Virtual screening and structure-based 3D pharmacophore approach to identify small-molecule inhibitors of SARS-CoV-2 M\textsuperscript{pro}

Samia A. Elseginy\textsuperscript{a,b}

\textsuperscript{a}Molecular Modelling Lab., Biochemistry School, Bristol University, Bristol, UK; \textsuperscript{b}Green Chemistry Department, Chemical Industries Research Division, National Research Centre, Egypt

Communicated by Ramaswamy H. Sarma

ABSTRACT
The outbreak caused by a coronavirus 2 has required quick and potential treatment strategies. The main protease enzyme M\textsuperscript{pro} plays an important role in the viral replication which renders it an important target for discovering SARS-CoV-2 inhibitors. In this study, 3D pharmacophore structure-based virtual screening and molecular docking were conducted using MOE and Bristol University Docking Engine (BUDE). Around 400,000 molecules of ZINC15 database were docked against the crystal structure of main protease, followed by 3D pharmacophore filtration. Six top-ranked hits (ZINC87517986, ZINC60399606, ZINC58662884, ZINC45988635, ZINC54706757 and ZINC17320595) were identified based on their strong spatial affinity and forming H-bonds with key residues H41, E166, Q189 and T190 of the binding pocket of M\textsuperscript{pro} SARS-CoV-2. The 6 hits subjected to molecular dynamics simulations for 100 ns followed by binding free energy calculations using MM-PBSA technique. Interestingly, three hits showed free binding energy ($\Delta G_{\text{binding}}$) lower than tert-butyl N-[1-[[2S]-1-[[2S]-4-(benzylamino)-3,4-dioxo-1-[[3S]-2-oxopyrrolidin-3-yl]butan-2-yl]amino]-3-cyclopropyl-1-oxopropan-2-yl]-2-oxopyridin-3-yl]carbamate (a-ketoamide 13 b) ($\Delta G_{\text{binding}}$) $-76.67 \pm 0.5$ kJ/mol which suggested their potential against SARS-CoV-2. The best binding free energy candidates, ZINC87517986 ($\Delta G_{\text{binding}}$) $-98.41 \pm 0.7$ kJ/mol. The second-best hit candidate, ZINC54706757 ($\Delta G_{\text{binding}}$) $-83.4 \pm 0.6$ kJ/mol and the third one ZINC17320595 ($\Delta G_{\text{binding}}$) $-78.85 \pm 0.5$ kJ/mol. Per residue decomposition free energy indicates H41, S46, H164, E166, D187, Q189 and Q192 are hot spot residues while residues M49, M165, L167 and P168 contribute to the hydrophobic interactions. The pharmacokinetic study suggests that the selected 6 hits possess drug-like properties. The 3D pharmacophore virtual screening, molecular dynamics and MM-PBSA approaches facilitated identification 3 promising hits with low free binding energy as SARS-CoV-2 inhibitors.
1. Introduction

The new pathogenic human coronavirus (CoV) was reported for the first time in December 2019 as the main cause of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Zhou et al., 2020). The virus was named SARS-CoV-2, because its RNA genome is 82% identical to the SARS coronavirus (SARS-CoV) (Zheng, 2020), both viruses and MERS-CoV belong to the coronaviridae family. The three viruses are known by their ability to cause severe respiratory syndrome in human in contrast to other human coronaviruses (HC0V-NL63, HCoV-229E, HCoV-OC43 and HCoV-HKU1) which cause mild respiratory infection (Gil et al., 2020). By March 2020, World Health Organisation declared that the coronavirus disease-2019 (COVID-19) is a world pandemic as around 73 million cases were confirmed over 225 countries (Cucinotta & Vanelli, 2020) (https://covid19.who.int/). In response to the social and economic crisis caused by COVID-19 outbreak, FDA authorized remdesivir for treatment COVID-19 (Thorlund et al., 2020) and clinical trials were evaluated for marketed drugs as lopinavir, ritonavir (Sanders et al., 2020) and hydroxychloroquine (Sanders et al., 2020). However, the efficacy of these drugs still needs to be further confirmed.

There are numerous COVID-19 vaccines currently underway research and development, until January 2021 three vaccines are approved by the Medicines and Healthcare products Regulatory Agency (MHRA); Pfizer/BioNTech vaccine, Moderna vaccine and Oxford University/Astra Zeneca vaccine (UK medicines regulator gives approval for first UK COVID-19 vaccine). The phase III data showed that the vaccines have more than 90% effectiveness in preventing COVID-19 (Kesselheim et al., 2021). However, the efficacy of vaccines measured in clinical trials sometimes imperfectly predict the effectiveness in real-world. Also, It is not known how long the vaccinated person is protected, and if the vaccine is safe for the pregnant and breastfeeding women (Caserotti et al., 2021). Besides, the growing vaccine hesitancy (Dror et al., 2020) and anti-vaccination views against the new vaccines may weaken vaccination uptake rates and reduce the chance of herd immunity (Murphy et al., 2021). For that, there is still a demand for specific and potent therapeutics against coronavirus-2 to reduce the severity of the deadly disease.

Coronaviruses express two proteases Main protease (M\textsuperscript{pro}) also known as 3-chymotrypsin-like protease (3CL\textsuperscript{pro}) and Papain-like protease (PL\textsuperscript{pro}) (Osipiuk et al., 2021). M\textsuperscript{pro} is an essential for processing the polyproteins that are translated from the viral RNA making it an attractive target for inhibition viral replication (Rut et al., 2020). Moreover, The M\textsuperscript{pro} operates at around 11 cleavage sites on the virus polyprotein, and until now there is no human proteases have a similar cleavage specificity are identified (Zhang et al., 2020), that makes developing M\textsuperscript{pro} inhibitors is likely to be effective and safe. The recently released SARS-CoV-2 M\textsuperscript{pro} crystal structures provides an insight information for rational drug design for SARS-CoV-2 (Wang, 2020), a peptidomimetic inhibitor (N3) (Jin et al., 2020) and tert-buty l N-[1-[(2S)-1-[(2S)-4-(benzylamino)-3,4-dioxo-1-[(3S)-2-oxopyrrolidin-3-yl]butan-2-yl]amino]-3-cyclopropyl-1-oxopropan-2-yl]-2-oxopyridin-3-yl)carbamate (α-ketoamide 13b) that has been recently crystallized with SARS-CoV-2 (Zhang et al., 2020) promote the virtual screening of available libraries for development SARS-CoV-2 M\textsuperscript{pro} inhibitors (Al-Khafaji et al., 2021). M\textsuperscript{pro} consists of two proteomes (A and B) associate to form a dimer, each protomer contains three domains: domain I (residues 8–101), domain II (residues 102–184) and domain III (residues 201–303) (Somboon et al., 2021). Domain III connected to domain II by a long loop region (residues 185 – 200), dimerization of M\textsuperscript{pro} regulates through salt bridge between Q290 of a proteome A and R4 of a promoter B, the M\textsuperscript{pro} dimerization is an important for the catalytic activity as it helps in shaping the S1 pocket (Shi & Song, 2006). The substrate-binding site is in a cleft between domain I and domain II, the binding site formed of four binding sites S1, S2, S3 and S4 (Anand et al., 2002). S1 subsite formed of F140, L141, N142, C145, H163, E166, H172 while S2 subsite which is hydrophobic consists of H41, M49, H164 and M165. S3 formed of Q189, and S4 formed of L167, F185 and Q192 and the main chain of Q189 (Jin et al., 2020). Computer-aided drug design (CADD) employing a molecular docking, virtual screening. 3D pharmacophore structure-based drug design and molecular dynamics have played an essential role for rapid discovering and developing SARS-CoV-2 M\textsuperscript{pro} inhibitors (Das et al., 2021).

The success in our previous work (Elseginy et al., 2021) in discovering SARS-CoV-2 Papain like protease inhibitor using 3D pharmacophore virtual screening technique encourage us to develop COVID-19 inhibitors by targeting M\textsuperscript{pro} of SARS-CoV-2 employing 3D pharmacophore structure-based virtual screening technique and molecular dynamics (MD) simulation. Molecular docking study were carried out to 400,000 commercial compounds from ZINC15 database (Irwin et al., 2012) then the compounds filtered through 3D pharmacophore using Bristol University Docking Engine (BUDE) (Mcintosh-Smith et al., 2015). Molecular docking performed prior to pharmacophore filtration to improve the efficiency and accuracy of filtration process. The compounds showed binding score better than tert-buty l N-[1-[(2S)-1-[(2S)-4-(benzylamino)-3,4-dioxo-1-[(3S)-2-oxopyrrolidin-3-yl]butan-2-yl]amino]-3-cyclopropyl-1-oxopropan-2-yl]-2-oxopyridin-3-yl)carbamate (α-ketoamide 13b) were selected and undergo visual inspection. MD for 100 ns was performed for the selected hits and the control compound (α-ketoamide 13b) and the free binding energy were calculated using molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) approach (Wang et al., 2019) to identify the hits with lowest binding energy. The ligand – residues interaction profile and the decomposition of binding free energy into different components were analyzed. The finding of this study provides insight into rationally designing potent inhibitors of SARS-CoV-2 main protease.

2. Computational method

2.1. Docking screening

2.1.1. Prepare the ligands

A library of 400,000 compounds were download from ZINC15 database, the ligands were protonated, partial
charges were assigned by MOE using Gastiger method and saved as mol2 files. The energy of ligands was minimized by MMFF94x force field until RMSD gradient of 0.05 kcal/mol Å was reached. The ligands were kept flexible during the docking simulation.

2.1.2. Prepare the protein
The crystal structures of APO (6M03) and SARS-CoV-2 main protease (Mpro) complexed with peptidomimetic inhibitor N3 and α-ketoamide 13b (PDB: 6LU7, 6Y2G, respectively) were downloaded from Protein Data Bank (www.rcsb.org), PDB structures employed for this study were selected based on their availability and their high resolution (6LU7: 2.16 Å and 6Y2G: 2.20 Å) and absence of the mutations.

Hydrogen atoms were added for the crystal structures, water molecules and ions were removed with MOE and the protonation state of each protein residue was determined at pH 7 using the PDB2PQR server (Dolinsky et al., 2004) (http://nbcr-222.ucsd.edu/pdb2pqr2.0.0/). The missing loop regions and missing residues were inserted with MODELLER program (Eswar et al., 2006) via the UCSF Chimera graphical interface (Pettersen et al., 2004). The protein was kept rigid during the docking.

2.1.3. Molecular docking
In this study, two docking studies were performed, first one, used before pharmacophore filtration, 400,000 commercial compounds were docked using BUDE programme against Mpro complexed with peptidomimetic inhibitor N3 (PDB code: 6LU7). The binding site is located at (X = −11.45, Y = 12.85, Z = 70.08) which is identified up on the center of N3. Before virtual screening, the accuracy of the screening protocol was validated by re-docking the peptidomimetic ligand N3 in the crystal structure, through which we found that the re-docked pose obtained a minimum RMSD < 1.0 Å against the original binding pose in the crystal structure which indicated the reliability of the screening method used in this study. For further confirmation of the accuracy of our docking protocol, several SARS-CoV-2 crystal structures (PDB codes: 6LZE, 6XQS, 6XQT, 6XQU, 7ABU, 7AKU, 7AQI, 7AWS, 7AXM and 7L0D) were downloaded from https://www.rcsb.org/ and prepared as mentioned before. The co-crystallized ligands were removed from the crystal structures and re-docked in the putative binding site of each protein. The second docking was carried using MOE against Mpro complexed with α-ketoamide 13b (PDB:6Y2G), the site map tool which implements in MOE was used to determine the binding site, α-ketoamide inhibitor 13b was re-docked with MOE against crystal structure (PDB:6Y2G). Placement tool were selected to be Triangle Matcher and London dG for scoring, the resoring affinity dG was used for refinement. The control compounds in virtual screening protocol (peptidomimetic N3 and α-ketoamide13b) were re-docked with BUDE and showing docking score (−64.4 and −65.32 kJ/mol), respectively, and re-docked with MOE showing docking score (−10.64 and −10.35 kcal/mol), respectively.

2.2. Pharmacophore – virtual screening
The crystal structure of Mpro complexed with the α-ketoamide 13b was used to generate 3D pharmacophore using Protein-Ligand Interaction Fingerprints (PLIF) tool which is implemented in MOE. The pharmacophore composed of 6 features, two H-bond donors (F1, F2), two H-bonds acceptors (F3, F4) and two hydrophobic centers (F5, F6). The resulting 3D pharmacophore validated its ability to distinguish true active compounds from decoys by screening a set of active compounds (30 compounds) (Fourches et al., 2010; Principi & Esposito, 2020; Zhang et al., 2021) and decoy compounds (300 compounds) obtained from Database of Useful Decoys (DUDe: http://dude.docking.org) (Mysinger et al., 2012). Molecules that came out from docking with BUDE and showed binding score ≤N3 were screening against the constructed 3D pharmacophore. Around 1000 molecules passed the 3D pharmacophore filtration, and 818 molecules were obtained after removing the duplicate. For consistency, the 818 were re-docked by BUDE and MOE against Mpro complexed with (α-ketoamide 13b) inhibitor, and the molecules that showed binding score lower than α-ketoamide 13b with MOE and BUDE (−10.35 kcal/mol, −65.35 kJ/mol, respectively) were selected and visualized using Pymol (DeLano, 2002). For accuracy and consistency as each software used different criteria to calculate the binding energy score, the BUDE docking score of the compounds were compared with BUDE docking score of native ligand (α-ketoamide 13b) and MOE docking score of the compounds were compared with MOE docking score of 13b. Six hits were selected upon their good interactions with the binding pocket residues. The shortlisted molecules were screened for Pan Assay Interference compounds (PAINS) using the online PAINS filters at ZINC (docking.org), and the 6 hit molecules passed the filter.

2.3. Drug likeness
The physicochemical and pharmacokinetic parameters of the 6 hits and control compound α-ketoamide 13b were calculated using SwissADME online tool (Kuhn et al., 2005). Lipinski’s parameters including molecular weight (MW), number of hydrogen bond donors, number of hydrogen bond acceptors, the partition coefficient log P and log S were calculated. Besides, topological polar surface area (TPSA), number of rotatable bonds, GI absorption, blood brain barrier (BBB) penetration and effect the 6 hits on CYP1A2. The toxicity study of the promising hits and control compound α-ketoamide 13b were carried out using pkCSM (Pires et al., 2015), and several predicators were calculated. The chemical structure of the hits was converted to SMILES format using MOE and sent to http://www.swissadme.ch/ and pkCSM (unimelb.edu.au) for analysis.

2.4. Molecular dynamics (MD) simulations
The top scoring docked poses obtained from docking against Mpro complexed with α-ketoamide 13b was taken as the initial poses for simulation. The MD simulations were carried
out with GROMACS 5.1.2 software (Berendsen et al., 1984). The topology file of protein was generated using Pdb2gmx and forcefield Amber99-SB-ILDN was employed (Lindorff-Larsen et al., 2010). The topology files of the 6 hits were created using Acype (Da Silva & Vranken, 2012) under the GAFF force field (Wang et al., 2010). A cubic periodic box was defined with a minimum margin of 1 nm, and each system was solvated and the box was filled with TIP3P water (Jorgensen et al., 1983). The sodium and chloride ions were added to the system to give an ionic strength of 0.15 M and an overall neutralize the system. The long-range electrostatics and Van der Waals (rvdw) interactions were calculated using the Particle Mesh Ewald (PME) method (Manhas et al., 2017), and the cut-off distance for the short-range Coulombic interactions and VDW was set to 1.2 nm (Wang et al., 2010). The steepest descent algorithm was used to minimized the energy of each system then each system was heated gradually from 0 to 310 K under NVT condition for 100 ps and then equilibrated with NPT conditions for 100 ps. Finally, 100 ns molecular dynamic simulations were performed for the 7 systems (6 hits and control compound complexes with SARS-CoV-2 Mpro) on Blue Crystal, the University of Bristol’s high-performance computing machine. All post MD analyses were performed using Gromacs tools and Xmgrace, molecular graphics operations and visualizations were executed using and Pymol (DeLano, 2002) and VMD-1.9.1 (Humphrey et al., 1996).

2.5. MM-PBSA binding free energy calculation

The binding free energies for the 6 hits and control compound α-ketoamide 13 b complexed with SARS-CoV-2 Mpro (PDB:6Y2G) were calculated using the MM/PBSA method (Kuhn et al., 2005). The g_mmpbsa implemented by GROMACS was used to calculate the binding free energy of the 6 hits and the control α-ketoamide 13 b complexed with receptors. The snapshots were collected every 100 ps over 100 ns MD starting from 40 ns giving 600 frames. ΔG and per residue decomposition calculations were determined using MMPBSA.py as input file (Miller et al., 2012). The ionic strength (istrng) was 0.1 and the external and internal dielectric constants were (exdi = 80, indi = 1.0), respectively. The method for calculating nonpolar solvation free energies was (inp = 2), cavity_surften =0.037 and cavity_offset = −0.569, fillratio = 4 and scale = 2.0. The number of iterations was 100 to perform of the linear PB equation (limit). A 1.4 Å probe radius (prbrad) was used for simulations carried out with water as the solvent to model the size of a typical water molecule. The solvation free energy was calculated using the default solver pbsa (Luo et al., 2002).

3. Results and discussion

3.1. Virtual screening procedure

In order to identify hits as promising inhibitors of SARS-CoV-2 Mpro target, virtual screening technique was applied followed by molecular dynamic simulation. Virtual screening strategy based on using molecular docking and pharmacophore-based structure, the database molecules were docked prior to screened by pharmacophore model to identify novel molecules in appropriate molecular shape and size for the target active site and hence improve the accuracy of screening procedure (Zhang et al., 2019).

3.1.1. Docking-based virtual screening

Validation results of the docking protocol by re-docking native ligands of several SARS-CoV-2 Mpro crystal structure (PDB: 6LU7, 6Y2G, 6LZE, 6XQS, 6XQT, 6XQU, 7ABU, 7AKU, 7AQI, 7AWS, 7AXM and 7LD) showed RMSD value between the native ligands and the re-docked poses < 1.0 Å, which indicates the reliability of the docking protocol. Next, around

---

Figure 1. Structure of APO SARS-CoV-2 Mpro and SARS-CoV-2 Mpro complexed with N3(PDB: 6LU7) and complexed with 13 b (PDB:6Y2G). (A) Superimposition of APO (dark khaki), 6LU7(red violet) and 6Y2G (purple). (B) binding site SARS-CoV-2 Mpro; compound N3 (orange red, stick) and 13 b (green, stick), N3 extend in the back hydrophobic pocket which highlighted with black circle. Residues colored :APO, dark kaki; 6LU7: red violet; and 6Y2G: purple.
400,000 commercially compounds downloaded from ZINC15 database (Irwin et al., 2012) were docked in the N3 ligand active site of SARS-CoV-2 M\textsuperscript{pro} (PDB:6LU7) using BUDE algorithm (Martineau et al., 2016; McIntosh-Smith et al., 2015) and ranked according their docking score. Around 170,000 compounds that showed binding score lower than N3 docked pose were selected for next step. The 170,000 compounds were screened against 3D pharmacophore. Around 1000 molecules were obtained and then were docked against SARS-CoV-2 M\textsuperscript{pro} complexed with $\alpha$-ketoamide 13 b (PDB:6Y2G).

The main reason for performing docking against 6LU7 and 6Y2G is that although the two structures are well aligned and both N3 and $\alpha$-ketoamide 13 b ligands showed that the same binding mode, the peptidomimetic inhibitor N3 ligand extends in the back hydrophobic pocket showing close contact to T190, A191 and A193 while tert-butoxycethenamine group of $\alpha$-ketoamide 13 b showed close contact to L167, P168 and D187 (Figure 1). In our screening, we took the advantage of binding mode of both ligands N3 and $\alpha$-ketoamide 13 b. Also it could be observed in Figure 1, APO SARS-CoV-2 M\textsuperscript{pro} (PDB: 6M03) aligned perfectly with SARS-CoV-2 complexed with N3 and $\alpha$-ketoamide 13 b (PDB: 6LU7 and 6Y2G, respectively) but it was observed that there a little difference in the position of some residues in the putative active site; T26, M49, M165, P168, Q189 and T190.

3.1.2. Pharmacophore-based virtual screening

3D structure-based pharmacophore consists of six features that was generated up on SARS-CoV-2 M\textsuperscript{pro} and $\alpha$-ketoamide 13 b interaction (PDB:6Y2G). Pharmacophore features including two H-bond centers; F1 and F2 corresponding to interaction with H41, H164 and E166, respectively. Two H-bonds acceptors; F3 and F4 corresponding to interaction with G143, C145, Q189 and Q192, two hydrophobic centers F5 and F6 corresponding to M165, L167 and P168 (Figure 2(A)). To validate the 3D constructed pharmacophore, a set of decoy compounds and active SARS-CoV-2 M\textsuperscript{pro} inhibitors were screening with the model, the sensitivity and specificity were calculated (Table 1), the ROC curve were plotted (Figure 2(B)) and area under the curve (AUC) was calculated. It is obvious from the good sensitivity and specificity values and an excellent AUC value of 0.9 indicating that the 3D pharmacophore model has the best ability to distinguish successfully active molecules from decoys. Subsequently, the pharmacophore model was used to screen the 170,000 hits, a total of 1000 molecules that best fit the pharmacophore model were obtained, to further narrow the compound, the duplicated structures were removed among 818 molecules. The remaining 818 compounds and $\alpha$-ketoamide 13 b ligand were docked within binding pocket of M\textsuperscript{pro} (PDB:6Y2G) using BUDE and MOE. The binding scores of re-docked $\alpha$-ketoamide 13 b with the BUDE and MOE (−65.32 kJ/mol and −10.35 kcal/mol), respectively, were used as a cut-off for selecting the promising hits. 100 promising hits showed (Supplementary Data, File S1) docking score better than the score of the re-docked $\alpha$-ketoamide 13 b conformation. The 100 molecules were inspected visually following the criteria; molecules that showed chemical structures diversity, followed rule of five, formed H-bond with one of Key residues H41, E166, C145, Q189 and Q192 and showed good fitting within M\textsuperscript{pro} active site. The visual inspection step gave short-list of 6 compounds (Figure 3) namely, ZINC58717986, ZINC60399606, ZINC58662884, ZINC45988635, ZINC54706757 and ZINC17320595 and their binding score by BUDE and MOE represented in Table 2. The virtual screening process was summarized in supplementary Figure S1.

3.2. Drug likeness

Physicochemical properties of the 6 molecules and control compound $\alpha$-ketoamide 13 b were evaluated using SwissADME web tool (Table 3). The molecules showed good physical parameters which indicate a good oral bioavailability, all the 6 hits follow Lipinski rule, the H-bond donor $\leq$ 5, H bonds acceptors $\leq$ 10, Molecular weight $\leq$ 500 and log $p<5.6$, the 6 hits showed log P range (1.02–3.09) which gave an indication the molecules have a good permeability across the cell membrane. Log S of 6 molecules $<−5$ which suggested they have good solubility and absorption. Furthermore the 6 molecules showed lower TPSA $\leq$ 140 $Å^{2}$.

| Table 1. Results of 3D-pharmacophore validation. |
|-----------------------------------------------|
| Total actives | 30 |
| Total inactive | 300 |
| True positive | 269 |
| True negative | 31 |
| False positive | 7 |
| Sensitivity | 0.76 |
| Specificity | 0.92 |

Figure 2. (A) Structure-based 3D pharmacophore model of SARS-CoV-2 M\textsuperscript{pro}. The pharmacophore color: pink for hydrogen bond donors (F1, F2), brown for hydrogen bond acceptor (F3), cyan for H-bond acceptor (F4), green for hydrophobic groups (F5, F6). (B) ROC curve of the 3D pharmacophore.
which suggested that they have better cellular internalization properties expect, the slight increase in TPSA of compound 5 (152 Å²) which do not expect to have a significant effect on the transportation of the molecules and transportation. The 6 hits showed an inhibitory effect on cytochrome CYP1A2, indicating that they have a prolonged action and good therapeutic effect. Also, the 6 molecules do not pass the blood brain barrier (Table 3). Control compound \( \alpha \)-ketoamide 13b follows Lipinski rule expect its molecular weight >500 which violated the role of five (Table 3). It could be concluded that the 6 molecules have a good pharmacokinetics and follow Lipinski rule, based on these results all 6 compounds are predicted to have a good absorption and higher bioavailability (Lipinski, 2004).

In this study, the toxicity profile of 6 promising hits and control compound \( \alpha \)-ketoamide 13b were predicted using pkCSM tool, several predictors including mutagenicity (AMES test), human maximum tolerated dose, oral rat acute toxicity (LD50), oral rat chronic toxicity, hepatotoxicity, human ether-a-go-go related gene (hERG) inhibitors and skin sensitization these parameters are presented in Table 4. The toxicity profile of the 6 hits is well-tolerated by rats and human. The AMES mutagenicity prediction revealed absence of induced mutagenic activity by the hits and the control compound 13b, as well as absence of skin sensitization. The hERG encoded voltage-dependent ion channel, blockage this ion channel may affect the cardiac repolarization, which is a critical side-effect of non-cardiovascular therapeutic agents (Vandenbarg et al., 2012). The pkCSM prediction of hERG of our compounds and the control is negative which indicates the 6 compounds have a safe profile. The pkCSM hepatotoxicity prediction of the 6 hits and control 13b were positive, but the hepatotoxicity prediction by pkCSM is based on the human liver-related side effects of 531 compounds, if the compound is related with at least one liver pathological or physiological event, the prediction come out positive result (Fourches et al., 2010; Pires et al., 2017). It could be concluded the 6 hits showed good toxicity profile except the hepatotoxicity still needs an experimental investigation. Acute toxicity describes the immediate adverse side effects occurring in short times after administration a chemical, calculating oral acute toxicity in vivo is very complicated, time consuming and costly (Li et al., 2014). In this study, the oral acute toxicity and chronic toxicity were predicted using pkCSM and the results are represented in Table 4.

### 3.3. Molecular docking

Docking results of 6 hits and control \( \alpha \)-ketoamide 13b were represented in Figure 4 and Table 5. \( \alpha \)-ketoamide 13b formed 4-H-bonds with H41, G143, H164 and E166 in distance 3.3 Å, 2.8 Å, 2.8 Å and 3.1 Å, and hydrophobic contact with M165, L167, P168, A191 and A193. ZINC5871886 formed 3 hydrogen bonds with NH of H41 (3.0 Å), C = O of E166 (2.9 Å) and NH of Q189 (3.2 Å), and hydrophobic contact with
Table 4. Results of toxicity profile of 6 hits and control compound α-ketoamide 13b calculated by pkCSM.

| Model name | ZINC 58717986 | ZINC 60399606 | ZINC 58662884 | ZINC 45988635 | ZINC 54706757 | ZINC 17320595 |
|------------|---------------|---------------|---------------|---------------|---------------|---------------|
| M.wt (g/mol) | 593.6 | 443.4 | 422.4 | 437.4 | 408.4 | 408.4 | 460 |
| TPSA (Å²) | 164.7 | 100.2 | 102.1 | 111.8 | 92.3 | 152.3 | 11.3 |
| logP | 2.39 | 3.57 | 2.59 | 1.22 | 3.00 | 1.02 | 3.09 |
| log S | −4.21 | −7.61 | −6.75 | −6.75 | −8.24 | −5.89 | −6.63 |
| Rotatable bonds | 12 | 11 | 11 | 10 | 10 | 10 | 8 |
| HBD | 4 | 3 | 2 | 4 | 3 | 4 | 2 |
| HBA | 7 | 6 | 7 | 5 | 7 | 5 | 7 |
| GI AB | low | high | high | high | high | high | high |
| BBB | NO | NO | NO | NO | NO | NO | NO |
| CYP1A2 | NO | NO | NO | NO | NO | NO | NO |
| RS | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

Figure 4. Binding mode of the promising hits within the binding site of SARS-CoV-2 Mpro (A) α-ketoamide 13b (green, stick), (B) ZINC58717986 (green, stick), (C) ZINC60399606 (green, stick) protein represented as purple cartoon. (D) Detailed interactions of α-ketoamide 13b with the key residues (pink, stick), (E) detailed interactions of ZINC58717986 with the key residues (pink, stick) and (F) detailed interactions of ZINC60399606 with the key residues (pink, stick). Binding mode of the promising hits within the binding site of SARS-CoV-2 Mpro (A) ZINC58662884 (green, stick), (B) ZINC45988635 (green, stick), (C) ZINC54706757 (green, stick), (D) ZINC17320595 (green, stick) protein represented as purple cartoon. (E) Detailed interactions of ZINC58662884 with the key residues, (F) detailed interactions of ZINC45988635 with the key residues, (G) detailed interactions of ZINC54706757 with the key residues, (H) detailed interactions of ZINC17320595 with the key residues. Ligands represented as green stick; key residues represented as pink stick.
L27, M49, M165, L167, P168 and A191. ZINC60399606, ZINC45988635 and ZINC54706757 formed H-bond with Q189 at distance 2.8 Å, 3.1 Å and 3.1 Å, respectively. ZINC60399606 formed hydrophobic interactions with L27, M49, Y118, N119 and N142 while ZINC45988635 and ZINC54706757 showed hydrophobic contact with M49, M165, L167 and P168. ZINC58662884 showed H-bonds with C=O of T26 and NH of Q192 at distance 3.1 Å and 2.9 Å, respectively, while ZINC17320595 formed H-bonds with C=O L167 and Q189 at distance 2.9 Å and 3.5 Å, respectively. Both ligands showed hydrophobic contact with M49, M165, L167, P168, A191. It could be seen that the 6 hits formed H-bond with at least one of the key residues and formed hydrophobic contact with M49, M165, L167, P168, Ala191 and Ala193 in similar manner to control compound α-ketoamide 13b.

### 3.4. Molecular dynamics simulations

#### 3.4.1. System stability

To explore the binding stability of the selected 6 hits-MPro complexes, the docked conformations of the 6 hits obtained from the docking, and control compound α-ketoamide 13b complexed with MPro protein (PDB:6Y2G) and APO protein (PDB:6M03) were applied for 100 ns MD simulations. Firstly, we investigated the conformations of APO SARS-CoV-2 MPro and the docked α-ketoamide 13b-MPro complex during MD simulation. It showed that at 20 ns the overall structure of both proteins is aligned perfectly expect the C-terminal part (Figure 5(A)), the binding site residues of APO and docked 13b-complex well aligned expect the residues T25, M49, N142, G143, M165, P168 and Q198 (Figure 5(A)). At 50 ns, the conformations of both proteins kept stable with slight changes in parts of β-sheets and the residues of binding site of APO and docked α-ketoamide 13b-complex (PDB:6Y2G) showed the same confirmations as in 20 ns expect N142, M165, P168 and Q189 showed slightly conformational changes (Figure 5(B)). Both APO and docked α-ketoamide 13b-complex kept the same conformations until almost end of the simulation. It could be observed at 80 ns (Figure 5(C)) both proteins kept the same conformations and biding site residues of both proteins kept aligned expect the same residues which identified in first 20 ns of the simulations (Figure 5(D–F)). It also could be observed that the stability of control compound α-ketoamide 13b within the binding site and the stability of 13b-complex model with minor conformational changes which indicates the reliability of the model and stability of the binding site.

The root mean square deviation (RMSD) values of the 6 hits and control compound 13b complexed with MPro protein were calculated and plotted (Figure 6). RMSD plot for APO vs. HOLO (13b-MPro complex) showed that APO system reached equilibrium at 20 ns while 13b-MPro reached the equilibrium at 40 ns, after that both systems showed stable trajectories over the simulation (Figure 6(A)). ZINC58717986-MPro showed stable RMSD pattern similar to 13b complex (Figure 6(B)) throughout the simulation. ZINC58717986,
ZINC60399606 and ZINC17320595 complexed with M$^{\text{pro}}$ reached the equilibrium at 40 ns and remain steady stable over the remaining simulation (Figure 6(B,C,G)), while ZINC58662884 and ZINC45988635 complexed to M$^{\text{pro}}$ reached the equilibrium at 25 ns (Figure 6(D,F)), compound ZINC54706757-M$^{\text{pro}}$ (Figure 6(F)) remained stable until the

### Table 5. Summary of the interactions $\alpha$-ketoamide 13 b (control compound) and promising SARS-CoV-2 M$^{\text{pro}}$ hit inhibitors.

| Compounds     | Interacting moiety of compound | Amino acid involved                          | Distance Å | Type of interaction               |
|---------------|--------------------------------|----------------------------------------------|------------|-----------------------------------|
| $\alpha$-ketoamide (13 b) | C == O                          | NH- H41                                      | 3.3        | H-bond                            |
|               | C == O                          | NH- G143                                     | 2.8        | H-bond                            |
|               | NH                              | NH- H164                                     | 2.8        | H-bond                            |
|               | NH                              | CO- E166                                     | 3.1        | H-bond                            |
| ZINC58717986  | N of indole                     | M165, L167, P168, A191, A193                 | 3.0        | Hydrophobic interactions          |
| ZINC60399606  | S == O                          | NH- H41                                      | 2.8        | H-bond                            |
| ZINC58662884  | NH                              | C == O E166                                  | 2.9        | H-bond                            |
| ZINC45988635  | NH                              | NH-Q189                                      | 3.2        | H-bond                            |
| ZINC54706757  | OH                              | L27, M49, M165, L167, P168, A191              | 3.2        | Hydrophobic interactions          |
| ZINC17320595  | NH                              | C == O Q189                                  | 3.1        | Hydrophobic interactions          |
|               | NH                              | L27, M49, M165, L167, P168                   | 2.9        | H-bond                            |
|               | NH                              | M49, L50, M165, L167, P168, Ala191, Ala193   | 3.5        | Hydrophobic interactions          |

Figure 5. Secondary structure conformations of APO SARS-CoV-2 M$^{\text{pro}}$ (PDB:6M03) dark kaki ribbon, and $\alpha$-ketoamide 13 b complex with M$^{\text{pro}}$ (PDB:6Y2G) purple ribbon. (A) conformations of APO protein and docked $\alpha$-ketoamide 13 b-complex at 20 ns, compound $\alpha$-ketoamide 13 b represented as green sphere. (B) Conformations of APO protein and docked 13 b-complex at 50 ns, compound $\alpha$-ketoamide 13 b represented as orange sphere. (C) Conformations of APO protein and docked 13 b-complex at 80 ns, compound $\alpha$-ketoamide 13 b represented as blue sphere. (D) Detailed binding site residues conformations of APO (dark kaki) and docked $\alpha$-ketoamide 13 b-protein complex (purple), (E) At 20 ns. (F) At 40 ns. (G) At 80 ns.
end of simulation while compound ZINC58662884-M<sub>pro</sub> (Figure 6(D)) showed sudden deviation at 62 ns and lasted until 65 ns and then remained stable till end of the of simulation. ZINC45988635 reached equilibrium at 10 ns and underwent an abrupt conformational change in 39 ns, returning to the stable conformation from 42 ns and continue smoothly until the end of the simulation (Figure 6(E)). The mean RMSD value for the 6 complexes around 0.24 nm except compound 3 is 0.26 nm (Supplementary Table S1) while RMSD value for the control 13b-M<sub>pro</sub> is 0.21 nm. RMSD values revealed that all 6 hits and the control 13b M<sub>pro</sub> complexes in the simulation were well-equilibrated and remained stable throughout 100 ns simulation. It was observed that the RMSD of 6 hit-complexes slightly higher than the control compound 13b complex, this may be attributed to the size of the hits which is smaller than the 13b and due to better conformational optimization of the ligands inside the binding site the RMSD is a little higher (Wang, 2020). To confirm the stability of 6 hits within the binding pocket of SARS-CoV-2 M<sub>pro</sub>, the distances between the 6 hits and the residues of the binding site were calculated as shown in (Figure 7(A–F)). ZINC58717986, ZINC58662884 and ZINC45988635 showed the most stable distance over the 100 ns simulation with average 0.2 nm, while compound ZINC60399606 showed great deviation from the binding pocket at 38 ns and return to the smooth pattern at around 58 ns which gave indications the ligand did not fit well within the binding pocket. Compound ZINC54706757 showed small distance from binding site around 0.2 nm but it adopted two conformations from 50 ns and until the end of the simulation. Compound ZINC17320595 showed stable conformation from beginning of the simulation with small distance around 0.2 nm but it adopted another conformation with distance 0.15 nm from the binding site, in the last 20 ns of the simulation ZINC17320595 kept moving between the two conformation while still stable within the pocket (Figure 7(A–F)).

To evaluate and compare the flexibility of 6 hits-M<sub>pro</sub> complexes, the root-mean-square fluctuations (RMSF) were calculated for each system (Figure 8). The RMSF analyses
highlighted similar fluctuations pattern in the 6 hits-M\textsuperscript{pro} system, compared to the control compound 13 b complexed to M\textsuperscript{pro}. From the RMSF plots it was observed that binding of 6 hits to the protein made SARS-CoV-2 M\textsuperscript{pro} protein less flexible as compared to 13 b-M\textsuperscript{pro} (Figure 8). The greatest observed fluctuation occurred in domain III, between residues 220 and 224, may be the reason related to involved domain III in the formation homodimer, second amplitude in 189–192 which form the loop region. It could be observed that the presence of hits caused a decrease in the residue’s amplitude of movement, in domain I the residues T25, T26, the residues within the range 41–46 and residue M49 showed a decrease in the movement with the 6 hits complexes in comparison to 13 b-M\textsuperscript{pro}. In domain II there are a decrease in residues fluctuation in the range between 141–146 and 161–166. This may be attributed to the interactions of the hits with the residues of the M\textsuperscript{pro} catalytic site (T25, T26, H41, S46, M49, N142, G143, S144, C145, H163, H164, E166) mean RMSF value for \(\alpha\)-ketoamide 13 b and ZINC17320595 were 0.18 and 0.16 nm, respectively, and for compounds ZINC58717986, ZINC60399606, ZINC45988635 and ZINC54706757 is 0.14 nm, compound ZINC58662884 showed the least fluctuation value 0.123 nm (Supplementary Table S1). MD simulation results indicate that the 6 hits-M\textsuperscript{pro} complexes were stable over 100 ns simulation and the binding pocket residues undergo little conformational changes.

Moreover, the stability of hits-M\textsuperscript{pro} complexes was evaluated by calculating the radius of gyration Rg (Figure 9(A)). All the 6 hits-M\textsuperscript{pro} systems showed protein compactness, Rg results indicate that all 6 systems are stable during MD simulations.

### 3.4.2. H-bonds analysis

To get insight of hydrogen bonds between the 6 hits and binding site of M\textsuperscript{pro}, the number of hydrogen bonds was monitored over the 100 ns simulation, it was found ZINC58717986 displayed an average 3 H-bonds, while compounds ZINC60399606, ZINC45988635 and ZINC54706757 displayed an average two H-bonds and compounds ZINC58662884 and ZINC17320595 displayed four H-bonds (Figure 9(B)).

To further analyze the interactions of the 6 molecules with M\textsuperscript{pro}, the ligands-protein interaction were monitored over 100 ns simulation. Figure 10 showed that the residue Q189 formed H-bonds with 5 hits, the highest occupancy was with ZINC58717986 (88%). Compound ZINC45988635 and ZINC54706757 showed continuous H-bond with Q189 for occupancy (77%, 65%, respectively) while ZINC60399606 and ZINC17320595 showed H-bonds with occupancy (25% and 30%, respectively). The key residue H41 involved in H-bond with ZINC58717986 for 60% of the simulation. Residue E166 involved in H-bond with ZINC17320595 for 62% of the dynamic time. ZINC58662884 stabilized within the binding pocket through H-bonds with T26, S46, T190 and Q192 with occupancy (71%, 33%, 43%, 49%, respectively). Residues T24, T25, S46, S144, C145, H164, E166, D187, R188, Q189, T190 and Q192 involved in salt bridge with the 6 hits with...
occupancy more than 50% over the 100 ns. Hydrophobic interactions were observed with L27, M49, M165, L167, P168 with occupancy more than 60%. It could be concluded that the key residues H41, E166, Q189, T190 and Q192 involved in H-bonds with the 6 molecules (Figure 10). The conserved residues M49, M165, L167, P168 involved in hydrophobic interactions. Most of the binding pocket involved in salt-bridge interactions with the residues in high occupancy.
3.5. Free energy binding for complexes

To provide insight into the binding mechanisms between the 6 hits and M\(^{\text{pro}}\), binding free energies were calculated by MM-PBSA (Table 6). The free binding energy values obtained for the 6 systems were as follows: 13\(b\), compound ZINC58717986, ZINC60399606, ZINC58662884, ZINC45988635, ZINC54706757, and ZINC17320595 are \(-76.67 \pm 0.5\), \(-98.41 \pm 0.7\), \(-50.40 \pm 1.0\), \(-50.12 \pm 0.5\), \(-58.06 \pm 1.0\), \(-83.4 \pm 0.6\) and \(-78.85 \text{ kJ/mol}\), respectively. It could be seen that compounds ZINC58717986, ZINC54706757 and ZINC17320595 showed free binding energy lower than 13\(b\) which reflected the strong binding of these three molecules to M\(^{\text{pro}}\) and suggested that the three hits have a promising potential activity against SARS-CoV-2. Interestingly, ZINC58717986 showed the lowest free binding energy \((-98.41 \text{ kJ/mol}\) among the 6 hits and the \(\alpha\)-ketoamide 13\(b\)
which gave an indication ZINC58717986 may have a high potential activity against SARS-CoV-2. For all the 6 complexes, the greatest contribution for binding to the protein active site occurred due to the contributions of van der Waals (ΔEvdw) (Table 6), this is due to the non-covalent interactions carried out by the inhibitors. The polar free solvation energy (ΔGpol) was unfavorable contributes to the free binding energy of the systems. The contributions of non-polar (ΔEnonpol) and gas phase electrostatic interactions (ΔEele) were favorable to the systems, but due to their low values, van der Waals interactions were the most important for inhibitor free binding affinity.

### Table 7. Ligand – residue MM-PBSA interaction energies (kJ/mol).

| Ligand | ZINC 58717986 | ZINC 60399606 | ZINC 58662884 | ZINC 4598635 | ZINC 54706757 | ZINC 17320595 |
|--------|---------------|---------------|---------------|---------------|---------------|---------------|
| Thr25  | −1.8          | −2.0          | −2.4          | −5.4          | −1.9          | 0.5           | −2.1          |
| Thr26  | −0.4          | −0.4          | −3.1          | −0.8          | −2.1          | −0.1          | −0.4          |
| Leu27  | −1.3          | −1.6          | −1.9          | −4.4          | −2.8          | −0.4          | −1.6          |
| His41  | −2.2          | −0.6          | −0.7          | −2.3          | −1.3          | −1.8          | −0.6          |
| Val42  | −0.7          | −0.3          | −0.7          | −0.5          | −0.8          | −0.2          | −0.3          |
| Ile43  | −0.6          | −0.1          | −0.2          | −0.3          | −0.1          | 0.1           | −0.1          |
| Thr45  | −0.5          | −0.5          | −0.2          | −0.7          | −1.3          | −0.4          | −0.5          |
| Ser46  | −2.1          | −0.3          | −0.5          | −1.4          | −0.2          | −0.8          | −0.3          |
| Met49  | −8.1          | −3.6          | −5.8          | −2.9          | −4.2          | −6.4          | −0.3          |
| Gly143 | −1.1          | −0.4          | −1.6          | −1.4          | −0.3          | −0.1          | −0.4          |
| Cys145 | −0.9          | −0.6          | −1.9          | −0.53         | −2.04         | −0.69         | −0.67         |
| His163 | −0.4          | −0.3          | −0.3          | −0.04         | −0.43         | −0.31         | −0.38         |
| His164 | −0.4          | −0.1          | −0.1          | −0.23         | −0.88         | −0.46         | −0.1          |
| Met165 | −3.3          | −4.8          | −4.7          | −2.0          | −1.59         | −6.02         | −4.8          |
| Glu166 | −1.9          | −0.1          | −0.1          | −0.59         | −1.99         | −3.36         | −0.1          |
| Leu167 | −0.4          | −2.1          | −0.7          | −0.07         | −0.38         | −2.5          | −2.1          |
| Pro168 | −0.7          | −2.4          | −0.6          | −0.05         | −1.67         | −4.20         | −2.47         |
| Asp187 | −0.5          | −1.1          | −1.6          | 0.01          | −2.30         | −1.66         | −1.09         |
| Gln189 | −2.1          | −1.2          | −1.8          | −0.14         | −0.50         | −6.41         | −1.17         |
| Thr190 | −0.1          | 1.4           | −0.3          | −0.003        | −0.66         | 0.18          | 1.48          |
| Ala191 | −0.1          | −0.8          | −0.2          | −0.015        | −2.36         | −0.73         | −0.79         |
| Gln192 | −0.2          | −1.2          | −0.01         | −0.07         | 0.31          | −1.32         | −1.17         |

![Figure 11](image)

**Figure 11.** Molecular contact of ZINC58717986 in different poses during the simulation. (A) Binding mode of ZINC58717986 (blue sphere) at 40 ns, (B) ZINC58717986 (green sphere) at 60 ns, (C) ZINC58717986 (purple sphere). Detailed Interactions of ZINC58717986 (D) at 40 ns, (E) at 60 ns and (F) 80 ns.

#### 3.6. Analysis protein structure inhibitors affinity relationship

To identify the key residues contribution in the binding pocket, the binding free energies were decomposed into residues using the mmpbsa.py. The importance of identification the hot spot residues is to help in rationally design potent and selective inhibitors of MPro. The key residues for the 6 hits and 13 b are summarized in Table 7. The most significant key residues for the 6 molecules are L27, H41, S46, M49, G143, C145, H164, M165, E166, L167, P168, D187, Q189, T190 and Q192. These results are in consistent with MD results of interaction analysis.
The identified hot spot residues play an important role in SARS-CoV-2 M\textsuperscript{pro} activity and in shaping the substrate-binding site, for example, E166 is curial for dimerization of SARS-CoV-2 M\textsuperscript{pro} which is necessary for the catalytic activity, as NH\textsubscript{2} terminal of each of the two protomers interacts with E166 of the other protomer and this helps in the shaping of the S1 substrate-binding site (Zhang et al., 2020). H41 and C145 are the catalytic dyad, N3 inhibitor formed a covalent bond with C145 while \(\alpha\)-keto group of \(\alpha\)-ketoamide (13b) inhibitor nucleophilic attack by the catalytic C145 and formed a thiohemiketal and the oxyanion group of the thiohemiketal is stabilized by a hydrogen bond with His41 (Zhang et al., 2020). Residues H41, M49, M165 and D187 formed S2 subsite while residues M165, L167, P168 and Q192 formed small hydrophobic pocket (Jin et al., 2020). The control inhibitor showed H-bonds with G143, H163 and E166 (Mittal et al., 2020), interactions the promising hits with these residues suggest their potential promising activity.

3.7. Trajectory analysis of ZINC58717986

ZINC58717986 showed the best binding free energy comparing with the control compound \(\alpha\)-ketoamide 13b, showed stable distance pattern with the binding pocket, and formed H-bonds with the key residues. These good results encourage us to study the molecular contact of the ligand in the last poses of the simulation and compare the interactions and residence of the ligand of last poses with the first pose of the molecular dynamics. It was showed at first 40 ns, ZINC58717986 showed H-bonds with H41, E166 and Q189, hydrophobic contact with M49, M165, E166, P168 and A191, and close contact to G143, C145, T190 and Q192 (Figure 11(A,D)). At 60 ns the ligand showed the same binding mode and formed H-bonds and showed hydrophobic interactions similar to the pose at 40 ns (Figure 11(B,E)). ZINC58717986 continuing showing the same binding mode and H-bonds profile at 80 ns similar to the poses at 40 and 60 ns, expect T190 showed slightly movement (Figure 11(C,F)). These results confirmed that ZINC58717986 has stable binding mode within the docked pocket and kept stable interactions with the key residues.

4. Conclusion

Regarding the importance of M\textsuperscript{pro} for SARS-CoV-2 survival in the host, inhibition of M\textsuperscript{pro} represents an effective approach for the treatment of COVID-19. In this study, with the aid of the released crystallographic structure of M\textsuperscript{pro} in complex with N3 and \(\alpha\)-ketoamide 13b ligands, structure-based 3D Pharmacophore based virtual screening were performed to identify novel SARS-CoV-2 inhibitors. ZINC database libraries consisting of diverse drug-like small molecules subjected to a molecular docking screening followed by filtration by the 3D pharmacophore, next was the visual examination. Finally, 6 hit compounds were selected namely, ZINC58717986, ZINC60399606, ZINC58662884, ZINC45988635, ZINC54706757 and ZINC17320595. Molecular dynamics simulation for 100 ns were performed for the selected hits. The hydrogen bonding scanning reveals that residues H41, E166, Q189, T190 and Q192 are the important residues for the binding of 6 hits to the binding pocket of M\textsuperscript{pro}, M49, M165, L167, P168 involved in a hydrophobic interaction while the residues T24, T25, C145, H164, E166, D187 and R188 involved in salt bridge Poisson-Boltzmann surface area (MM-PBSA) binding energy calculations were performed. ZINC58717986, ZINC54706757 and ZINC17320595 showed free binding energy (98.41 ± 0.7, −83.4 ± 0.6, −78.85 ± 0.5 kJ/mol) lower than 13b (−76.67 ± 0.5 kJ/mol). ZINC58717986 showed the lowest free binding energy which suggested the potential activity of this molecule against SARS-CoV-2. The per-residue free energy decomposition analysis suggested that the hot spot residues that contributed to the hits binding are L27, H41, S46, M49, H164, M165, E166, L167, P168, D187, Q189 and Q192. The pharmacokinetics of the selected hits indicated that those molecules possess drug like properties. Finally, 6 hits were selected as the candidates for SARS-CoV-2 inhibitors. ZINC58717986 is the most potential candidate as SARS-CoV-2 inhibitor. The promising 3 hits with lowest free binding energy and good stability within the binding pocket are available and ready to purchase which facilitates further investigation of their inhibitory activity in the future and carrying out further lead optimization and developing selective and potent derivatives as COVID-19 inhibitors.

Acknowledgements

This work used computational facilities of the Advanced Computing Research Centre, and Blue Crystal. University of Bristol, http://www.bris.ac.uk/acrc/. I thank the Advanced Computing Research Centre at the University of Bristol for access to High-performance Computing. The author is grateful to Dr. Richard Session at molecular modelling laboratory, University of Bristol for providing the computing resources and the access to BUDE program. Also, for his valuable scientific discussion with me about the virtual screening process.

Disclosure statement

The author declares no competing financial interest.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Samia A. Elseginy http://orcid.org/0000-0002-0040-0288

References

Al-Khafaji, K., Al-Dhaidahawi, D., & Taskin Tok, T. (2021). Using integrated computational approaches to identify safe and rapid treatment for SARS-CoV-2. Journal of Biomolecular Structure & Dynamics, 39(9), 3387–3395.

Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J., & Hilgenfeld, R. (2002). Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical
papain-like protease: A framework for anti–COVID-19 drug design. *Science Advances*, 6(42), eabd4596. https://doi.org/10.1126/sciadv.abd4596

Sanders, J. M., Monogue, M. L., Jodlowski, T. Z., & Cutrell, J. B. (2020). Pharmacologic treatments for coronavirus disease 2019 (COVID-19): A review. *JAMA*, 323(18), 1824–1836. https://doi.org/10.1001/jama.2020.6019

Shi, J., & Song, J. (2006). The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *The FEBS Journal*, 273(5), 1035–1045. https://doi.org/10.1111/j.1742-4658.2006.05130.x

Somboon, T., Mahalapbutr, P., Sanachai, K., Maitarad, P., Lee, V. S., Hannongbua, S., & Rungrotmongkol, T. (2021). Computational study on peptidomimetic inhibitors against SARS-CoV-2 main protease. *Journal of Molecular Liquids*, 322, 114999. https://doi.org/10.1016/j.molliq.2020.114999

Thorlund, K., Dron, L., Park, J., Hsu, G., Forrest, J. I., & Mills, E. J. (2020). A real-time dashboard of clinical trials for COVID-19. *The Lancet. Digital Health*, 2(6), e286–e287.

Vandenberg, L. N., Colborn, T., Hayes, T. B., Heindel, J. J., Jacobs, D. R., Lee, D.-H., Shioda, T., Soto, A. M., Vom Saal, F. S., Welshons, W. V., Zoeller, R. T., & Myers, J. P. (2012). Hormones and endocrine-disrupting chemicals: Low-dose effects and nonmonotonic dose responses. *Endocrine Reviews*, 33(3), 378–455.

Wang, J. (2020). Fast identification of possible drug treatment of coronavirus disease-19 (COVID-19) through computational drug repurposing study. *Journal of Chemical Information and Modeling*, 60(6), 3277–3286. https://doi.org/10.1021/acs.jcim.0c00179

Wang, H., Dommett, F., & Holm, C. (2010). Optimizing working parameters of the smooth particle mesh Ewald algorithm in terms of accuracy and efficiency. *The Journal of Chemical Physics*, 133(3), 034117.

Zhang, D., Lin, Y., Chen, X., Zhao, W., Chen, D., Gao, M., Wang, Q., Wang, B., Huang, H., Lu, Y., & Lu, Y. (2019). Docking-and pharmacophore-based virtual screening for the identification of novel Mycobacterium tuberculosis protein tyrosine phosphatase B (MptpB) inhibitor with a thiobarbiturate scaffold. *Bioorganic Chemistry*, 85, 229–239.

Zhang, L., Lin, D., Kusov, Y., Nian, Y., Ma, Q., Wang, J., von Brunn, A., Leyssen, P., Lanko, K., Neyts, J., de Wilde, A., Snijder, E. J., Liu, H., & Hilgenfeld, R. (2020). α-Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: Structure-based design, synthesis, and activity assessment. *Journal of Medicinal Chemistry*, 63(9), 4562–4578.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science (New York, N.Y.)*, 368(6489), 409–412.

Zhang, C.-H., Stone, E. A., Deshmukh, M., Ippolito, J. A., Ghahremanpour, M. M., Tirado-Rives, J., Spasov, K. A., Zhang, S., Takeo, Y., Kudalkar, S. N., Liang, Z., Isaacs, F., Lindenbach, B., Miller, S. J., Anderson, K. S., & Jorgensen, L. L. (2021). Potent noncovalent inhibitors of the main protease of SARS-CoV-2 from molecular sculpting of the drug perampanel guided by free energy perturbation calculations. *ACS Central Science*, 7(3), 467–475. https://doi.org/10.1021/acscentsci.1c00039

Zheng, J. (2020). SARS-CoV-2: An emerging coronavirus that causes a global threat. *International Journal of Biological Sciences*, 16(10), 1678–1685.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., … Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7