Drug repurposing for SARS-CoV-2 main protease: Molecular docking and molecular dynamics investigations

Samia E. Omer ¹, Tawasol M. Ibrahim ¹, Omer A. Krar, Amna M. Ali, Alaa A. Makki, Walaa Ibraheem, Abdulrahim A. Alzain ¹*

Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Gezira, Gezira, Sudan

ARTICLE INFO

Keywords:
COVID-19
Main protease
Drug repurposing
Molecular dynamics

ABSTRACT

The current novel corona virus illness (COVID-19) is a developing viral disease that was discovered in 2019. There is currently no viable therapeutic strategy for this illness management. Because traditional medication development and discovery has lagged behind the threat of emerging and re-emerging illnesses like Ebola, MERS-CoV, and, more recently, SARS-CoV-2. Drug developers began to consider drug repurposing (or repositioning) as a viable option to the more traditional drug development method. The goal of drug repurposing is to uncover new uses for an approved or investigational medicine that aren’t related to its original use. The main benefits of this strategy are that there is less developmental risk and that it takes less time because the safety and pharmacologic requirements are met. The main protease (Mpro) of corona viruses is one of the well-studied and appealing therapeutic targets. As a result, the current research examines the molecular docking of Mpro (PDB ID: 5R81) conjugated repurposed drugs. 12,432 approved drugs were collected from ChEMBL and drugbank libraries, and docked separately into the receptor grid created on 5R81, using the three phases of molecular docking including high throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP). Based on docking scores and MM-GBSA binding free energy calculation, top three drugs (kanamycin, sulfonalol and carvedilol) were chosen for further analyses for molecular dynamic simulations.

1. Introduction

SARS-CoV-2 (severe acute respiratory syndrome-Coronavirus-2) is a highly infectious and pathogenic virus that was initially discovered in Wuhan city of China. This COVID-19 outbreak has prompted the development of new therapeutic techniques [1]. There are presently no target-specific medications available for the coronavirus that causes the fatal respiratory disease of COVID19, however, new treatment candidates that target the viral replication cycle are being investigated [2]. The major protease (Mpro) enzyme is a promising therapeutic target because of its critical involvement in viral replication and gene expression [3].

The three major domains of the primary protease correspond to sites 8–101, 102–184, 201–303, respectively. A connecting loop corresponding to locations 185–200 connects Domains II as well as Domain III. His41 and Cys145 form an essential catalytic dyad in the structure of Mpro [4]. Once SARS-CoV-2 has been internalized into the cell, genomic RNA sequence is used as a mold (template) to the forward translation of the pp1a as well as pp1ab polyproteins, which encode several important proteins that are not structural (nsp). Proteases such as main protease (Mpro) and papain-like a protease (Ppro), maintain this translation by processing the 1a and 1ab polypeptide (pp) in a sequentially specified manner to produce 16 distinct nonstructural proteins [5]. The polypeptide is broken down by a papain-like protease to generate nonstructural proteins 1–4. Simultaneously, the Mpro produces the remaining essential nonstructural proteins, such as RNA-dependent RNA polymerase, helicase, and methyltransferase by accurately identifying the sequence Leu-Gln*Ser-Ala-Gly (* denotes the cleavage point) (RdRp). They all play an important part in the cycle of viral infection by producing a replication-transcription complex (RTC) [6]. As a result, the primary protease is a promising therapeutic target for inhibiting the generation of nonstructural viral components and hence the virus life cycle’s replication event. Furthermore, no human protease with identical cleavage specificity has been identified, ruling out the likelihood of

---

¹ Corresponding author. 21111 Barakat Street, Medani, Gezira, Sudan.
E-mail addresses: abdulrahim.altoam@uofg.edu.sd, awh3134@hotmail.com (A.A. Alzain).

Authors contributed equally.
cellular toxicity as a result of the major viral protease’s suppression [7].

Identification of medications for newly discovered diseases are becoming challenging nowadays. Drug repurposing is the most recent technique. The term “drug repurposing” refers to the process of repurposing existing medications for novel restorative motives [8]. It can also provide therapies with well-understood preclinical, pharmacokinetics, pharmacodynamics, and toxicity characteristics that can enter clinical trials promptly [9].

Drug repurposing has long been recognized as a precise computational method for achieving quick and consistent outcomes. This work studied a therapeutic repurposing technique aimed at screening suitable inhibitors of approved medications against COVID 19 main protease, taking into account the structural proteins of the virus as well as their interconnection with the host’s cell-specific receptors.

In this study, a virtual screening of approved drug libraries form ChEMBL and drugbank was performed against the main protease structure of SARS-CoV-2. Moreover, MM-GBSA binding free energy calculations and molecular dynamics (MD) simulations were conducted.

2. Material and method

Schrodinger software was used for the computational investigations (Maestro 12.8, Schrodinger 2021). Academic Desmond v6.5 by D.E. Shaw Research on a Linux system was used to run molecular dynamics simulations.

2.1. Protein preparation

SARS-CoV-2 Mpro (PDB ID: 5R81) crystallographic structure has been obtained from the Protein Data Bank (PDB) as seen in Fig. 1. In order to prepare the protein structure for the docking study, the protein preparation wizard (PPW) module from Schrodinger software package was used. PPW includes three steps; import and refine, review and modify, and protein minimization. Initially, the proteins were pre-processed by adding hydrogen atoms, removal of water molecules beyond 5 Å from the active site and fixing the unresolved residues/side chains. The protein was then energy-optimized, assigning RMSD of 0.30 Å utilizing the OPLS3e force field [10].

2.2. Receptor grid generation

Identification of the binding cavity is crucial in the molecular docking study. The Receptor Grid generation module of Schrodinger suite used with the default option. The centroid of the bound compound was considered for Grid generation [11].

2.3. Ligand preparation

The ligand libraries of 12,432 of approved drugs were downloaded in SDF Format from two approved drugs containing libraries ChEMBL (https://www.ebi.ac.uk/chembl/) and Drugbank (https://go.drugbank.com/) which contain 9923 and 2509 drugs, respectively. To perform the computational studies, LigPrep module of Schrodinger was used to prepare the ligands. Charge neutralization to obtain the biological relevant pH (pH 7), desalting was carried on using Epik v4.7, which depends on Hammett and Taft methodologies. The polar residues’ protonation statuses were adjusted using PROPKA. A maximum one stereoisomer per compound was produced. Energy minimization of the compounds was conducted using OPLS3e force field in order to obtain the low energy conformers [10].

2.4. Ligand docking

The Glide module of Schrodinger has been employed to perform the docking. The grid generation tool was used for the identification of the centroid based on the co-crystallized ligand. The docking procedure was conducted in three steps, which include high throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP). HTVS, and SP use the self-same scoring function, while XP uses extensive sampling and sophisticated scoring function penalize compounds with reduced shape complementarity with the protein binding cavity [12]. Initially, utilizing HTVS, 12,432 approved drugs were docked individually into the receptor grids generated on 5R81. On top molecules, the Standard Precision (SP) docking method was used afterwards. Extra precision (XP) was applied to the top molecules from (SP).

2.5. MM-GBSA binding energies

To assess the strength of interaction between the SARS-CoV-2 Mpro (5R81) and the top selected drugs from (XP) docking. The binding free energy of association has been calculated through Molecular Mechanics energies combined with Generalized Born and Surface Area (MM-GBSA) continuum of Prime module using the default settings [13]. The binding free energies were calculated with OPLS3e force field using the MM-GBSA continuum solvent and VSG 2.0 solvation model utilizing the following formula:

$$
\Delta E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}}
$$

Where, $E_{\text{complex}}$ is energy of the protein-inhibitor complex, $E_{\text{protein}}$ is energy of protein and $E_{\text{ligand}}$ is energy of ligand.

2.6. Molecular dynamics simulations

Academic Desmond v6.5 was used to run MD simulations on selected top docking scored drugs [14]. MD simulation study has used the Glide XP output files as an input. The system was neutralized by the addition of Na+ and Cl- ions, and solvated using the TIP3P water model in an orthorhombic box (10 × 10 × 10). The system was situated at a distance of 10 Å from the edge of the box and LBFGS minimization was conducted with 3 vectors and minimum 10 steepest descent steps until a gradient threshold of 25 kcal/mol/Å was achieved. The maximum iterations during minimization were 2000 and convergence was set at 1.0 kcal/mol/Å. For long-range electrostatic interactions Smooth Particle Mesh Ewald method was used at a tolerance of 1e-09 and a cut-off radius of 9 Å was selected for short-range electrostatic interactions. Before equilibration and MD, the system was minimized and pre-equilibrated using the default settings. NPT ensemble at temperature of 300 K and constant pressure of 1 atm was used throughout the process by Nose-Hoover thermostat and Martyna-Tobias-Klein barostat. The simulation time period was kept for 100 ns. 10,000 frames for each system during the simulations were collected. Experiments were replicated three times.
Lastly, the interaction analysis was performed. The root mean square deviation (RMSD) was calculated for complexes. RMSD shows the average change in the position of the selected atoms in a molecule in comparison to the reference trajectory frame.

3. Results

3.1. Docking studies and MM-GBSA analysis

Glide module for the docking studies uses different scores including docking score, glide score and Emodel to determine the interaction strength between ligand and receptor. Docking score is a scoring function which used to predict the binding affinity of both ligand and target and the more negative the docking score the higher the binding affinity. GlideScore is an empirical scoring function that approximates the ligand binding free energy, it has been optimized for docking accuracy, database enrichment, and binding affinity prediction. Emodel which has a more significant weighting of the force field components (electrostatic and Van der Waals energies), which makes it well suited for comparing conformers, but much less so for comparing chemically distinct species [15]. Therefore, Glide uses Emodel to pick the “best” pose of a ligand (pose selection), and then ranks these best poses against one another with GlideScore. MM-GBSA is used to estimate the relative binding affinity for a list of ligands. As the MM-GBSA binding energies are approximate free energies of binding, a more negative value indicates stronger binding.

As there is no available anti-COVID-19 drugs, molecular docking studies were carried out over 12,432 approved drugs from both ChEMBL database and drugbank on the binding pocket of enzyme COVID-19 (PDB ID: 5R81) in attempt to find suitable candidates for treating COVID-19. Using HTVS mode, a total of 29,765 conformations reflecting various ligand ionization states were screened for the substrate binding site of MPro. The HTVS study revealed that 3258 ligands with docking energies ranging from $-7.001$ to $-8.838$ kcal/mol could bind to Mpro. SP mode was used to re-dock these ligands to the substrate binding site. With docking scores ranging from $-7.000$ to $-9.400$ kcal/mol, 660 ligands demonstrated significant binding affinity for the MPro binding site. XP docking was also done on these top binding ligands with docking scores $\leq -7.000$ kcal/mol. The XP analysis betrayed top fifteen compounds with docking energies score $\leq -7.000$ kcal/mol as depicted in Table 1.

Furthermore, the top docked fifteen compounds were evaluated for their binding free energy using MM-GBSA calculations in order to predict the strength of interactions between the top docked compounds and the MPro. In comparison to the reference values (RZJ; i.e., pre-existing ligand with MPro), all 15 compounds showed good binding free energy score with the MPro binding site ranging from $-35.05$ to $-58.22$ kcal/mol.

Carvedilol, Kanamycin and Sulfinalol were selected for further analysis according to their highest docking score, and low MM-GBSA which were $-52.90$ kcal/mol, $-35.05$ kcal/mol, and $-49.01$ kcal/mol respectively.

3.1.1. Carvedilol

The interaction analysis of Carvedilol-Mpro has revealed the key amino acid residues that found at the substrate-binding site which interacted with the Carvedilol. Anisoole and carbazole rings of Carvedilol are responsible of two Pi-Pi stacking interactions with amino acid HIE41 which are shown in Fig. 2. Also, Carvedilol interacted with two water molecules in the binding site through hydrogen bonds.

3.1.2. Kanamycin

The interaction analysis of Kanamycin with Mpro disclosed that it formed five H-bonds with amino acid residues HIE41, CYS44, SER46, THR190 and ARG188 and two hydrogen bonds with water molecules. Moreover, Kanamycin with HIE41 through formed pi-cation interaction (Fig. 3).

3.1.3. Sulfinalol

Sulfinalol interaction with MPro revealed that it formed two H-bonds with SER46 and GLY143, while forming a Pi-Pi stacking interaction with HIE41 (Fig. 4).

3.2. Molecular dynamics simulations

After the docking tests, the top three medicines (Carvedilol, Kanamycin, and Sulfinalol) were screened for molecular dynamics to determine the binding stability of the docked complexes. To examine the conformational stability of the complexes, the simulation was run for 100 ns. By showing Root Mean Square Deviation (RMSD), which displays the values of bound and unbound ligands in different time intervals and summarizing the conformational changes of the ligands in 100 ns, the information obtained by this trajectory was used to examine the stability of the complexes. Generally, all the three drugs were found forming hydrogen bonds with a key amino acid residue at the substrate binding site of the Mpro (PDB ID: 5R81); the analysis of the interaction pattern covered that this interaction, i.e., the hydrogen bond, is playing a significant role in the stabilization of these drugs during the simulation time.

| No. | Name of compounds | Classification | Docking score (Kcal/mol) | Glide g-score (Kcal/mol) | Glide e-model (Kcal/mol) | Glide energy (Kcal/mol) | MM-GBSA (Kcal/mol) |
|-----|------------------|----------------|--------------------------|--------------------------|--------------------------|-------------------------|-------------------|
| 1   | Carvedilol       | B-blocker      | -5.132                   | -5.132                   | -32.414                  | -41.378                 | -35.85            |
| 2   | Kanamycin        | Aminoglycoside | -5.861                   | -7.984                   | -59.834                  | -50.308                 | -35.05            |
| 3   | Sulfinalol       | B-blocker      | -7.825                   | -7.825                   | -66.545                  | -51.064                 | -49.01            |
| 4   | Arbutamine       | B-blocker      | -7.711                   | -7.718                   | -57.788                  | -43.112                 | -45.76            |
| 5   | Valopicitabine   | Antiviral      | -7.550                   | -7.794                   | -56.803                  | -44.327                 | -23.58            |
| 6   | Midodrine        | Alpha Ad agonist | -7.535                  | -7.643                   | -53.224                  | -43.259                 | -43.49            |
| 7   | Bifonazole       | Imitazole Antifungal | -7.507              | -7.843                   | -59.307                  | -46.301                 | -58.22            |
| 8   | Carazolol        | B-blocker      | -7.460                   | -7.462                   | -54.353                  | -44.252                 | -45.76            |
| 9   | Demeclocycline   | Tetracycline Antibiotic | -7.273              | -7.469                   | -61.837                  | -46.714                 | -43.94            |
| 10  | Butocrolol       | B-blocker      | -7.250                   | -7.259                   | -60.709                  | -48.004                 | -52.40            |
| 11  | Tobramycin       | Aminoglycoside | -7.157                   | -7.428                   | -68.735                  | -51.345                 | -32.59            |
| 12  | Flavoxate        | Anticholinergic | -7.169                   | -7.126                   | -58.002                  | -45.733                 | -43.46            |
| 13  | Carazolol        | B-blocker      | -7.084                   | -7.085                   | -66.035                  | -48.570                 | -52.99            |
| 14  | Dapaconzole      | Antifungal     | -7.041                   | -7.248                   | -60.414                  | -45.343                 | -52.79            |
| 15  | Kalafungin       | Antibiotic     | -7.020                   | -7.053                   | -55.283                  | -40.041                 | -43.87            |
3.2.1. Carvedilol

The RMSD plot of Mpro and Carvedilol is shown in (Fig. 5A). Carvedilol-Mpro complex remained stable for 22ns after which the protein undergoes structural conformation while the ligand was more stable with RMSD ligand value below 2 Å. The interaction pattern of Carvedilol with Mpro that occurred during the simulation revealed that Carvedilol formed various interactions such as hydrogen bonds hydrophobic interactions, ionic interactions and water bridges with different substrates in the binding sites (Fig. 5B). During the simulation, the amino acid residues involved in the Mpro-Carvedilol complex’s stability were THR-24, THR-25, HIS-41, THR-45, SER-46, MET-49, LEU-50, ASN-142, GLU-166, PRO-168, GLN-189, THR-190, ALA-191, GLN-192 and ALA-193. Importantly, PRO-186, GLN-189, THR-190, ALA-191, GLN-192 and ALA-193 were found to form hydrophobic, hydrogen bonds and water bridge interactions.

3.2.2. Kanamycin

The RMSD plot of Mpro and Kanamycin is shown in (Fig. 6A). Kanamycin-Mpro complex remained stable from 8 ns to 22 ns. Thereafter, kanamycin was more stable than the protein with ligand RMSD value of 1.2 Å for the rest of the simulation. Kanamycin generated multiple interactions with different substrate binding site residues, including hydrogen bonds, hydrophobic interactions, ionic contacts, and water bridges, according to the study of the interaction pattern between Kanamycin and Mpro (Fig. 6B). During the period of simulation, amino acid residues that which stabilized the Mpro-Kanamycin complex in various ways were THR-24, THR-25, THR-26, HIS-41, CYS-44, THR-45, SER-46, GLU-47, PHE-140, LEU-141, ASN-142, GLY-143, SER-144, HIS-163, HIS-164, GLU-166, LEU-167, PRO-168, HIS-172, ARG-188, GLN-189, THR-190 and GLN-192. Notably, GLU-47, ASN-142, SER-144, GLU-166 and GLN-189 were the most stable interaction with kanamycin. Moreover, GLU-166 formed hydrogen bonds and Water Bridge with kanamycin of 40% during simulation time.

3.2.3. Sulfinalol

The RMSD plot of Mpro and Sulfinalol is shown in (Fig. 7A). The
RMSD graph of Sulfinalol shows sharp deviation during the initial phase of simulation, however, it was stabilized from 7ns to 28 ns. Moreover, the RMSD value of the Sulfinalol was more stable than the protein with ligand RMSD value below 2.3 Å during the simulation period. Sulfinalol interacted through hydrogen bonds, hydrophobic contacts, ionic interactions, and water bridges with the residues of distinct substrate binding sites, according to the examination of the interaction between

Fig. 4. 2D and 3D docking interaction of Sulfinalol with 5R81.

Fig. 5. (A) RMSD calculations showing the conformational deviation of drugs-protein complexes: the drugs were represented in different colors as (Carvedilol (red), and protein (blue)). (B) Participation of different amino acid residues during the simulation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 6. (A) RMSD calculations showing the conformational deviation of drugs-protein complexes: the drugs were represented in different colors as (Kanamycin (red), and protein (blue)). (B) Participation of different amino acid residues during the simulation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
Sulfinalol and Mpro that occurred during the simulation (Fig. 7B). The important amino acid residues involved during the period of simulation include THR-24, THR-25, HIS-41, THR-45, SER-46, GLU-47, MET-49, LEU-50, ASN-142, MET-165, GLU-166, GLN-189, THR-190 and GLN-192. GLU-47 played a key role in the formation of a stable Mpro-Sulfinalol complex through hydrogen bonds and Water Bridges with the protein.

4. Discussion

Molecular docking, molecular dynamics and other in silico techniques are widely used approaches to understand ligand-receptor interactions during drug discovery process [16–20]. Virtual screening and drug repurposing have been used in drug discovery against emerging and fatal diseases including SARS CoV proteases [21–26]. Also, drug repurposing acts as a very effective drug discovery technique as it decreases the costs and shortens the time required compared with de novo drug discovery [15]. In this study, we utilized several in silico approaches to identify potential drugs that could inhibit the Mpro of SARS-CoV-2 including the use of molecular dynamics and quantum mechanics. In this study, three drugs were selected namely carvedilol, kanamycin and sulfinalol with docking score values of −7.995, −7.836 and −7.825 kcal/mol, respectively. Moreover, the three displayed favorable MM-GBSA free binding energy values ranging from −35.05 to −52.9 kcal/mol. The molecular dynamics showed stable interaction for 100 ns.

Kanamycin is an antibiotic that belongs to the aminoglycoside’s family, that function to suppress the synthesis of the protein by firmly adhering to the preserved A16S rRNA region found in the 30S ribosomal subunit. The limitations presented by multidrug-resistant organisms on the treatment of severe bacterial infections have resulted in a rise in clinical usage of aminoglycosides in current years [27]. Kanamycin has been found to possess a favorable profile of interaction the receptor binding domain (RBD) of spike protein S1 and ACE2 which is vital to the integration of coronavirus RNA into the host cell [28]. Moreover, kanamycin reported interacted COVID-19 main protease (PDB ID: 6IU7) [29,30].

Sulfinalol is a beta-adrenoceptor antagonist which represents the most important class of medications in the treatment of cardiovascular disorders (such as ischemic heart diseases, hypertension, and particularly in heart failure) [31]. The mechanism of action in cardiovascular disorders is related to the blockade of catecholamine endogenous responses (mainly at beta1-adrenoceptors in the heart), whereas the worrying side effects of bronchoconstriction resulting from antagonism or blockade of airway beta2-adrenoceptors [32]. Beta blockers have been found to reduce the coronavirus cellular invasion by reducing angiotensin converting enzyme 2 receptor expression as well as cluster of differentiation 147 in different cells in the human body [33].

Carvedilol is a drug that belongs to the nonselective beta-adrenergic blocker class that have no intrinsic sympathomimetic activity and with alpha 1-adrenergic receptors antagonistic activity. Moreover, carvedilol was found to inhibit RNA-dependent RNA polymerase of SARS-CoV-2 [34].

5. Conclusion

A library of FDA-approved drugs from: ChEMBL and drug bank was screened against the main protease enzyme, Mpro (PDB ID: 5R81) of the virus in order to block the virus’s reproduction using in silico techniques. Among these drugs, carvedilol, kanamycin and sulfinalol yield the best docking score of −7.995, −7.836 and −7.825 kcal/mol, respectively. Also, the three drugs showed favorable MM-GBSA free binding energy values ranging from −35.05 to −52.9 kcal/mol. The protein-ligand interactions obtained from docking were validated using molecular dynamics (MD) study. The MD simulation was performed for 100 ns using Academic Desmond. The fluctuations during the whole simulation were within the standard range of RMSD 1–3 Å, thus indicating their stability. Further, more in vitro and in vivo studies are required to validate the results.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We acknowledge Mme Katia Dekimeche from Schrodinger for her support and help.

References

[1] N. Kapoor, S.M. Ghorai, P.K. Kushwaha, R. Shukla, C. Aggarwal, R. Bandlichor, Plausible mechanisms explaining the role of cucurbitacin as potential therapeutic drugs against coronavirus 2019, Inf. Med. Unlocked 21 (Jan. 2020) 100484, https://doi.org/10.1016/j.imu.2020.100484.
[2] R. Banerjee, L. Perera, L.M.V. Tillekeratne, Potential SARS-CoV-2 main protease inhibitors, Drug Discov. Today 26 (3) (2021), https://doi.org/10.1016/j.drudis.2020.12.005. Elsevier Ltd, pp. 804-816, Mar. 01.
[3] A. Galalawat, et al., Structure-based virtual screening to discover potential lead molecules for the SARS-CoV-2 main protease, J. Chem. Inf. Model. 60 (12) (Dec. 2020) 5781-5793, https://doi.org/10.1021/acs.jcim.0c00546.
[4] A. Khan, et al., Combined drug repurposing and virtual screening strategies with molecular dynamics simulation identified potent inhibitors for SARS-CoV-2 main protease (3CLpro), J. Biomol. Struct. Dyn. (2020), https://doi.org/10.1080/07391102.2020.1779126.
