further subjected to MD simulation to study the stability and dynamic of protein-ligand complexes.

3.2. Molecular dynamic simulation

The molecular dynamics simulation provides the information about the stability of the protein-ligand complexes in a time dependent manner. In this study, we employed 100 ns simulation of 4 compounds: SKS-01, SKS-02, SKS-03 and SKS-04 in complex with 3CLPro of SARS-CoV-2. The stability of these ligands within the active site of target protein was analyzed by plotting root mean square deviation (RMSD) of the backbone atoms and root mean square fluctuation (RMSF) of amino acid residues. In order to determine the compactness of the protein upon binding of the ligands, radius of gyration (RoG) was also calculated.

3.2.1. Root mean square deviation (RMSD)

MD trajectories were analyzed to ensure the binding stability of all the four protein-ligand complexes by calculating the deviation of backbone carbon atom. All the systems showed inconsiderable fluctuation with the deviation of less than 0.4 nm during the whole simulation Fig. 3). Initially SKS-01 showed variable fluctuation but after 16 ns, RMSD become quite stable. Slight fluctuations were observed till the end of the simulation with the inconsiderable deviation of less than 0.35 nm. SKS-02 has represented the most stable RMSD below 0.3 nm with no major variations. However, the system converged after 27 ns of the simulation. SKS-03 showed slightly fluctuated RMSD in comparison to other systems. Therefore, marginal elevation was observed throughout the simulation with the deviation of less than 0.4 nm. In case of SKS-04 bound to the protein, insignificant variations were observed throughout the simulation with the deviation of less than 0.35 nm. The backbone profile of all the systems was almost similar with no conformational changes.

3.2.2. Root mean square fluctuation (RMSF)

RMSF represented the average displacement of each residue upon binding of all the four ligands (Fig. 4). Interestingly, conserved residues (His41, Asn142, Gly143, Ser144, Cys145 and Met165) involved in the ligands binding were stable throughout the whole simulation. The average value of RMSF for all the systems was around 0.3 nm. Protein-ligand complexes showed similar pattern of RMSF expect SKS-04. The RMSF analysis of SKS-01, SKS-02 and SKS-03 showed higher fluctuation in three loop regions; Thr45-Leu50, Ile152-Cys156 and Met275-Glu288. In addition to these regions, SKS-04 showed fluctuation in the region of loop connecting Domain II and III. Moreover, His41, a catalytic

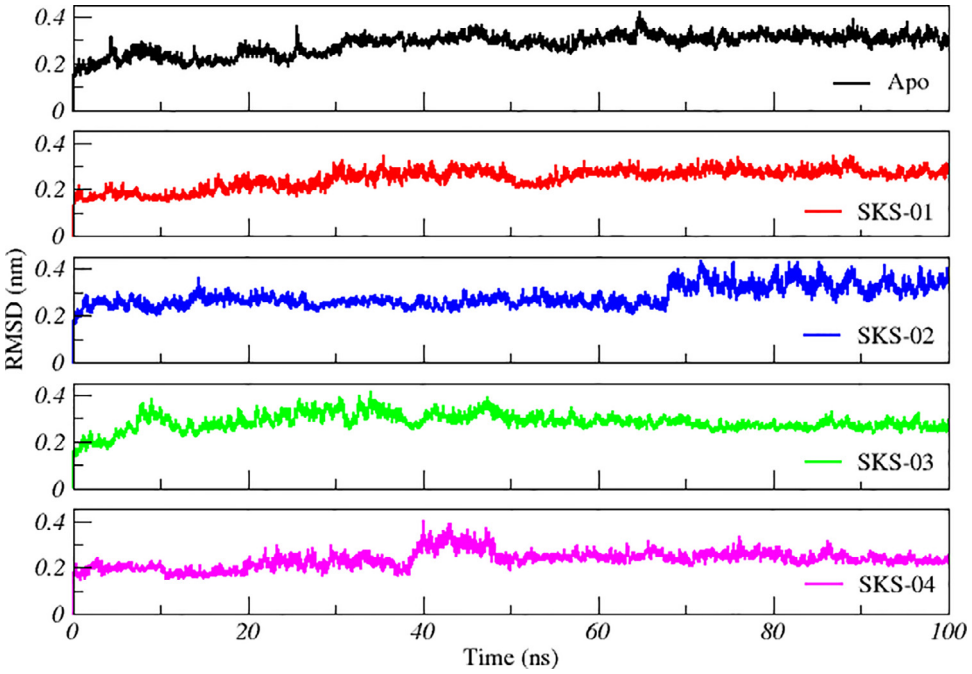

Fig. 3. Time dependent RMSD of all the four simulated systems.

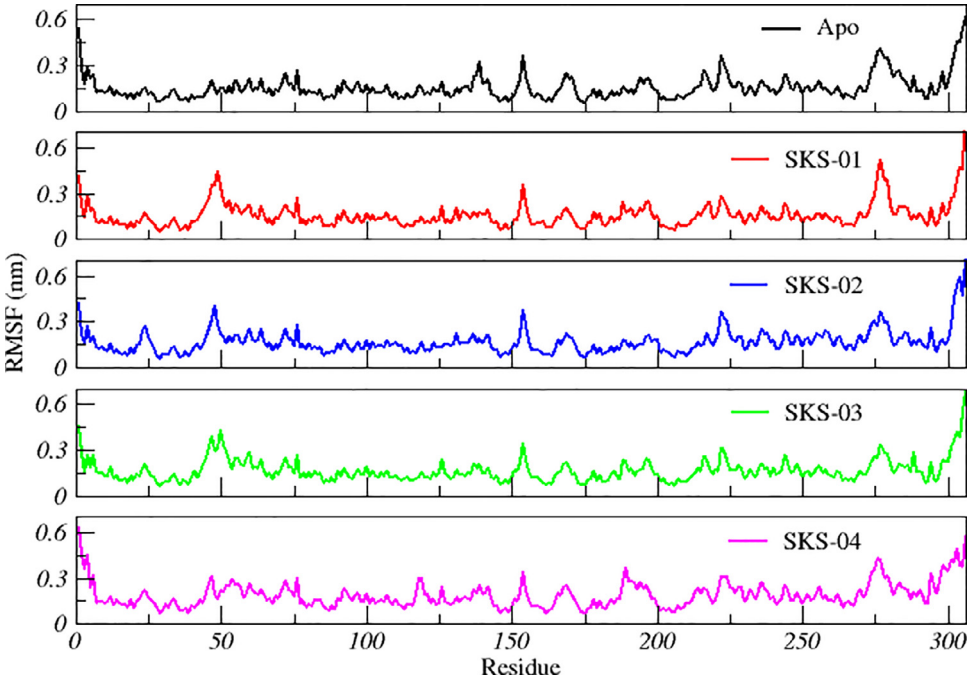

Fig. 4. Time dependent RMSF of all the four simulated systems.

dyad amino acid residue, showed higher fluctuation in all the systems except apo protein which may attribute to the involvement of His41 in the accommodation of ligands into the binding site. Similarly, all the other amino acid residues of the simulated systems showed significantly stable RMSF.

3.2.3. Radius of gyration (RoG)

In order to determine the compactness or folding of the target protein during 100 ns simulation, radius of gyration was calculated. Gyration score for all the systems oscillated between 2.15 nm to 2.30 nm (Fig. 5). The average gyration score of Apo, SKS-01, SKS-02, SKS-03 and SKS-04 was found to be 2.20 ± 0.04, 2.21 ± 0.02, 2.24 ± 0.01, 2.23 ± 0.01 and 2.21 ± 0.01 nm respectively. RoG analysis suggested that all the simulated systems more or less showed similar degree of compactness.

3.2.4. Binding mode analysis

Comprehensive interactions profile before and after MD simulation, between the selected compounds and the important binding residues of target protein is presented in Table 2. Before and after MD simulation all the compounds were positioned at the cleft of Domain I and II of 3CLPro. Thr25, His41, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, Met165, Glu166, and Gln189 contribute significantly in the binding of the compounds. All these

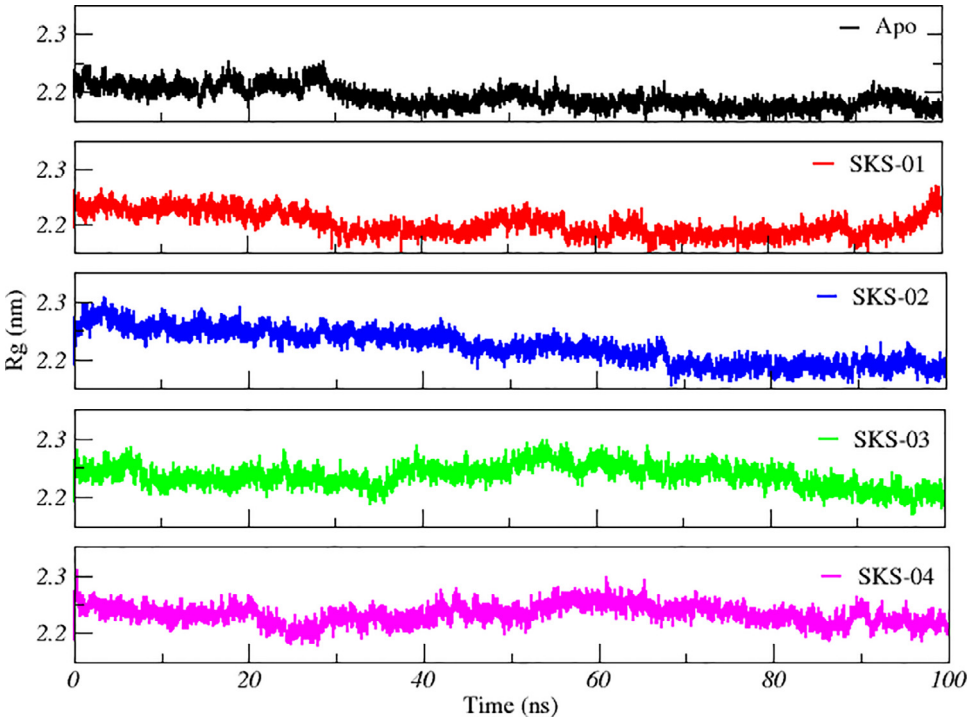

Fig. 5. Time dependent RoG of all the four simulated systems.

Table 2
Detailed protein-ligand interactions before and after MD simulation.

Compound Name	Docking score (kcal/mol)	Vann der Waal Interactions After Docking	Hydrogen Bond Contacts After Docking			Vann der Waal Interactions After MD Simulation	Hydrogen Bond Contacts After MD Simulation	
			Acceptor	Donor	Distance		Acceptor	Donor
SKS-01	−11.49	His41, Met165, Gln189	Lig: Furan=*O* Lig: Methoxy=*O* Lig: Oxirane=*O* Lig: Carbonyl=*O* Lig: Carbonyl=*O* Lig: Triazole=*N*	His41-NH Asn142-NH Gly143-NH Ser144-OH Glu166-NH Gln189-NH	2.18 2.17 2.08 3.11 2.15 3.44	Thr25, Leu141	Lig: Carbonyl=*O* Lig: Carbonyl=*O* Lig: Carbonyl=*O* His164: Carbonyl=*O* His164: Amide=*N* Lig: Carbonyl=*O*	His41-NH Gly143-NH Ser144-NH Lig: Amino=NH Lig: Amino=NH Met165-NH
SKS-02	−15.32	Thr25, Met165, Gln189	Lig: Carbonyl=*O* Asn142: Carbonyl=*O* Lig: Amide=*N* Lig: Sulfonyl=*O*	His41-NH Lig: Methoxy=*O* Gly143-NH Glu166-NH	2.73 2.53 2.89 2.27	His41, Met165, Gln189	Lig: Hydroxyl=*O* His41: Amide=*O*	Thr26: Carbonyl=*O* Lig: Amide-NH-Thr25
SKS-03	−13.67	Thr25, Leu27, Gln189	Lig: Carbonyl=*O* Lig: Pyrimidine=*N* Lig: Carbonyl=*O*	His41-NH His163-NH Glu166	2.33 2.82 2.02	His41, Met49, Met165, Gln189	Lig: Purine-N Gly143: Carbonyl=*O* Lig: Oxycarbonyl=*O* Lig: Carbonyl=*O* Lig: Carbonyl=*O*	Asn144-NH Lig: Amide-NH Ser144-NH Cys145-NH Met165-NH
SKS-04	−15.22	–	Lig: Hydroxyl=*O* Lig: Carbonyl=*O* Lig: Nitro=*O* Lig: Triazole=*N* Lig: Hydroxyl=*O*	His41-NH Gly143-NH Glu166-NH His163-NH Gln198-NH		–	Lig: Hydroxyl=*O* Lig: Hydroxyl=*O* Lig: Nitro=*O*	His41-NH His164-NH Glu166-NH

residues are critically important for substrate binding. Insight docking modes of the selected ligands showed that all the compound found to mediate van der Waals interactions as well as hydrogen bond contact with the afore-mentioned residue. Results of MD simulation demonstrated that all the docking interactions didn't persist throughout simulation, however some new interactions were also observed during the course of simulation. For time dependent analysis of protein-ligand interactions, residues within the 5 Å of ligands were selected. All the four ligands showed similar type of interactions as like N3 (co-crystallized ligand).

Upon examining the MD trajectories, SKS-01 found to form six hydrogen bonds contacts with significant persistence (Fig. 6A). Oxygen of benzoyl oxycarbonyl moiety in the ligand mediate hydrogen bond with the backbone amide of Met165 (Table 2). Similarly, alanine carbonyl group establishes two hydrogen bonds with the nitrogen of amino group of Gly143 and Ser144. Moreover, amide group adjacent to triazole ring in the ligand mediate one hydrogen bond with His41 and two other hydrogen bonds with His164. Additionally, methyl group in the ligand interact with the side chain of Thr25 via alkyl-alkyl interaction producing hydropho-

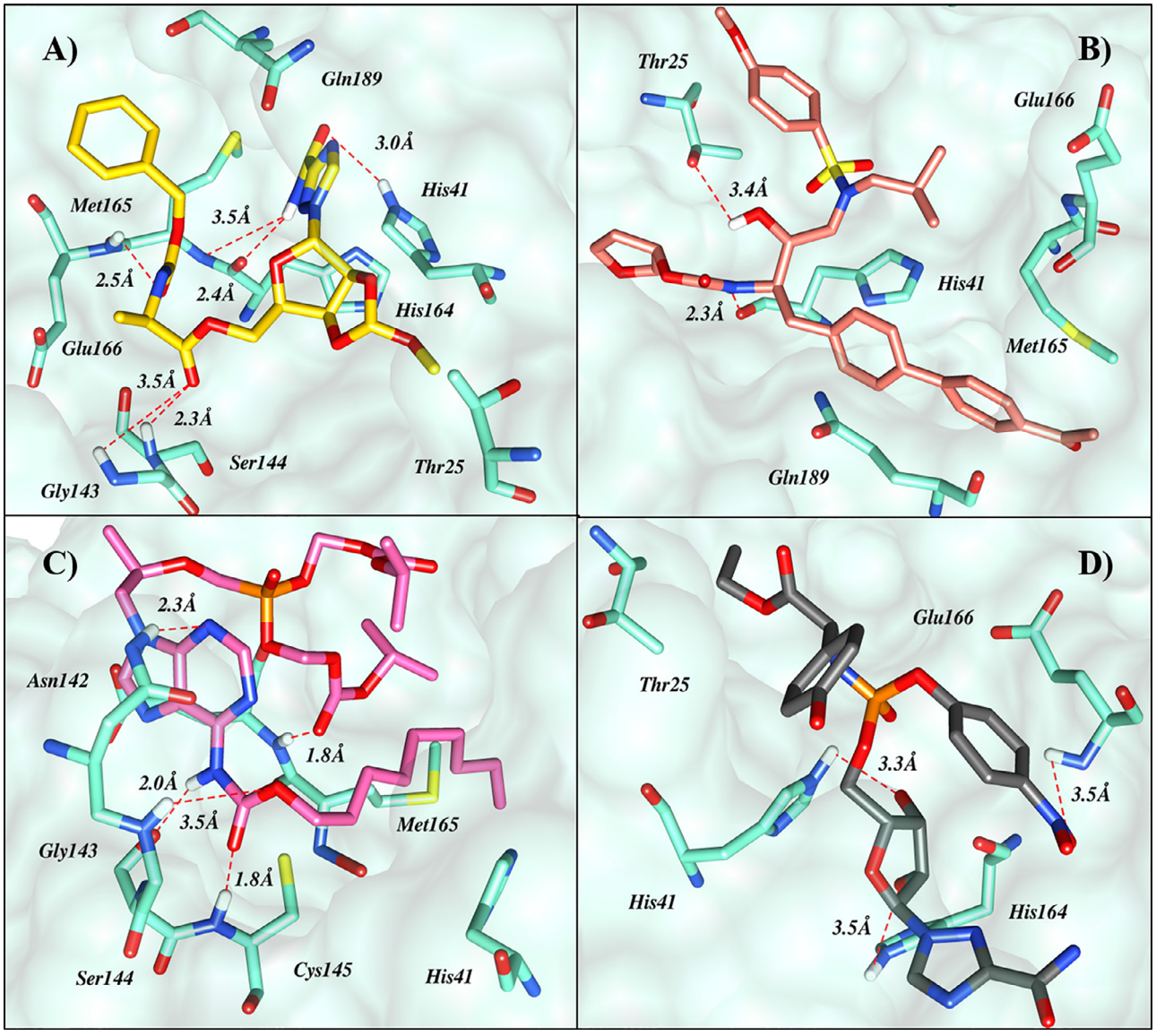

Fig. 6. Binding mode of selected hits retrieved from score based screening. (A) SKS-01, (B) SKS-02, (C) SKS-03 and (D) SKS-04. Interacting residues presented as aquamarine sticks while red dotted lines presented the hydrogen bonds.

bic effect. SKS-01 mediate highest number of hydrogen bonds as compare to the other ligands.

Insight into the MD trajectories of SKS-02 (Fig. 6B) indicated that most of the hydrogen bonds observed in docked pose were not persistent during simulation. However, Van der Waals interactions were conserved throughout the simulation. Biphenyl ring in the ligand establish π-stacking with imidazole ring of His41. Similarly, Gln189 and Met165 also found to interact with ligand by mediating π-alkyl interactions. Moreover, hydroxyl group in the ligand act as proton donor and establish hydrogen bond with the Thr25. Nitrogen of amide group near to tetrahydrofuran ring in the ligand form hydrogen bond with backbone amide of His41. All these interactions persist significantly throughout the simulation.

In case of SKS-03 some new interactions were observed other than docking interactions (Fig. 6C). Carbamate group in the ligand establish three hydrogen bonds with Gly143, Ser144 and Cys145. Moreover, nitrogen of purine ring in the ligand form a hydrogen bond with Asn142 while oxycarbonyl group mediate hydrogen bond with Met165. The protein-ligand contacts are further stabilized by His41 and Met165 by producing hydrophobic effect and π-stacking.

The analysis of time dependent molecular interactions of SKS-04 demonstrated that most of the docking interactions persisted throughout the MD trajectory (Fig. 6D). His41 and His164 act as hydrogen bond donor and form a hydrogen bond with the hydroxyl group attached to tetrahydrofuran ring in the ligand. Similarly, oxygen of nitrophenyl group in the ligand establishes a hydrogen bond with the backbone amide of Glu166. Similarly, Thr25 contributing the binding of ligand by mediating hydrophobic interactions with the ethoxy group in the ligand.

3.2.5. Protein-ligand interactions energy

A quantitative knowledge of receptor–ligand binding affinities is important in understanding of molecular recognition; hence, the calculation of effective and accurate binding free energy (ΔGbind) plays a central roal in computer-aided drug design. Therefore, in this study to compute the strength of protein-ligand interactions, short range Coulombic (Coul-SR) and lennard-Jones (LJ-

SR) interactions energies were calculated. The sum of Coul-SR and LJ-SR for SKS-01, SKS-02, SKS-03 and SKS-04 was found to be −283.1973 kJ/mol, −342.338 kJ/mol, −238.1212 kJ/mol and −311.7273 kJ/mol respectively. The data indicate that non-bonded interactions energies of all four complexes were fairly enough to explain the strenght of protein-ligand interactions.

4. Conclusion

The SARS-CoV-2 chymotrypsin like protease, plays a pivotal role in the virus maturation and life cycle, therefore it is considered as an attractive drug target to combat COVID-19. Due to lack of antiviral therapy and vaccine under development, the identification of chemical probe inhibitors of 3CLPro that can potentially be optimized as therapeutic agent seems to be highly desirable. In the present study structure-based virtual screening of CAS antiviral database containing ~50,000 compounds was carried against 3CLPro of SARS-CoV-2. Virtual screening campaign utilize the consensus scoring strategy in combination with flexible docking which led the extraction of potential hit compounds. The extracted compounds were further subjected to molecular interactions analysis. In the successive screening, 4 hits were shortlisted and subjected to dynamic study. The results of MD simulation suggested that all the selected compounds significantly stabilize the protein structure and have encouraging potential to inhibit the protein activity. We believe that our findings provide suitable chemical candidates for further detailed in vitro and in vivo analyses.

Declaration of Competing Interest

The authors declare no competing financial interests.

CRediT authorship contribution statement

Komal Zia: Conceptualization, Methodology, Visualization, Data curation, Writing - original draft. **Salman Ali Khan:** Conceptualization, Methodology, Writing - original draft. **Sajda Ashraf:** Writing - review & editing, Data curation. **Mohammad Nur-e-Alam:** Funding acquisition. **Sarfaraz Ahmed:** Funding acquisition. **Zaheer Ul-Haq:** Supervision, Project administration, Writing - review & editing, Software, Funding acquisition.

Acknowledgement

The authors also thank the Deanship of Scientific Research at King Saud University, Saudi Arabia for funding through the research group project no. RGP-1438-043.

References

[1] W. Wang, J. Tang, F. Wei, Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China, J. Med. Virol. 92 (2020) 441–447, doi:10.1002/jmv.25689.
[2] A.E. Gorbalenya, S.C. Baker, R.S. Baric, R.J. de Groot, C. Drosten, A.A. Gulyaeva, B.L. Haagmans, C. Lauber, A.M. Leontovich, B.W. Neuman, D. Penzar, S. Perlman, L.L.M. Poon, D.V. Samborskiy, I.A. Sidorov, I. Sola, J. Ziebuhr, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nat. Microbiol. 5 (2020) 536–544, doi:10.1038/s41564-020-0695-z.
[3] S.A. Khan, K. Zia, S. Ashraf, R. Uddin, Z. Ul-Haq, Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach, J. Biomol. Struct. Dyn. (2020) 1–10, doi:10.1080/07391102.2020.1751298.
[4] N. Jebril, World health organization declared a pandemic public health menace: a systematic review of the coronavirus disease 2019 "COVID-19, SSRN Electron. J. (2020), doi:10.2139/ssrn.3566298.
[5] D.A. Schwartz, A.L. Graham, Potential maternal and infant outcomes from coronavirus 2019-NCOV (SARS-CoV-2) infecting pregnant women: lessons from SARS, MERS, and other human coronavirus infections, Viruses 12 (2020) 194, doi:10.3390/v12020194.
[6] Z.W. Ye, S. Yuan, K.S. Yuen, S.Y. Fung, C.P. Chan, D.Y. Jin, Zoonotic origins of human coronaviruses, Int. J. Biol. Sci. 16 (2020) 1686–1697, doi:10.7150/ijbs.45472.
[7] X. Jiang, S. Rayner, M.H. Luo, Does SARS-CoV-2 has a longer incubation period than SARS and MERS? J. Med. Virol. 92 (2020) 476–478, doi:10.1002/jmv.25708.
[8] M. Shammi, M. Bodrud-Doza, A.R.M.T. Islam, M.M. Rahman, Strategic assessment of COVID-19 pandemic in Bangladesh: comparative lockdown scenario analysis, public perception, and management for sustainability, Environ. Dev. Sustain. (2020) 1–44, doi:10.1007/s10668-020-00867-y.
[9] B. Paital, K. Das, S.K. Parida, Inter nation social lockdown versus medical care against COVID-19, a mild environmental insight with special reference to India, Sci. Total Environ. 728 (2020) 138914, doi:10.1016/j.scitotenv.2020.138914.
[10] S. Kaur, H. Bherwani, S. Gulia, R. Vijay, R. Kumar, Understanding COVID-19 transmission, health impacts and mitigation: timely social distancing is the key, Environ. Dev. Sustain. (2020) 1–17, doi:10.1007/s10668-020-00884-x.
[11] A.R. Fehr, S. Perlman, Coronaviruses: an overview of their replication and pathogenesis, in: Coronaviruses Methods Protoc, Springer, 2015, pp. 1–23, doi:10.1007/978-1-4939-2438-7_1.
[12] Y. Yan, W.I. Shin, Y.X. Pang, Y. Meng, J. Lai, C. You, H. Zhao, E. Lester, T. Wu, C.H. Pang, The first 75 days of novel coronavirus (SARS-CoV-2) outbreak: recent advances, prevention, and treatment, Int. J. Environ. Res. Public Health. 17 (2020) 2323, doi:10.3390/ijerph17072323.
[13] E. Ortiz-Prado, K. Simbaña-Rivera, L. Gómez- Barreno, M. Rubio-Neira, L.P. Guaman, N.C. Kyriakidis, C. Muslin, A.M.G. Jaramillo, C. Barba-Ostria, D. Cevallos-Robalino, H. Sanches-SanMiguel, L. Unigarro, R. Zalakeviciute, N. Gadian, A. López-Cortés, Clinical, molecular, and epidemiological characterization of the SARS-CoV-2 virus and the coronavirus disease 2019 (COVID-19), a comprehensive literature review, Diagn. Microbiol. Infect. Dis. 98 (2020) 115094, doi:10.1016/j.diagmicrobio.2020.115094.
[14] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E.C. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269, doi:10.1038/s41586-020-2008-3.
[15] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: genome structure, replication, and pathogenesis, J. Med. Virol. 92 (2020) 418–423.
[16] H. Yang, M. Bartlam, Z. Rao, Drug design targeting the main protease, the achilles heel of coronaviruses, Curr. Pharm. Des. 12 (2006) 4573–4590, doi:10.2174/138161206779010369.
[17] K. Anand, H. Yang, M. Bartlam, Z. Rao, R. Hilgenfeld, Coronavirus main proteinase: target for antiviral drug therapy, in: Coronaviruses with Spec. Emphas. First Insights Concern. SARS, Springer, 2005, pp. 173–199, doi:10.1007/3-7643-7339-3_9.
[18] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L.W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, Nature 582 (2020) 289–293, doi:10.1038/s41586-020-2223-y.
[19] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors, Science (80-.) 368 (2020) 409–412, doi:10.1126/science.abb3405.
[20] Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, F. Cheng, Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2, Cell Discov. 6 (2020) 1–18, doi:10.1038/s41421-020-0153-3.
[21] L. Caly, J.D. Druce, M.G. Catton, D.A. Jans, K.M. Wagstaff, The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro, Antiviral Res. 178 (2020) 104787, doi:10.1016/j.antiviral.2020.104787.
[22] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, Acta Pharm. Sin. B. 10 (2020) 766–788, doi:10.1016/j.apsb.2020.02.008.
[23] B.R. Beck, B. Shin, Y. Choi, S. Park, K. Kang, Predicting commercially available antiviral drugs that may act on the novel coronavirus (2019-nCoV), Wuhan, China through a drug-target interaction deep learning model, BioRxiv. (2020). https://doi.org/10.1101/2020.01.31.929547.
[24] X. Yao, F. Ye, M. Zhang, C. Cui, B. Huang, P. Niu, X. Liu, L. Zhao, E. Dong, C. Song, S. Zhan, R. Lu, H. Li, W. Tan, D. Liu, In vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), Clin. Infect. Dis. 71 (2020) 732–739, doi:10.1093/cid/ciaa237.
[25] M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong, G. Xiao, Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro, Cell Res. 30 (2020) 269–271, doi:10.1038/s41422-020-0282-0.
[26] F. Touret, M. Gilles, K. Barral, A. Nougairède, J. van Helden, E. Decroly, X. de Lamballerie, B. Coutard, In vitro screening of a FDA approved chemical library reveals potential inhibitors of SARS-CoV-2 replication, Sci. Rep. 10 (2020) 1–8, doi:10.1038/s41598-020-70143-6.
[27] F. Batool, E.U. Mughal, K. Zia, A. Sadiq, N. Naeem, A. Javid, Z. Ul-Haq, M. Saeed, Synthetic flavonoids as potential antiviral agents against SARS-CoV-2 main protease, J. Biomol. Struct. Dyn. (2020) 1–12.
[28] A.F. Marmolejo, J.L. Medina-Franco, M. Giulianotti, K. Martinez-Mayorga, Interaction fingerprints and their applications to identify hot spots, in: Methods Mol. Biol., Springer, 2015, pp. 313–324, doi:10.1007/978-1-4939-2914-6_20.
[29] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25, doi:10.1016/j.softx.2015.06.001.

[30] N. Schmid, A.P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A.E. Mark, W.F. Van Gunsteren, Definition and testing of the GROMOS force-field versions 54A7 and 54B7, Eur. Biophys. J. 40 (2011) 843–856, doi:10.1007/s00249-011-0700-9.
[31] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38, doi:10.1016/0263-7855(96)00018-5.
[32] J.C. Baber, W.A. Shirley, Y. Gao, M. Feher, The use of consensus scoring in ligand-based virtual screening, J. Chem. Inf. Model. 46 (2006) 277–288, doi:10.1021/ci050296y.
[33] F. Berenger, O. Vu, J. Meiler, Consensus queries in ligand-based virtual screening experiments, J. Cheminform. 9 (2017) 60, doi:10.1186/s13321-017-0248-5.