



OPEN

Novel sofosbuvir derivatives against SARS-CoV-2 RNA-dependent RNA polymerase: an in silico perspective

Abdulwahed Alrehaily¹, Abdo A. Elfiky^{2✉}, Ibrahim M. Ibrahim², Mohamed N. Ibrahim³ & Amr Sonousi^{4,5}

The human coronavirus, SARS-CoV-2, had a negative impact on both the economy and human health, and the emerging resistant variants are an ongoing threat. One essential protein to target to prevent virus replication is the viral RNA-dependent RNA polymerase (RdRp). Sofosbuvir, a uridine nucleotide analog that potently inhibits viral polymerase, has been found to help treat SARS-CoV-2 patients. This work combines molecular docking and dynamics simulation (MDS) to test 14 sofosbuvir-based modifications against SARS-CoV-2 RdRp. The results reveal comparable (slightly better) average binding affinity of five modifications (compounds 3, 4, 11, 12, and 14) to the parent molecule, sofosbuvir. Compounds 3 and 4 show the best average binding affinities against SARS-CoV-2 RdRp (-16.28 ± 5.69 and -16.25 ± 5.78 kcal/mol average binding energy compared to -16.20 ± 6.35 kcal/mol for sofosbuvir) calculated by Molecular Mechanics Generalized Born Surface Area (MM-GBSA) after MDS. The present study proposes compounds 3 and 4 as potential SARS-CoV-2 RdRp blockers, although this has yet to be proven experimentally.

SARS-CoV-2 has infected an estimated 769 million people, with about 6.9 million deaths since its outbreak in 2019 (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019>). The disease symptoms can range from mild symptoms such as cough, fever, fatigue, loss of smell, and loss of taste to more severe complications such as severe pneumonia, dyspnea, and organ dysfunction¹. SARS-CoV-2 (also known as Covid-19) is a single-stranded RNA virus in the Coronaviridae family². SARS-CoV-2 has many known variants, such as alpha, beta, gamma, delta, and omicron variant³. In addition, other variants of interest include, epsilon, zeta, eta, theta, iota, kappa, lambda, and mu. Although disease mortality and morbidity have decreased with the development of vaccination, the emergence of different variants able to overcome the vaccines has caused recent waves and breakthrough infections⁴.

Several viral targets have been used in COVID-19 drug development, and these encompass crucial stages of the virus replication cycle⁵⁻⁹. For example, the spike protein is necessary during viral recognition and host-cell entry, the main protease (M^{pro}) during the proteolytic activation stage, and RNA-dependent RNA polymerase (RdRp) during the transcription stage⁵. Of these targets, RdRp has been shown to be the most appealing for drug discovery as it is a very conservative non-structural protein functions as RNA polymerase, which either copies the RNA of the virus for replication or produces a messenger RNA¹⁰. The consecutive aspartates active site were found and targeted in many viral RNA polymerases. Therefore, inhibition of RdRp prevents these essential steps in the lifecycle of SARS-CoV-2 and, thus, prevents replication and propagation of the virus.

Sofosbuvir is a uridine nucleotide analog that potently inhibits viral polymerases. It is an anti-RdRp that is found useful in treating SARS-CoV-2 patients¹¹. Sofosbuvir is a prodrug that is hydrolyzed by liver enzymes after absorption to form the monophosphate uridine analog, which is further phosphorylated to form the active triphosphate form. Sofosbuvir was approved by FDA against hepatitis C Virus (HCV) in the year 2013. It is administered orally in combination with ribavirin to treat HCV of genotypes 2 and 3¹². We choose this drug as it was deeply studied with millions of HCV patients worldwide during the past 10 years.

¹Biology Department, Faculty of Science, Islamic University of Madinah, 42351 Madinah, Saudi Arabia. ²Biophysics Department, Faculty of Science, Cairo University, Giza, Egypt. ³Clinical Laboratories Department, College of Applied Medical Sciences, Jof University, Qurrayat, Saudi Arabia. ⁴Pharmaceutical Organic Department, Faculty of Pharmacy, Cairo University, Giza, Egypt. ⁵University of Hertfordshire Hosted By Global Academic Foundation, New Administrative Capital, Cairo, Egypt. ✉email: abdo@sci.cu.edu.eg

The SARS-CoV-2 RNA genome frequently mutates due to its rapid replication and lack of error proofreading by the viral RNA polymerase, which results in the emergence of drug resistance (e.g., the P323L mutation where the substitution of proline [Pro] to leucine [Leu] occurred at amino acid position 323 of RdRp)^{13,14}. Therefore, continuous development of drugs is required to overcome emerging viral resistance.

In this study, our strategy was to carry out rational modifications to the general features of the sofosbuvir scaffold and to model the effect of such alterations on RdRp binding. Modifications were categorized into two groups based on the site of transformation: sugar modifications and uracil modifications. Sugar modifications include studying the effect of extending position 3 with the ethyl amino group or changing position 3 to the fluoro group (structures 1 and 5, Fig. 1)¹⁵. To mimic the binding of the triphosphate, an amino group, or polyhydroxylated-substituted amino groups, substitutes the hydroxyl of the ribosyl 5-position. The uracil modifications of sofosbuvir include attaching the ethylamino group at uracil N-3 and substituting the carbonyl oxygen at position 4 with sulphur or amino or aminoethyl amino groups (structures 6–9, Fig. 1)¹⁶. Other modifications involve substitutions at the carbonyl oxygen at position 2 with thio, benzylthio, and dihydroxy benzylthio groups (structures 10–12, Fig. 1). The effect of attaching hydrazine or phenylhydrazine at position 6 was also explored (structures 13 and 14, Fig. 1)¹⁷. To evaluate the proposed modifications, all analogs were docked in the catalytic site of RdRp, and the molecular dynamics of the best hits were determined to study the stability of the analogs in the gorge of the enzyme.

Molecular dynamics simulation (MDS) was successful in repurposing and prediction of new therapeutics against SARS-CoV-2 proteins^{18,19}. The RdRp was previously studied using the MDS technique and was successful in predicting the potential usefulness of many drugs such as remdesivir against COVID-19 pandemic^{20–22}. In the current study we use the same technique to test new scaffold molecules based on sofosbuvir.

Materials and methods

Protein and ligands preparation

As previously reported, the PDB: ID 7BV2 was downloaded from the Protein Data Bank (www.pdb.org)²³ and was then prepared by removing unwanted molecules and adding missing Hydrogen atoms using PyMOL software^{24,25}. As a result, the prepared protein target contained only nonstructural protein 12, nonstructural protein 7, and nonstructural protein 8 as a complex, the active complex of the viral polymerase²⁶. We choose this structure as it is of moderate resolution (2.5 Å) and containing the polymerase in conjunction with nsp7 and nsp12, representing the complete system.

The drug was modified and optimized using SCIGRESS 3.2 software²⁷. First, the structures were minimized using the classical MM3 force field, followed by the semiempirical parameterization method (PM6) in water²⁸. Next, the vibrations were calculated at the same level of computation to ensure their stability (not a transition

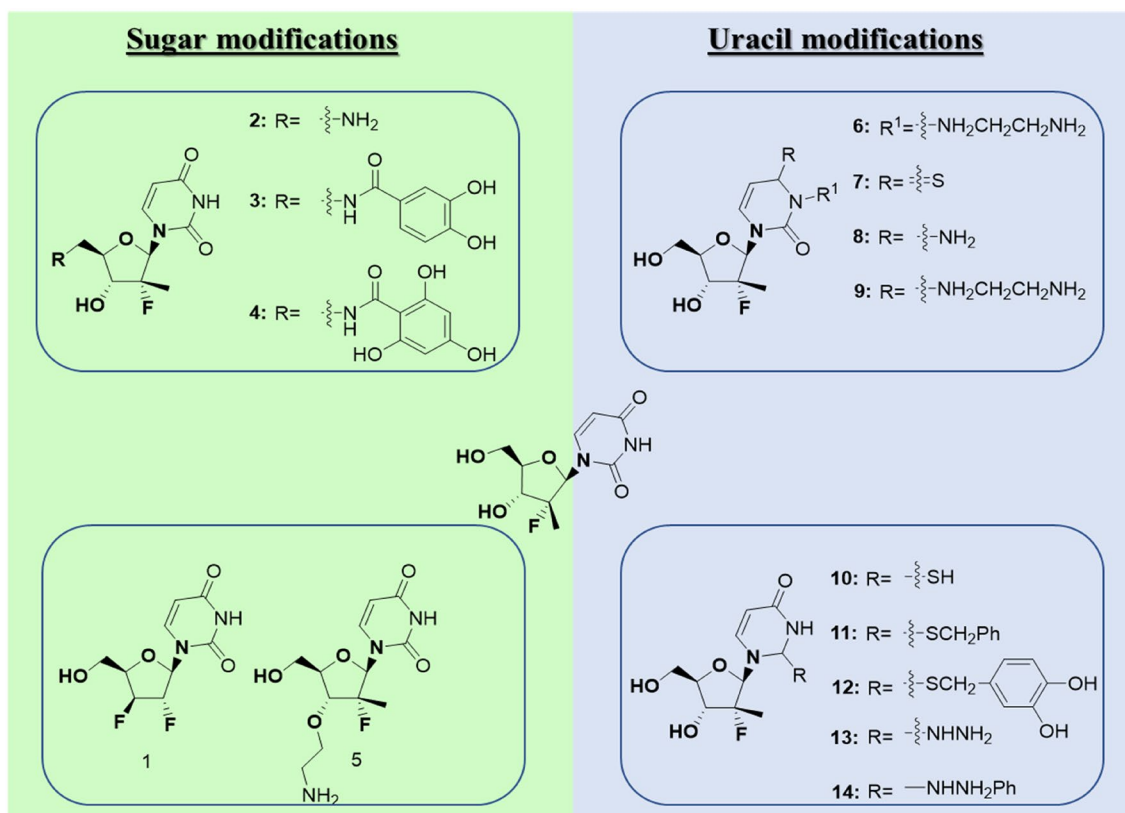


Figure 1. Sugar and uracil modifications to the sofosbuvir drug.

state). Finally, structures were optimized using the quantum mechanical method, which is based on density functional theory (DFT), using the B3LYP functional^{29,30}. Again, geometry was judged by their infra-red spectrum calculated at the same level of theory. We ensure no negative vibrations (transition states) are found in the optimized compounds.

Molecular docking

The clustered trajectories (three clusters) obtained from the previous study were used as possible protein conformations for docking²⁴. The clustering was performed using Chimera using default parameters. Prior to the simulation any water, ions, or ligands were removed. AutoDock Vina (7) was used for the flexible docking of the modified compounds with each cluster representative. The compounds were prepared by adding Gasteiger charges and converted to PDBQT file format using AutoDock Tools software³¹. The protein was also prepared by adding partial charges and converted to PDBQT file format. The grid box dimensions were set to 40 Å in each direction for each docking experiment to cover the active site residues. The box centers for the docking with cluster 1, cluster 2, and cluster 3 representative structures (after clustering the trajectories of the protein target) were set to (−5.194, −9.118, 0.000), (−6.871, −10.988, −0.146), and (−5.558, −9.933, −0.090), respectively. According to the calculated binding affinity, the best complex (having the lowest binding energy values) and the parent compound, sofosbuvir, were subjected to another MDS run of 100 ns. We compared the binding behaviour of the tested compounds to the reference drug (sofosbuvir) using the same protocol.

Molecular dynamic simulation (MDS)

The files were prepared for MDS using the CHARMM-GUI web server^{32,33} (<https://www.charmm-gui.org/>). The system was solvated using the TIP3P water model in a cubic box with a padding of 10 Å and the box size was 121 × 121 × 121 Å³ then the systems were neutralized by adding 0.154 M NaCl which resulted in a total of around 166,000 Atoms. The temperature was set to the physiological value of 37 °C (310 K), and the CHARMM36 force field was utilized. CHARMM36 accurately predict protein parameters and is utilized by many MDS software when dealing with proteins. It could predict backbone & side-chain scalar couplings, residual dipolar couplings, and relaxation & side-chain order parameters³⁴.

Nanoscale molecular dynamics (NAMD) 2.13 software was used to perform the MDS calculation³⁵. Before the production run, the minimization and equilibration steps were completed. The system was minimized for 10,000 steps and then equilibrated for 0.25 ns in an ensemble of a constant number of atoms, a constant volume, and a constant temperature (NVT). The temperature was maintained using Langevin dynamics^{35,36}. Then, the production run was started for 100 ns in the NVT ensemble. Finally, the MDS trajectories were analyzed using in-house codes and visualizing molecular dynamics (VMD) 1.9.3 software³⁷. In addition, the trajectory of the best two complexes and the reference compound were clustered using TtClust and representative frames for each cluster were obtained. Protein–Ligand Interaction Profiler (PLIP) was used with these frames to detect the interacting amino acids^{38–40}.

Binding energy calculations during the molecular dynamics simulation for the best hits

MMPBSA.py module implemented in the AmberTools 21 package was utilized to determine the binding affinity for the best-hit compounds (3 and 4) and the positive control sofosbuvir^{41,42}. In addition, decomposition analysis was calculated to determine the contribution of amino acids to total binding energy. The salt concentration and solvation method (igb) were set to 0.154 M and 5, respectively. The internal and external dielectric constants were set to 1.0 and 80.0, respectively, and other options were assigned to the default values. The molecular mechanics-generalized Born surface area (MM-GBSA) approach is depicted in Eq. (1).

$$\Delta G = \langle G_{\text{complex}} - (G_{\text{receptor}} + G_{\text{ligand}}) \rangle \quad (1)$$

where $\langle \rangle$ represents the average of the enclosed free energies of complex, receptor, and ligand over the frames used in the calculation. In our approach, we used the whole trajectory (1000 frames). Different energy terms can be calculated according to Eqs. (2) – (6) as follows:

$$\Delta G_{\text{binding}} = \Delta H - T\Delta S \quad (2)$$

$$\Delta H = \Delta E_{\text{gas}} + \Delta E_{\text{sol}} \quad (3)$$

$$\Delta E_{\text{gas}} = \Delta E_{\text{ele}} + \Delta E_{\text{vdw}} \quad (4)$$

$$\Delta E_{\text{solv}} = E_{\text{GB}} + E_{\text{SA}} \quad (5)$$

$$E_{\text{SA}} = \gamma \cdot \text{SASA} \quad (6)$$

where ΔH is the enthalpy calculated from gas-phase energy (E_{gas}) and solvation-free energy (E_{sol}), $-T\Delta S$ is the entropy contribution to the free binding energy, and E_{gas} is composed of electrostatic term E_{ele} and van der Waals term E_{vdw} . E_{sol} can be calculated from the polar solvation energy (E_{GB}) and nonpolar solvation energy (E_{SA}), which are estimated from the solvent-accessible surface area. The entropy contribution was calculated using the interaction entropy method using the last 90 ns^{43,44}.

Results and discussion

Sofosbuvir is a uridine nucleotide analog that potently inhibits HCV NS5B polymerase. Sofosbuvir is a prodrug that is hydrolyzed by liver enzymes after absorption to form the monophosphate uridine analog, which is further phosphorylated to form the active triphosphate form. Sofosbuvir received FDA approval for use against HCV in 2013; it is administered orally in combination with ribavirin for the treatment of genotypes 2 and 3⁴⁵.

Molecular docking

The average docking scores of the suggested modifications of sofosbuvir (Fig. 1) and the positive control compound (sofosbuvir triphosphate), calculated with AutoDock Vina, are presented in the bar graphs in Fig. 2. The docking experiments were performed with the three different conformations of the SARS-CoV-2 RdRp after 100 ns MDS²⁴. The error bars represent the standard deviations. As shown in Fig. 2, the fourteen compounds showed average binding energies to the SARS-CoV-2 RdRp active site comparable to that of the positive control and parent compound, sofosbuvir. Additionally, five compounds (3, 4, 11, 12, and 14) showed better average values than sofosbuvir in binding SARS-CoV-2 RdRp, while compounds 3 and 4 were the best hits in binding the viral polymerase.

Table 1 summarizes the formed interactions that were established upon docking these fourteen compounds in the active site of the SARS-CoV-2 RdRp. Two types of interaction stabilized the complexes; Hydrogen bonding (H-bonds) and hydrophobic interactions. On average, 2.9 ± 1.8 H-bonds and 1.8 ± 0.8 hydrophobic contacts were established upon docking the fourteen ligands against the SARS-CoV-2 RdRp active site. In contrast, five H-bonds were formed for sofosbuvir, and only one hydrophobic contact was established upon docking. For the best two compounds, 3 and 4, seven interactions were established between the ligands and the protein residues. Additionally, the following six residues were the most significant contributors in the fourteen compounds against the SARS-CoV-2 RdRp active site: D618 (8), D761 (5), A762 (7), W800 (5), E811 (4), and C813 (4), with the numbers in brackets, indicating the total number of interactions (as shown in Table 1, bolded residues). All these residues were found in the active site pocket of the RdRp, as previously reported⁴⁶.

Molecular dynamics of the sofosbuvir- and the best compounds -RdRp complexes

MDS was used to assess protein system stability and to assess the binding affinities during system equilibration⁴⁷. We simulated the protein–ligand complexes for both sofosbuvir (positive control) and compounds 3 and 4 for 100 ns MDS each. Figure 3 shows the root-mean-square deviation (RMSD) in Å (A), the radius of gyration (RoG) in Å (B), the surface accessible surface area (SASA) in Å² (C), and the total number of formed H-bonds (D) versus the simulation time in nanoseconds. Sofosbuvir–RdRp complexes are shown in red, while compounds 3 and 4–RdRp complexes are shown in gray and blue, respectively. Based on the RMSD curves (Fig. 3A), all the complexes were stable, and the systems were slowly equilibrated during the simulation period with RMSD values between 2 and 3.2 Å during the second half of the simulation. Additionally, the systems were stable during the simulation with RoG values around 31 Å (Fig. 3B), SASA values between 44,000 and 47,000 Å² (Fig. 3C), and total system H-bonds between 1600 and 1750 (Fig. 3D). The per-residue root-mean-square fluctuations

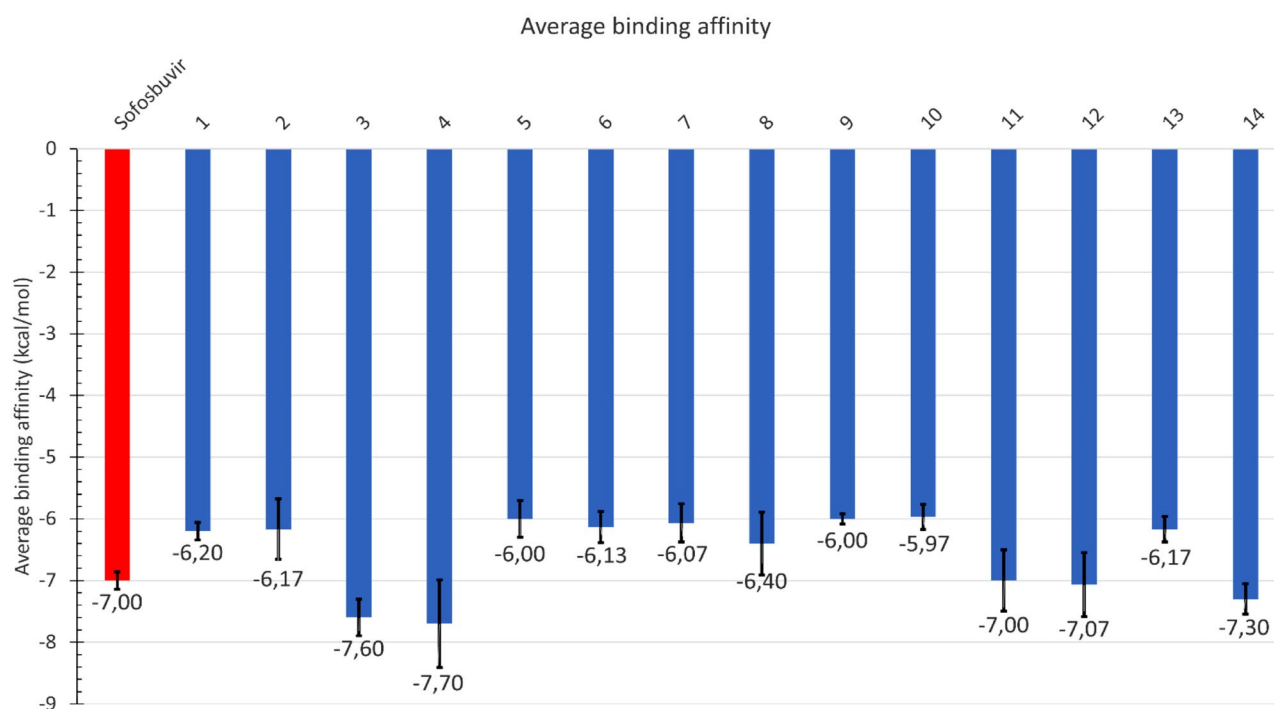


Figure 2. The average binding affinities (in kcal/mol) of sofosbuvir (red) and its fourteen modifications (blue) against SARS-CoV-2 RdRp.

| Ligand | Hydrogen bonds | | Hydrophobic interactions | |
|------------|----------------|---|--------------------------|---|
| | No | Residues involved | No | Residues involved |
| Sofosbuvir | 5 | A550, K551, W800 , H810, and C813 | 1 | D618 |
| 1 | 3 | W800 , and C813(2) | | |
| 2 | 1 | C813 | 2 | D618 and D761 |
| 3 | 6 | S549, A550, K551, W800 , C813, and R836 | 1 | D618 |
| 4 | 7 | K545, R555, Y619, R624, S759, D760 , and D761 | | |
| 5 | 2 | A762 and W800 | 1 | D618 |
| 6 | 1 | A762 | 1 | D618 |
| 7 | 1 | I548 | 3 | I837, A840, and L862 |
| 8 | 3 | A762(2) and W800 | 2 | D618 and D761 |
| 9 | 3 | A762 , H810, and R836 | 2 | D618 and E811 |
| 10 | 2 | W800 and R836 | 1 | D761 |
| 11 | 2 | S814 and R858 | 3 | I548, R836, and D865 |
| 12 | 3 | R555 and A762(2) | 1 | D618 |
| 13 | 5 | W617, K798(2), C799, and E811 | 3 | D618 , D761 , and E811 |
| 14 | 2 | K551 and K798 | 1 | E811 |

Table 1. Detailed interactions established upon docking the sofosbuvir and its derivatives (1–14) against SARS-CoV-2 RdRp in complex with NSP8 and NSP12 retrieved from PLIP webserver and PyMOL software. Bold residues indicate the residues most reported to form interactions with the ligands.

(RMSF) in Å are shown in Fig. 3E. The active site residues, which are marked on the RMSF curve, were stable (RMSF > 2Å). In general, the RMSF showed stable protein with minimal high fluctuation regions (RMSF < 3Å), as reported previously⁴⁸. This indicates the stability of the complexes during the simulation period. The ligand-RMSD (Fig. 3F) also indicate stable complexes during the simulation period.

The Fig. 4 shows the results of trajectory clustering and the interacting amino acids as detected by PLIP. As can be seen in the representative frames of sofosbuvir, D760 and D761 are nearly common in all representative frames and forming three to four H-bonds. Moreover, K621 forms a salt bridge with sofosbuvir. For compound 3, the first two cluster representatives show V495 and Y516 forming one hydrophobic contact and one H-bond, respectively. In addition, the Fluorine atom form a halogen bond with Q492. For compound 4, there is no common amino acids in the representative frames, however, K545 forms a π -cation interaction.

The binding energies of the top hits' (compounds 3 and 4) complexes with RdRp were compared to the sofosbuvir-RdRp complexes during the simulation period using the molecular mechanics- generalized Born surface area (MM-GBSA) method. All the equilibrated region of trajectories (the last 90 ns) were used for the MM-GBSA calculations (stride = 1), and the values are listed in Table 2. The residual contribution to binding is listed in the table, where highly contributed residues (binding energy > -0.85 kcal/mol) are shown in bold. In all complexes, the ligand (LIG) was the most contributed element for the binding with values between -7.78 and -10.10 kcal/mol. The average total binding energy for the compounds 3- and 4-RdRp complexes (-16.28 ± 5.69 and -16.25 ± 5.78 kcal/mol) is comparable to that of the average binding energy of sofosbuvir-RdRp complexes (-16.20 ± 6.35 kcal/mol). However, the entropy contribution of the reference compound is higher than that of the two proposed compounds. This indicates the feasibility of using these candidates as anti-SARS-CoV-2 RdRp after toxicity testing and experimental validation.

ADMET prediction

With the help of the pkCSM web server⁴⁹, the ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties of the three compounds were predicted. Table 3 shows the predicted values for each property. The three compounds show similar water solubility (log(mol/L)), with compound 4 being the more soluble one (-2.84 log(mol/L)), followed by compound 3 (-2.524 log(mol/L)), and then sofosbuvir (-2.271 log(mol/L)), which are in the same region as other druggable compounds⁵⁰. On the other hand, the Caco2 permeability shows a similar trend, with compound 4 as the most permeable (58.2%), followed by compound 3 (49.9%), and then sofosbuvir (28.7%).

The fraction unbound indicates the predicted fraction of compound that will be unbound to serum proteins. High unbound fraction, low blood-brain barrier (BBB) permeability (log BBB), and low central nervous system (CNS) permeability (log CNS) indicate a good distribution and low permeability to the BBB and CNS. All three compounds show low BBB permeability as they have values < -1 and low CNS permeability (values < -3). However, the fraction unbound shows a better value for the reference compound in comparison with the two compounds.

Inhibitors of Cytochrome P450 can activate the drug metabolism and, therefore, can remove the compound from the market. All compounds were predicted to not act as inhibitors of different isoforms (CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4).

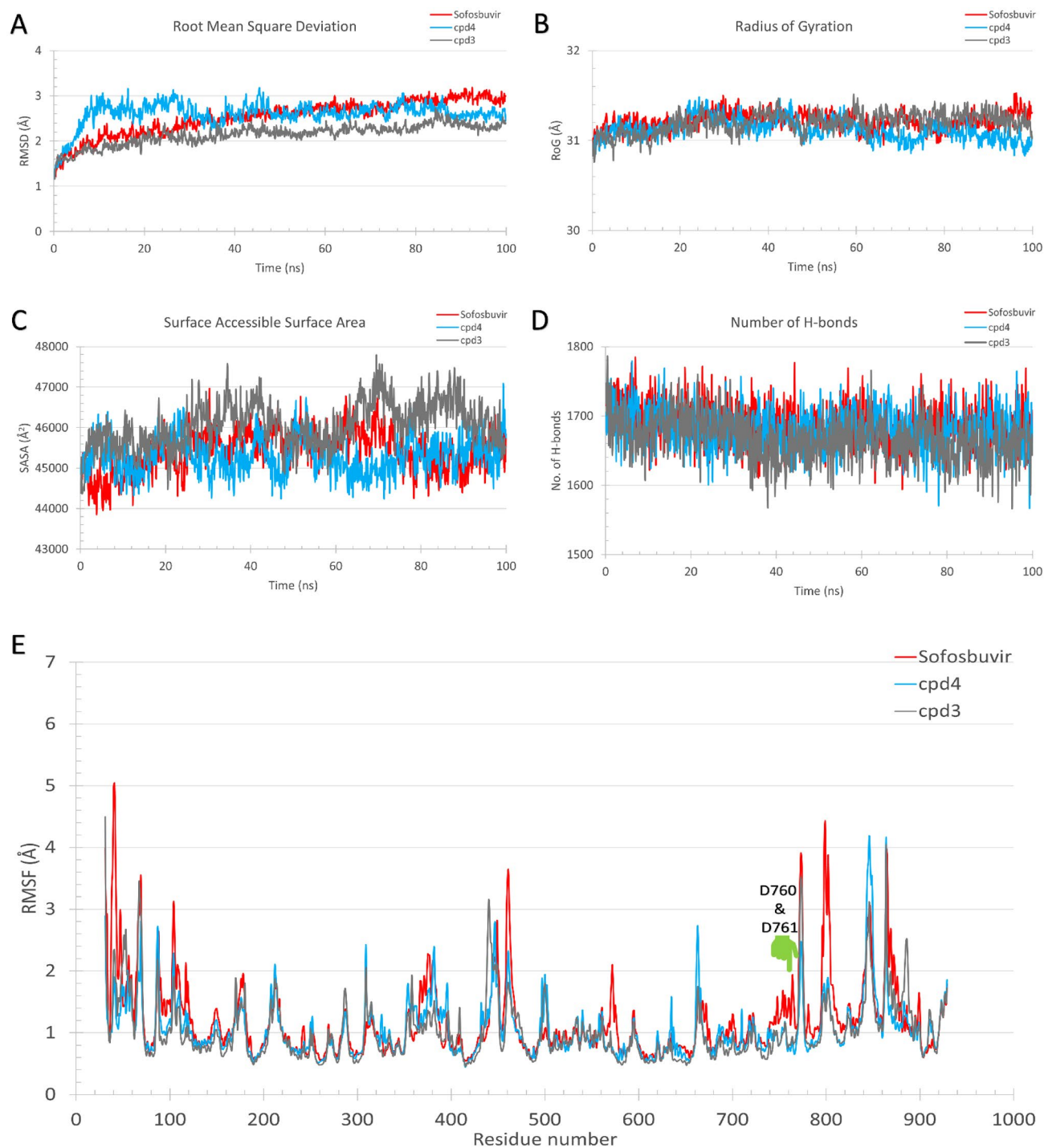


Figure 3. The molecular dynamics simulation analysis of the three best-hit complexes (4-RdRp) and the three sofosbuvir-RdRp complexes. (A), (B), (C), and (D) are the root-mean-square deviation (RMSD), the radius of gyration (RoG), the surface accessible surface area (SASA), and the total number of formed H-bonds versus the simulation time in ns. (E) The pre-residue root-mean-square fluctuations (RMSF) of the six complexes during the simulation period. The position of active site residues (D760 and D761) is marked in green. (F) The ligand RMSD versus the simulation time in ns.

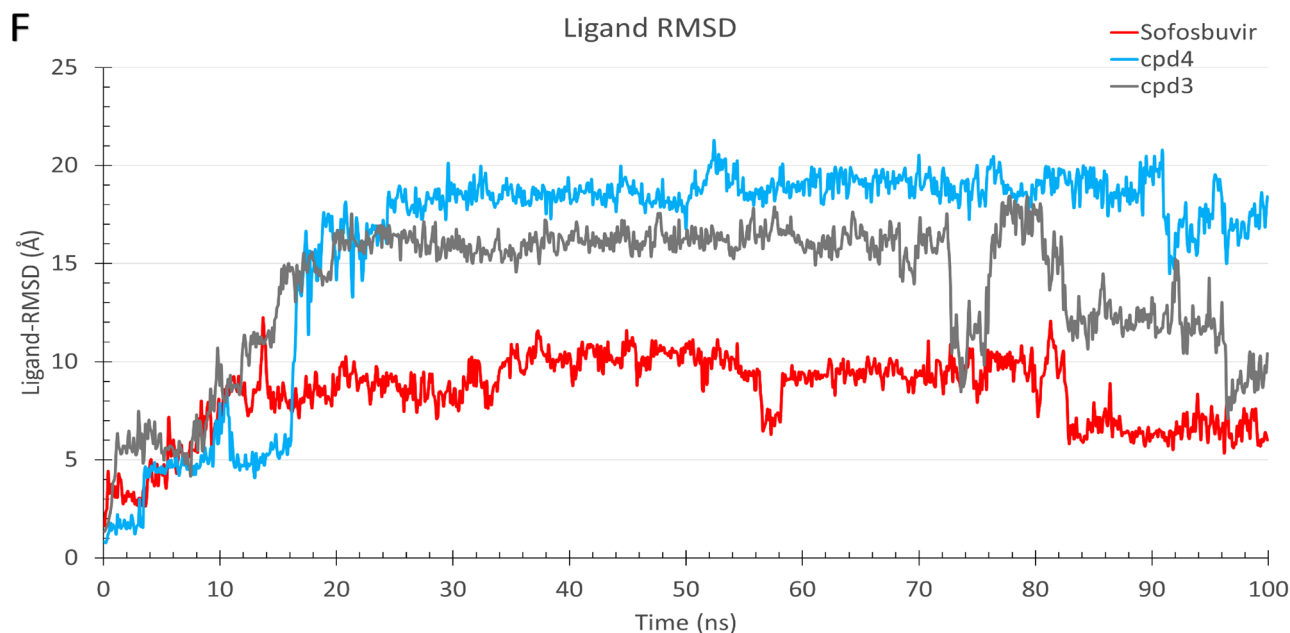


Figure 3. (continued)

For excretion, the model predicts that all compounds are not substrates for renal organic cation transporter 2. The interaction with this transporter helps in the clearance of the compound and may produce adverse interactions; therefore, a negative prediction is considered good.

Finally, four indicators are used to predict the toxicity of the compounds. AMES toxicity is a test that indicates whether the compound is a carcinogen. Inhibition of hERG I/II is the principal cause of fatal ventricular arrhythmia and has resulted in the withdrawal of many substances. As its name implies, hepatotoxicity indicates whether the compound may disrupt the normal function of the liver. The server predicted that all compounds do not act as carcinogens or inhibitors for hERG I/II. On the other hand, the two compounds were predicted to have a hepatotoxic effect, while the reference does not.

Conclusion

SARS-CoV-2 RdRp is a well-known viral protein target for drug designers. It has been studied for possible inhibition by previously approved drugs, such as remdesivir and sofosbuvir. In this study, we tested 14 novel modifications of sofosbuvir against the RdRp of SARS-CoV-2 *in silico*. Our docking study revealed the potential of five modifications (compounds **3**, **4**, **11**, **12**, and **14**), as they obtained binding comparable to that of the parent compound, sofosbuvir, against the RdRp active site. Additionally, the molecular dynamics simulation revealed the effectiveness of compounds **3** and **4** as the best RdRp binders. It forms stable complexes with SARS-CoV-2 RdRp mainly through V415, V477, R489, and Q493 with binding energies of -16.28 and -16.25 for compounds **3** and **4**, respectively. This study paves the way for the laboratory verification of novel SARS-CoV-2 RdRp blockers.

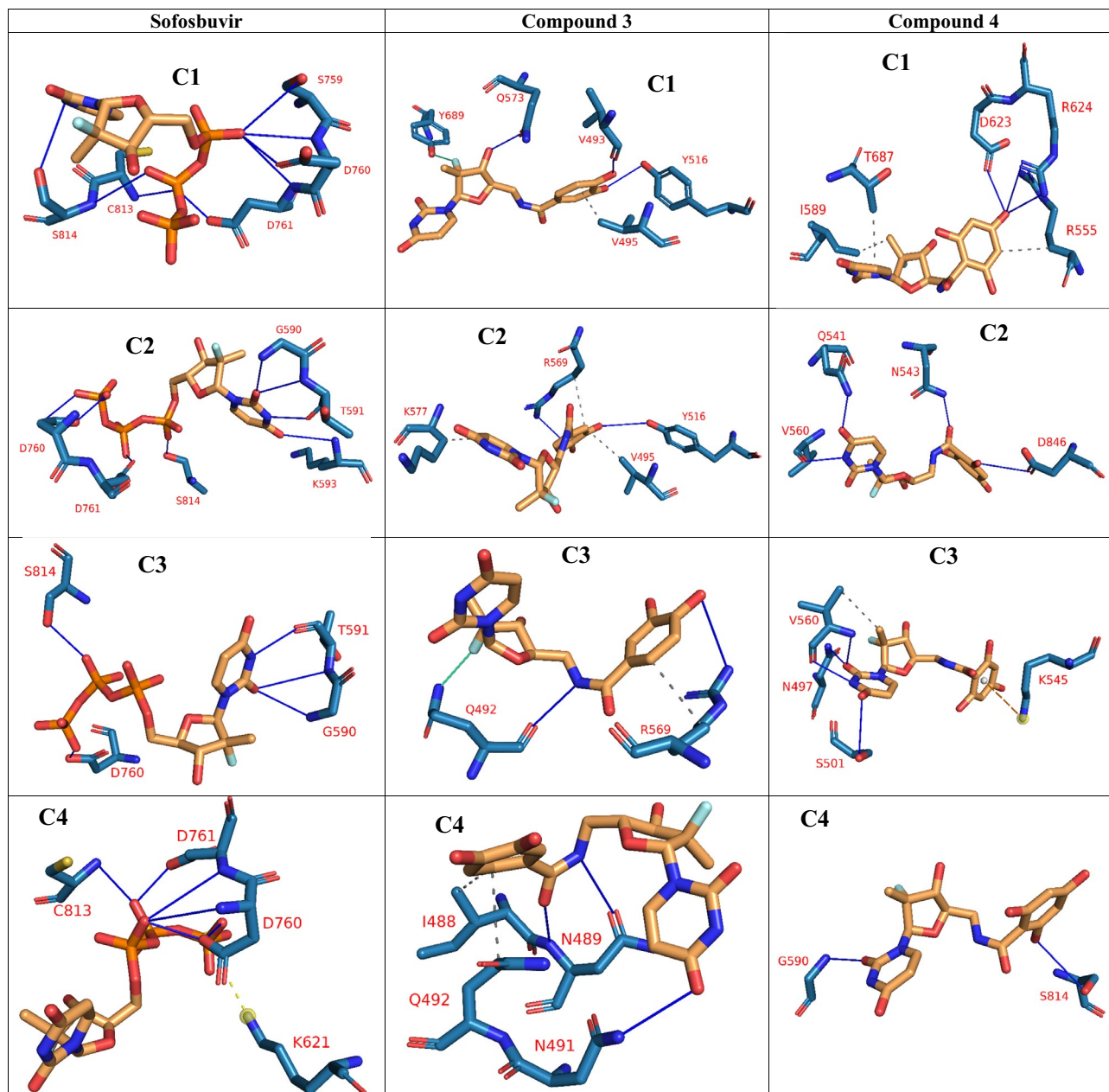


Figure 4. Shows the interactions detected using PLIP webserver (<https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index>) and then depicted using PyMOL 2.0.4 software on the representative frames obtained from the clustering of each trajectory after MDS. H-bonds are shown in blue lines, while hydrophobic contacts are shown in dashed-grey lines.

| RDRP complex | Sofosbuvir | | Compound 3 | | Compound 4 | |
|----------------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------|------------|---------------------------|
| | Residue | Binding energy (kcal/mol) | Residue | Binding energy (kcal/mol) | Residue | Binding energy (kcal/mol) |
| Residual contribution to binding | LIG | -10.10103 | LIG | -7.78 | LIG | -8.49 |
| | C733 | -2.04689 | R489 | -1.66 | V477 | -0.88 |
| | D680 | -1.21198 | Q493 | -1.1 | L678 | -0.35 |
| | L678 | -0.98482 | V415 | -0.85 | D543 | -0.34 |
| | S734 | -0.908 | K497 | -0.31 | C733 | -0.34 |
| | I509 | -0.56127 | L418 | -0.30 | K465 | -0.33 |
| | S679 | -0.51494 | T485 | -0.18 | I509 | -0.32 |
| | R475 | -0.31803 | L496 | -0.13 | R475 | -0.18 |
| | A608 | -0.26359 | Y436 | -0.08 | A478 | -0.15 |
| | F732 | -0.25081 | A500 | -0.08 | G603 | -0.15 |
| | N611 | -0.22928 | S421 | -0.05 | S679 | -0.15 |
| | Q735 | -0.11987 | N488 | -0.05 | G510 | -0.14 |
| | T607 | -0.11078 | N417 | -0.05 | N463 | -0.13 |
| | D543 | +0.08207 | I509 | -0.05 | S602 | -0.12 |
| | E731 | +0.19694 | N416 | -0.04 | A467 | -0.09 |
| | D538 | +0.21456 | R574 | +0.01 | K541 | +0.03 |
| | K465 | +0.35558 | E578 | +0.02 | D680 | +0.13 |
| D681 | +0.44948 | D419 | +0.03 | D681 | +0.14 | |
| ΔE_{VDW} (kcal/mol) | -19.84 ± 5.33 | -25.55 ± 7.08 | -26.88 ± 5.98 | | | |
| ΔE_{ELE} (kcal/mol) | -118.75 ± 26.37 | -10.18 ± 8.60 | -24.25 ± 13.21 | | | |
| ΔG_{GB} (kcal/mol) | 125.85 ± 22.72 | 23.19 ± 7.60 | 38.75 ± 12.74 | | | |
| ΔG_{SA} (kcal/mol) | -3.46 ± 0.58 | -3.74 ± 0.94 | -3.87 ± 0.88 | | | |
| ΔG_{GAS} (kcal/mol) | -138.59 ± 24.01 | -35.73 ± 9.79 | -51.13 ± 16.01 | | | |
| ΔG_{SOLV} (kcal/mol) | 122.39 ± 22.80 | 19.44 ± 7.44 | 34.88 ± 12.18 | | | |
| $-T\Delta S$ (kcal/mol) | 104.2 \pm 13.21 | 29.27 \pm 1.94 | 37.31 \pm 4.89 | | | |
| ΔG total (kcal/mol) | -16.20 ± 6.35 | -16.28 ± 5.69 | -16.25 ± 5.78 | | | |

Table 2. The MM-GBSA calculations for sofosbuvir and the best hits (compounds 3 and 4) complexes with SARS-CoV-2 RdRp after 100 ns MDS. Bold residues indicate the most contributed residues (binding energy < -0.85 kcal/mol), while red-colored residues have a negative contribution to the binding (positive binding energies). The average binding free energies and their individual terms are shown at the bottom for each complex, along with its standard deviations. Entropy ($-T\Delta S$) was calculated using the interaction entropy using the last 90 ns of each trajectory.

| | Sofosbuvir | Compound 3 | Compound 4 |
|--------------------------|------------|------------|------------|
| Absorption | | | |
| Water solubility | -2.271 | -2.524 | -2.84 |
| Caco2 permeability | 0.287 | 0.499 | 0.582 |
| Distribution | | | |
| Fraction unbound (human) | 0.451 | 0.292 | 0.291 |
| BBB permeability | -2.577 | -1.38 | -1.647 |
| CNS permeability | -3.974 | -4.083 | -4.886 |
| Metabolism | | | |
| CYP1A2 inhibitor | No | No | No |
| CYP2C19 inhibitor | No | No | No |
| CYP2C9 inhibitor | No | No | No |
| CYP2D6 inhibitor | No | No | No |
| CYP3A4 inhibitor | No | No | No |
| Excretion | | | |
| Renal OCT2 substrate | No | No | No |
| Toxicity | | | |
| AMES toxicity | No | No | No |
| hERG I inhibitor | No | No | No |
| hERG II inhibitor | No | No | No |
| Hepatotoxicity | No | Yes | Yes |

Table 3. The ADMET predictions performed by pkCSM web server.

Data availability

Data is available upon request from the corresponding author.

Received: 26 October 2023; Accepted: 11 December 2023

Published online: 27 December 2023

References

- Zhang, J. J. *et al.* Risk and protective factors for COVID-19 morbidity, severity, and mortality. *Clin. Rev. Allergy Immunol.* **64**(1), 90–107 (2023).
- Fan, Y. *et al.* Bat coronaviruses in China. *Viruses* **11**(3), 210 (2019).
- Noor, R., Shareen, S. & Billah, M. COVID-19 vaccines: Their effectiveness against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and its emerging variants. *Bull. Natl. Res. Cent.* **46**(1), 96 (2022).
- Kumar, A. *et al.* Emerging SARS-CoV-2 variants can potentially break set epidemiological barriers in COVID-19. *J. Med. Virol.* **94**(4), 1300–1314 (2022).
- Krumm, Z. A. *et al.* Precision therapeutic targets for COVID-19. *Virol. J.* **18**(1), 66 (2021).
- Nadeem, M. S. *et al.* Origin, potential therapeutic targets and treatment for coronavirus disease (COVID-19). *Pathogens* **9**(4), 307 (2020).
- Amin, F. G., Elfiky, A. A. & Nassar, A. M. In silico targeting of SARS-CoV-2 spike receptor-binding domain from different variants with chaga mushroom terpenoids. *J. Biomol. Struct. Dyn.* **12**, 1–9 (2023).
- Gomha, S. M. *et al.* Synthesis and in silico study of some new bis-[1,3,4]thiadiazolines and bis-Thiazolines as potential inhibitors for SARS-CoV-2 main protease. *Curr. Issues Mol. Biol.* **44**(10), 4540–4556 (2022).
- Almutairi, F. M. *et al.* Synthesis, molecular docking, c-met inhibitions of 2,2,2-trichloroethylidene- cyclohexane-1, 3-dione derivatives together with their application as target SARS-CoV-2 main protease (Mpro) and as potential anti-COVID-19. *Comb. Chem. High Throughput. Screen.* **26**(7), 1437–1449 (2023).
- Vicenti, I., Zazzi, M. & Saladini, F. SARS-CoV-2 RNA-dependent RNA polymerase as a therapeutic target for COVID-19. *Expert Opin. Ther. Pat.* **31**(4), 325–337 (2021).
- Roosbeh, F. *et al.* Sofosbuvir and daclatasvir for the treatment of COVID-19 outpatients: A double-blind, randomized controlled trial. *J. Antimicrob. Chemother.* **76**(3), 753–757 (2021).
- Gentile, I. *et al.* A novel promising therapeutic option against hepatitis C virus: An oral nucleotide NS5B polymerase inhibitor sofosbuvir. *Curr. Med. Chem.* **20**(30), 3733–3742 (2013).
- Ilmjarv, S. *et al.* Concurrent mutations in RNA-dependent RNA polymerase and spike protein emerged as the epidemiologically most successful SARS-CoV-2 variant. *Sci. Rep.* **11**(1), 13705 (2021).
- Goldswain, H. *et al.* The P323L substitution in the SARS-CoV-2 polymerase (NSP12) confers a selective advantage during infection. *Genome Biol.* **24**(1), 47 (2023).
- Gillis, E. P. *et al.* Applications of fluorine in medicinal chemistry. *J. Med. Chem.* **58**(21), 8315–8359 (2015).
- Palasz, A. & Ciez, D. In search of uracil derivatives as bioactive agents. Uracils and fused uracils: Synthesis, biological activity and applications. *Eur. J. Med. Chem.* **97**, 582–611 (2015).
- Basyouni, W. M. *et al.* Synthesis and antiviral screening of 2-(propylthio)-7-substituted-thiazolo[5,4-d]pyrimidines as anti-bovine viral diarrhea virus agents. *J. Heterocycl. Chem.* **58**(9), 1766–1774 (2021).
- Abduljalil, J. M. *et al.* How helpful were molecular dynamics simulations in shaping our understanding of SARS-CoV-2 spike protein dynamics?. *Int. J. Biol. Macromol.* **242**, 125153 (2023).
- Subramaniam Sivakumar, S.M., Narasimhan, R., Venkatesan, S., Maderi Velayutham D.P., *In Silico Prediction of Interactions and Molecular Dynamics Simulation Analysis of Mpro of Severe Acute Respiratory Syndrome Caused by Novel Coronavirus 2 With the FDA-Approved Nonprotein Antiviral Drugs* **12**(5), 104–119 (2022).

20. Raimundo e Silva, J. P. *et al.* Natural products as potential agents against SARS-CoV and SARSCoV-2. *Curr. Med. Chem.* **28**(27), 5498–5526 (2021).
21. Elfiky, A. A. *et al.* Molecular dynamics simulations and MM-GBSA reveal novel guanosine derivatives against SARS-CoV-2 RNA dependent RNA polymerase. *RSC Adv.* **12**(5), 2741–2750 (2022).
22. Elfiky, A. A., Azzam, E. B. & Shafaa, M. W. The anti-HCV, sofosbuvir, versus the anti-EBOV remdesivir against SARS-CoV-2 RNA dependent RNA polymerase in silico. *Mol. Divers.* **26**(1), 171–181 (2022).
23. Sussman, J. L. *et al.* Protein Data Bank (PDB): Database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr. D Biol. Crystallogr.* **54**(Pt 6 Pt 1), 1078–1084 (1998).
24. Sonousi, A. *et al.* Novel adenosine derivatives against SARS-CoV-2 RNA-dependent RNA polymerase: An in silico perspective. *Pharmacol. Rep.* **73**(6), 1754–1764 (2021).
25. Seeliger, D. & de Groot, B. L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Des.* **24**(5), 417–422 (2010).
26. Yin, W. *et al.* Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **368**(6498), 1499–1504 (2020).
27. Ezat, A. A. *et al.* Novel inhibitors against wild-type and mutated HCV NS3 serine protease: An in silico study. *Virusdisease* **30**(2), 207–213 (2019).
28. Bikadi, Z. & Hazai, E. Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. *J. Cheminform.* **1**(1), 15 (2009).
29. Stewart, J. J. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **13**(12), 1173–1213 (2007).
30. Leach, A. *Molecular Modelling: Principles and Applications* 2nd edn. (Prentice Hall, 2001).
31. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**(16), 2785–2791 (2009).
32. Lee, J. *et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**(1), 405–413 (2016).
33. Jo, S. *et al.* CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**(11), 1859–1865 (2008).
34. Huang, J. & MacKerell, A. D. Jr. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**(25), 2135–2145 (2013).
35. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**(16), 1781–1802 (2005).
36. Wang, J. *et al.* Development and testing of a general amber force field. *J. Comput. Chem.* **25**(9), 1157–1174 (2004).
37. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**(1), 338–278 (1996).
38. Adasme, M. F. *et al.* PLIP 2021: Expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Res.* **49**(W1), W530–W534 (2021).
39. Salentin, S. *et al.* PLIP: fully automated protein–ligand interaction profiler. *Nucleic acids Res.* **43**(W1), W443–W447 (2015).
40. Tubiana, T. *et al.* TTClust: A versatile molecular simulation trajectory clustering program with graphical summaries. *J. Chem. Inf. Model.* **58**(11), 2178–2182 (2018).
41. Miller, B. R. 3rd. *et al.* MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**(9), 3314–21 (2012).
42. Michaud-Agrawal, N. *et al.* MDAAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**(10), 2319–2327 (2011).
43. Xue, W. *et al.* What contributes to serotonin-norepinephrine reuptake inhibitors' dual-targeting mechanism? The key role of transmembrane domain 6 in human serotonin and norepinephrine transporters revealed by molecular dynamics simulation. *ACS Chem. Neurosci.* **9**(5), 1128–1140 (2018).
44. Tuccinardi, T. What is the current value of MM/PBSA and MM/GBSA methods in drug discovery?. *Expert. Opin. Drug Discov.* **16**(11), 1233–1237 (2021).
45. Lawitz, E. *et al.* Sofosbuvir for previously untreated chronic hepatitis C infection. *N. Engl. J. Med.* **368**(20), 1878–1887 (2013).
46. Aftab, S. O. *et al.* Analysis of SARS-CoV-2 RNA-dependent RNA polymerase as a potential therapeutic drug target using a computational approach. *J. Transl. Med.* **18**(1), 275 (2020).
47. Lindahl, E., *Molecular Dynamics Simulations*, in *Molecular Modeling of Proteins*, A. Kukol, Editor. 2015, Springer, New York, p. 3–26.
48. Elfiky, A. A. *et al.* Molecular dynamics simulations and MM-GBSA reveal novel guanosine derivatives against SARS-CoV-2 RNA dependent RNA polymerase. *RSC Adv* **12**(5), 2741–2750 (2022).
49. Pires, D. E. V., Blundell, T. L. & Ascher, D. B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **58**(9), 4066–4072 (2015).
50. Wang, J. *et al.* Development of reliable aqueous solubility models and their application in druglike analysis. *J. Chem. Inf. Model.* **47**(4), 1395–1404 (2007).

Author contributions

A.E. owns the research idea, writes the manuscript, and draws figures. I.I. performed the calculations and drafted the methods. M.I. and A.A. revised the manuscript, A.S. wrote the introduction and suggested modifications to sofosbuvir. All the authors approve the final version of the document.

Funding

Shaheen HPC in King Abdallah University for Science and Technology (KAUST), Thuwal, Saudi Arabia, is utilized to perform the MDS calculations (project k1482).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023