



OPEN Discovering novel multi-target compounds against aspartyl viral polymerases via ligand-based pharmacophore and structure-based screening

Armin Zarei¹, Mohammad Ezati^{1,2}, Marc T. Morris³, Gerhard Schenk³ & Hadi Adibi^{1,4}✉

Viral polymerases are key targets for antiviral drug development due to their crucial role in viral replication. This study seeks to discover novel inhibitors using virtual screening of a 200,583-member ZINC-based library, targeting three aspartyl-based viral polymerases: SARS-CoV-2 RdRp, HIV-1 RT, and HCV NS5B. In this line, applying RO5, drug-likeness, and PAINS filters, leading to a pharmacophore model with Ribavirin as the benchmark. Four ligands (Ligands 1–4) with the highest binding affinities were identified via molecular docking and ADME/T studies. These candidates were then subjected to 200 ns molecular dynamics (MD) simulations and binding free-energy calculations to evaluate binding stability. The computational analyses demonstrated that all four ligands stably bound the polymerase active sites and induced some conformational changes. According to MM/PBSA results, Lig-2 and Lig-1 surpass Ribavirin in SARS-CoV-2 RdRp and HIV-1 RT complexes, while Ribavirin showed the strongest binding in HCV NS5B. ADME/T studies supported the pharmacological potential of the ligands, with Lig-2 emerging as a promising lead against all targets, and Lig-1, Lig-3, and Lig-4 showing potential for multi-target antiviral development.

Keywords Viral polymerases, Pharmacophore, Virtual screening, Multi-targeted agent, MD simulations

Viral disorders pose a significant risk to public health, impacting global populations physically and mentally¹. HIV-1 (human immunodeficiency virus type 1) and HCV (hepatitis C virus) are among the most significant threats to human health. HCV, an RNA virus targeting the liver, can lead to hepatocellular carcinoma and cirrhosis², and affects ~50 million people globally, and still leads to over 240,000 deaths annually despite continued global efforts to develop treatments^{3,4}. Likewise, HIV-1 is a retrovirus that targets CD4 cells, and ultimately resulting in the development of acquired immunodeficiency syndrome (AIDS), has caused ~40 million deaths over the last four decades⁵. RNA-based viral strains, like SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), the causative agent of COVID-19, similarly pose a tremendous threat to humans, devastating lives and challenging global healthcare organizations⁶. Since November 2023, the WHO reports >771 million reported positive cases and above 6.9 million total fatalities attributed to SARS-CoV-2⁷. This viral disorder induces thrombosis, endothelial dysfunction, immune suppression, and ACE-2 pathway disruption, leading to tissue damage and extrapulmonary manifestations. Severe cases may progress to acute respiratory distress syndrome (ARDS) via inflammation, causing lung fluid accumulation and tissue damage^{6,8,9}. Fortunately, international vaccination and therapeutics targeting spike proteins or viral RNA offer promising strategies to mitigate coronavirus mortality¹⁰, despite challenges such as limited efficacy, side effects due to spike protein variations, and potential mitochondrial RNA dysfunction in humans⁷.

Viral polymerases like SARS-CoV-2 RdRp (also known as nsp12), HCV NS5B, and HIV-1 RT are crucial for viral replication, making them key targets in antiviral drug development. Targeting RdRps is particularly advantageous due to their low similarity with host cell proteins, thus minimizing the potential for drug side

¹Pharmaceutical Sciences Research Center, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran. ²Bioinformatics Laboratory, Department of Biology, School of Sciences, Razi University, Kermanshah, Iran. ³School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia. ⁴Faculty of Health and Life Sciences, INTI International University, Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia. ✉email: hadi.adibi@yahoo.com

effects¹¹. Cryogenic electron microscopy studies reveal that the SARS-CoV-2 RdRp bears a cup-shaped structure with an N-terminal extension and a NiRAN domain typical of nidoviruses linked by an interface domain (Leu251-Arg365) (Fig. 1). The interface domain facilitates the connection of N-terminal and NiRAN domains with the C-terminal domain^{12,13}. The right-hand RdRp polymerase domain is divided into three subdomains: the thumb (His816-Glu0), palm (Thr582-Pro620, Thr680-Gln81592), and finger (Leu366-Ala581, Lys621-Gly679)^{12,14}. The catalytic pocket is composed of conserved motifs (A to G) and include key aspartate residues (Asp618, Asp623, Asp760, and Asp761) which govern enzymatic function. Motif A (611–626) contributes a ligand (Asp618) that coordinate the two essential divalent metal cation [$Mg^{2+}(A)$, $Mg^{2+}(B)$], while motif C (753–767) contains the catalytic amino acids (Asp759, Asp760, and Asp761), two of which (Asp760 and Asp761) also coordinate two metallic cofactors $Mg^{2+}(A)$ and $Mg^{2+}(B)$ ¹². These catalytic amino acids are highly preserved in most viral polymerases, *e.g.*, Asp318 and Asp319 in HCV NS5B^{12,15}. The enzymatic activity of the HCV NS5B polymerase and HIV-1 RT are also governed by aspartate residues. The enzymatic activity of HCV NS5B polymerase, with a divalent cationic center in the catalytic pocket, is facilitated through highly conserved aspartate residues (Asp220, Asp318, and Asp319)¹⁶, while HIV-1 RT polymerase activity is facilitated by Asp110, Asp185, and Asp186¹⁷. The lack of effective anti-HCV and anti-HIV vaccines, and the loss of efficiency of current anti-SARS-CoV-2 vaccines due to possible mutations are serious obstacles for the treatment of these infections, the reliance of these viral strains on catalytically-similar polymerases thus offers a possible common therapeutic target. Indeed, the polymerases of the three viral strains are aspartate-based and could thus rationally be targeted by the same or

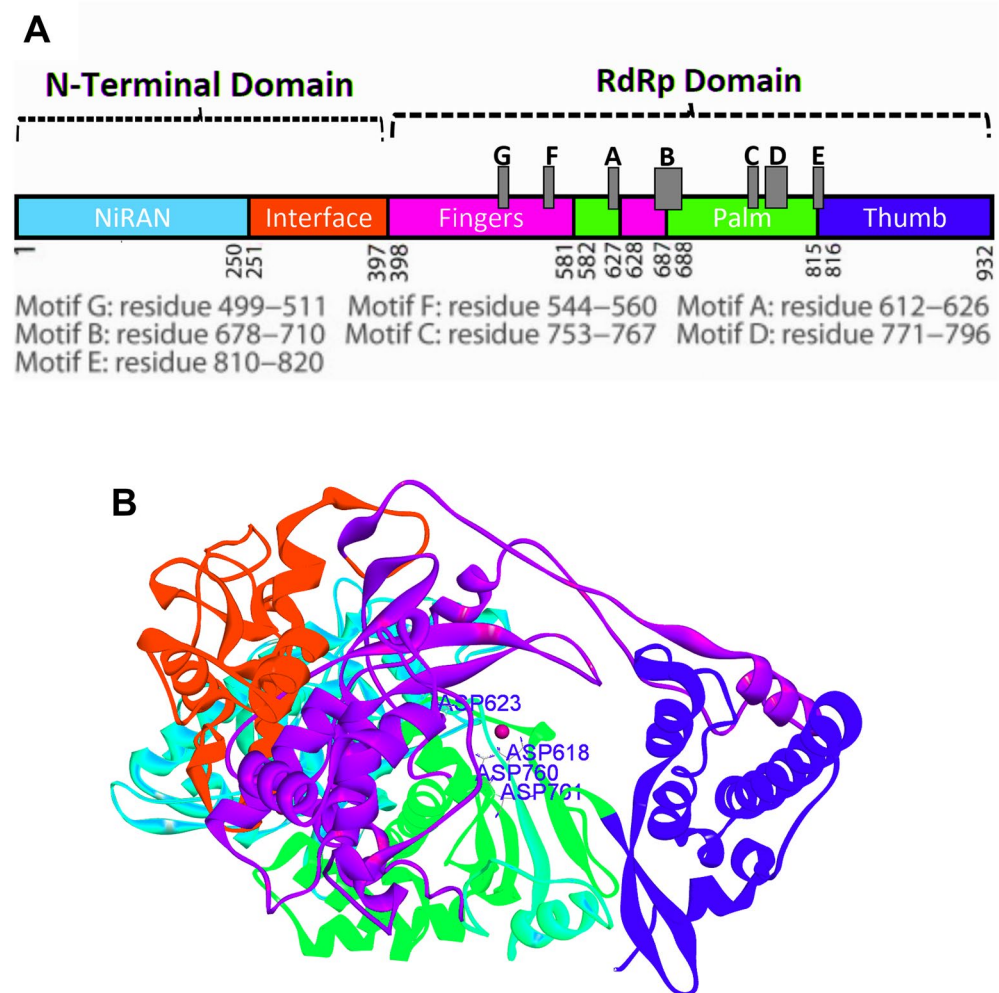


Fig. 1. The SARS-CoV-2 RNA-dependent RNA polymerase (RdRp), in complex with non-structural cofactors nsp7 and nsp8, comprises a multidomain architecture. Its N-terminal segment includes the nidovirus RdRp-associated nucleotidyltransferase (NiRAN; residues 1–250) and Interface domain (251–397), linked to the catalytic RdRp core (398–932). This core comprises three structural regions: Fingers (397–581, 627–687), Palm (582–627, 688–815), and Thumb (816–932). Within the polymerase, seven conserved sequence motifs—A (612–626), B (678–710), C (753–767), D (771–796), E (810–820), F (544–560), and G (499–511)—support template recognition and catalysis. Structural representation (ribbon model) highlights NiRAN (light blue), Interface (red), Fingers (purple), Palm (green), and Thumb (blue). The catalytic core harbors four essential aspartate residues (Asp618, Asp623, Asp760, Asp761) coordinating two Mg^{2+} ions (A and B), critical for polymerase function.

similar compounds, opening up the exploration of a novel, multi-purpose antiviral therapeutic. However, since the process of discovering new drugs involves complex procedures¹⁴, theoretical approaches have increasingly been used as the first and rapid step towards identifying promising leads for subsequent testing and refinement¹⁸. Recently, a series of concerted efforts shows that target the replication and transcription of SARS-CoV-2 by using the drugs Molnupiravir and Remdesivir^{19,20}. These studies demonstrate that nucleotide-based inhibitors have potential as antivirals against future coronaviruses¹⁰. Substrate binding in the SARS-CoV-2 RdRp may occur if motif A stabilizes substrate binding, thus facilitating the coordination of Asp618 to both metallic cofactors, leading to substrate recognition through H-bond formation between catalytic Asp623 and the 2'-OH of the ribose ring of the substrate¹⁰. This hydroxyl moiety at the 2' position of the substrate ribose is a pharmacophore of high importance in antiviral drug development¹⁰. The structures containing different domains and motifs of HIV-1 RT and HCV NS5B are also illustrated in Fig. S1 (Supporting information).

To highlight the research gap, it should be noted that polymerase inhibitors like sofosbuvir and remdesivir (targeting SARS-CoV-2 and HCV) are efficient but virus-specific, showing limited cross-viral activity and vulnerability to resistance²¹. Besides, no drug is approved to simultaneously inhibit the aspartate-based polymerases of HIV-1, HCV, and SARS-CoV-2²². Thus, multi-target in silico strategy seeks compounds that bind the conserved polymerase active sites across these viruses, addressing a clear gap and offering a novel paradigm for broad-spectrum antiviral development. To rationalize the importance of this study, given the importance of the ribose moiety and the common reliance on aspartate residues in all three target viral polymerases (HIV-1 RT, HCV NS5B, and SARS-CoV-2 RdRp), we employed a rational pharmacophore-based process, using the ribose-containing Ribavirin as a control to explore naturally-occurring ligands as potential inhibitors for SARS-CoV-2 RdRp, HCV NS5B, and HIV-1 RT. Firstly, a workflow was designed to produce leads by applying a number of filters (drug-likeness, PAINS, Lipinski's rule of five, toxicity, and ADME properties) on potential hits. Secondly, those natural compounds meeting the requirements of all the filters were introduced to a molecular docking pipeline. Finally, to shed light on the dynamic nature of these ligands, those with the strongest binding energies were introduced to MD simulations.

Methods

Formation of the library of natural ligands

A collection of 200,583 commercially-available natural compounds was obtained from the ZINC webserver²³. To ensure dataset quality, a rigorous cleanup process was applied, which included; removal of duplicate entries, incomplete structures, and isotopic species and filtering based on physicochemical properties using the FAF-Drugs4 web server²⁴. The following RO5 parameters were applied to ensure drug-likeness: Molecular Weight ≤ 500 , Hydrogen Bond Acceptors ≤ 10 , Hydrogen Bond Donors ≤ 5 , LogP ≤ 5 , Polar Surface Area ≤ 140 , and Rotatable Bonds ≤ 10 . Then, pan-assay interference (PAINS) filtration was employed to remove compounds susceptible to such effects²⁴. Subsequently, the structures underwent energy minimization using the MMFF94 force field through Open Babel 3.0.1²³.

Ligand-based pharmacophore modelling

Ligand-based pharmacophore modeling and database screening were conducted using the Pharmit online tool (<http://pharmit.csb.pitt.edu>)²⁴. This online platform facilitates virtual screening of both pre-existing and user-generated databases by employing molecular configuration, pharmacophore models, and energy minimization techniques. In the present study, the main pharmacophore was set on Ribavirin since this ligand has defined functional groups that interact effectively with the active sites of polymerases, e.g., a nucleobase-like moiety, which allows this ligand to interact with nucleotide-binding sites^{25,26}, a ribose ring that is a vital moiety in interactions with the catalytic residue Asp623¹⁰, and two hydroxyl groups at the 2' and 3' positions of the ribose ring, which are critical in the substrate recognition within the active site of SARS-CoV-2 RdRp¹⁰. Last but not least, hydrogen bond-forming potential to anchor the molecule in the active site. Importantly by far, Ribavirin has demonstrated strong antiviral potential against different viral strains^{27–29}. Thus, these excellent pharmacophoric features of Ribavirin have prompted us to generate our pharmacophoric model based on this ligand. In this line, a ligand-based pharmacophore model was generated using the Pharmit online webserver. To do this, only the 3D structure of Ribavirin (not in complex with the protein) was uploaded in the webserver (Fig. 2), and specific pharmacophores were selected to screen the generated library in order to find natural ligands with the potential of targeting multiple viral enzymes. Ligand-based pharmacophore screening was subsequently carried out against the compounds identified in this screening step.

Rationale for Ribavirin as the Template

Ribavirin was selected as the foundational pharmacophore for this multi-target drug discovery campaign based on a dual rationale: (i) Structural-Mechanistic: its 2'-hydroxyl group is critically involved in substrate recognition by forming a hydrogen bond with the catalytic Asp623 in SARS-CoV-2 RdRp¹⁰, a key interaction in the aspartate-dependent polymerase mechanism common to all three targets. (ii) broad-spectrum validation: it is a rare example of a clinically relevant agent with documented in vitro inhibitory activity, e.g. Ribavirin inhibits the replication of HCV (IC₅₀ of 11.5 μ M)²⁷, Influenza A (IC₅₀ of 30 μ M)²⁸, and HIV-1 RT (IC₅₀ values of 81 and 112 μ M for its di- and triphosphate forms)²⁹. Besides, its key pharmacophoric feature (2'-OH) is essential for interaction with SARS-CoV-2 RdRp¹⁰, making it an ideal prototype for discovering novel broad-spectrum inhibitors.

Rationale for pharmacophore model design and feature selection

The pharmacophore model was explicitly derived from Ribavirin for a rational, structure-based strategy. First, the 2'-hydroxyl group of Ribavirin's ribose moiety is mechanistically critical, as it forms a key hydrogen bond

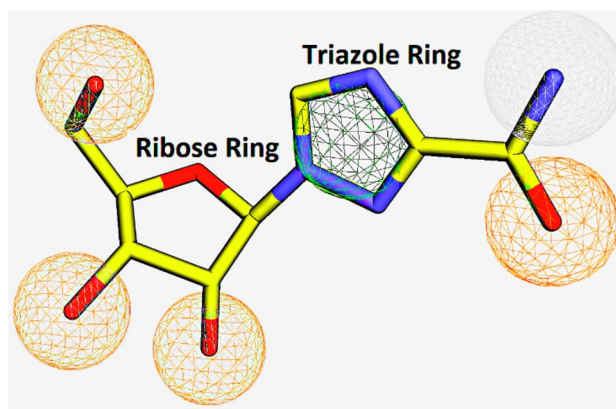


Fig. 2. Crucial pharmacophoric characteristics of Ribavirin obtained evaluated by using the Pharmit server, all within a radius of 1 Å. Four hydrogen bond acceptors in orange color (the hydroxyl groups of the ribose ring and the carbonyl moiety), one hydrogen bond donor in white color (amine group), and a hydrophobic moiety colored in green (the 1,2,4-triazole ring).

with the catalytic Asp623 residue during substrate recognition in SARS-CoV-2 RdRp¹⁰. Second, Ribavirin exhibits documented in vitro inhibitory activity against multiple viral polymerases including HCV NS5B ($IC_{50} = 11.5 \mu M$)²⁷, Influenza A²⁸, and HIV-1 RT²⁹, making it a proven multi-target template. The specific pharmacophoric features extracted were: four hydrogen bond acceptors (corresponding to the ribose hydroxyls and the carbonyl group), one hydrogen bond donor (the amine group), and one hydrophobic region (the 1,2,4-triazole ring), which directly map to these essential functional groups. This feature set was designed to mimic the core interaction pattern of a nucleoside analog within the aspartate-rich catalytic pockets of the target polymerases, thereby enriching the library for compounds with inherent broad-spectrum potential.

Structure-based virtual screening

Preparation of the viral receptors

The 3D building blocks of the viral polymerases SARS-CoV-2 nsp12 (PDB ID: 7ED5), HCV NS5B polymerase (PDB ID: 2XI3), and HIV-1 reverse transcriptase (1S9G) with resolutions of 2.5 Å, 1.7 Å, and 2.8 Å, respectively, were obtained from PDB RCSB webserver (<https://www.rcsb.org>). The 3D structures were examined to ensure no atoms are missed, following which water and other nonessential molecules were then discarded. Comparative analyses were then carried out to explore the composition, quantity, and amino acid sequences, as well as to assess the structural resemblance amongst these enzymes. The initial 3D structures were analyzed through the ESPript 3.0 online tool³⁰, whereas their spatial resemblance was assessed via the Chimera 1.13.1 software³¹.

Docking studies

Prior to commencing the docking studies, the coordinates of the viral proteins underwent energy minimization using the GROMACS 2022 package³² to guarantee their thermodynamic structural stability. The minimization involved employing the steepest descent approach initially, followed by conjugate gradient energy minimization, aiming to reduce the maximum force to $< 1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ and thus remove any disruptive forces. The docking procedures were performed through Autodock Vina 2.1.1 to estimate the binding affinities of the ligands³³. The library of the ligands is in pdbqt format and was produced through Open Babel 3.0.1. Initially, hydrogen atoms were added to the ligands, and the structures were subsequently energy-minimized using the MMFF94 force field implemented in the Open Babel toolbox. The compounds were then converted into the required file formats sequentially using Open Babel 3.0.1 and AutoDock Tools (ADT). The protonation states of the targets have been checked out by Chimera 1.8.1. The coordinates of the viral enzymatic receptors were generated using ADT 1.5.6. (Autodock Tools) through the incorporation of Kollman charges and polar hydrogen atoms, the assignment of suitable atom types, and the merging of nonpolar hydrogens. The grid box was then set to cover all of the significant amino acids involved in the catalytic pocket map. For the molecular docking protocol, target viral proteins were treated as rigid structures, while ligand conformational flexibility was permitted. Grid box centers were assigned as follows: SARS-CoV-2 RdRp ($x = 128, y = 134, z = 127$), HIV-1 reverse transcriptase ($x = 154, y = -17, z = 80$), and HCV NS5B polymerase ($x = 8, y = 6.567, z = 13$). A uniform grid box size of $40 \times 40 \times 40 \text{ Å}$ was employed across all systems, covering the active site of the targets. The exhaustiveness parameter was set to 24, with all remaining configuration settings retained at default values. 20 binding poses were generated for each ligand. AutoDock Vina's default empirical scoring function was used to estimate binding affinities, and random seeds were automatically assigned by the program. Also, for the validation of molecular docking procedure, the crystallographic inhibitor of each receptor (for SARS CoV-2; AT9 in the PDB ID: 7ED5, for HCV NS5B; GTP in the PDB ID: 2XI3, and for HIV-1 RT; ABZ in the PDB ID: 1S9G) was extracted and redocked into its active site. Then, the RMSD between the crystallographic ligand and redocked ones was calculated by Chimera 1.8.1 (see Table S1, Supporting Information, and results in section).

Library screening by vina scoring and pose

Binding affinity scores generated by AutoDock Vina were ranked based on their predicted interaction energies. In the presented binding score plots, the top 25% of the score range was used as the selection criterion for hits. Ligands exhibiting pose redundancy upon structural inspection were shortlisted for further analysis. Molecular interaction profiling was performed using LigPlot⁺ for visualization.³⁴

Docking and virtual screening validation and enrichment studies

Docking validation was conducted using a robust computational framework in Python, employing a tenfold cross-validation strategy and enrichment factor (EF) calculations. This process applied the scientific ecosystem of Python libraries, including scikit-learn (version 1.2.2) for data partitioning and RDKit (version 2022.09.5) for molecular descriptor calculations and similarity filtering. Enrichment analysis was performed by curating high-quality, enzyme-specific benchmarking sets. To ensure biological relevance, a total of 8 active compounds with high-confidence experimental data (IC_{50} or K_i values) were chosen from peer-reviewed literature and reputable databases, including BindingDB, ChEMBL, and PubChem. These actives were categorized based on their target specificity:

1. SARS-CoV-2 RdRp inhibitors: Remdesivir (CID: 121,304,016), Molnupiravir (CID: 145,996,610), and Ribavirin (CID: 37,542).
2. HCV NS5B polymerase inhibitors: Sofosbuvir (CID: 45,375,808), Dasabuvir (CID: 56,640,146), and Ribavirin (CID: 37,542).
3. HIV-1 RT inhibitors: Tenofovir (CID: 464,205), Zidovudine (CID: 35,370), Nevirapine (CID: 4463), and Ribavirin (CID: 37,542).

Ribavirin can act as the principal broad-spectrum positive control across all sets due to its documented inhibitory activity against all three target viral polymerases^{10,27,29}. The decoy sets were generated using a property-matched approach to maintain comparable physicochemical features (MW, logP, TPSA, and rotatable bonds) relative to the actives. A total of 1085, 1022, and 1235 decoy molecules were curated for SARS-CoV-2 RdRp, HCV NS5B, and HIV-1 RT, respectively, from the ZINC15 database (accessed October 2023). To minimize analog bias, any decoy with a Tanimoto coefficient > 0.35 relative to any active was excluded using Morgan fingerprints (radius 2).

Validation metrics and procedure

The tenfold cross-validation was implemented using the KFold class from scikit-learn with a fixed seed (random_state = 42). To comprehensively evaluate the screening performance and the early recognition problem, the following metrics were calculated:

- AUC-ROC and Log-AUC: To assess overall discriminative power and early enrichment on a logarithmic scale.
- Enrichment Factor ($EF_1\%$ and $EF_{10\%}$): To quantify the concentration of actives within the top-ranked results.
- BEDROC Score ($\alpha = 20$): To evaluate the protocol's ability to prioritize actives at the very top of the list.
- Confusion Matrix: To determine Precision, Recall, and F1-score at optimized docking score cut-offs.

To account for statistical variance and ensure the robustness of the metrics, 95% confidence intervals (95% CI) were calculated for AUC-ROC using the DeLong method. For EF and BEDROC, a bootstrapping procedure with 1000 replicates was employed to estimate the standard error and confidence limits. The Enrichment Factor (EF) was calculated as:

$EF\ X\% = (\text{Number of actives found in top } X\% / \text{Total number of compounds in top } X\%) / (\text{Total number of actives} / \text{Total number of compounds})$. This comprehensive and enzyme-specific validation approach ensures the robustness of the virtual screening protocol and its ability to distinguish true polymerase inhibitors from inactive decoys.

Prediction of ADMET factors and physicochemical properties

The pkCSM³⁵, SwissADME³⁶, and ProTox-II web³⁷ servers were utilized to analyze various factors for the selected hits and Ribavirin (control drug). These factors included Lipinski's rule of five and physicochemical peculiarities including water solubility, topological polar surface area, and rotatable bonds count. Pharmacokinetic features such as human intestinal absorption and P-glycoprotein inhibition were predicted, along with distribution parameters including central nervous system permeability, blood-brain barrier permeability, and volume of distribution. Additionally, metabolism-related factors, e.g. inhibition potential against different cytochrome P450 subtypes, were calculated. Furthermore, excretion factors such as the inhibition of human ether-related gene channels, total clearance, renal OCT2 substrate potential, and toxicity parameters like AMES toxicity, acute and chronic oral toxicity in rats, LD50, and skin sensitization were examined.

MD (molecular dynamics) simulations

MD simulations were applied via the GROMACS 2022 software³² to shed more light on the dynamic behavior of the viral receptors in free states and in complexes with the ligands of the highest binding energies. These computational modelling studies can provide valuable information on the paradigm, the interactions stability, and possible dynamic structural alterations. The topologies of the viral proteins were generated using the OPLS-AA force field, while the LigParGen was utilized for the topologies of the ligands^{38,39}. Simulation boxes were filled with water (SPC/E model) before the addition of adequate Na^+ and Cl^- ions for neutralization of the simulation systems. Energy minimization was supplemented using the steepest descent algorithm until energy gains below

10 kJ/mol⁴⁰. Periodic boundaries were applied in all directions (X, Y, Z). Following this, the pressure was set at 1.0 bar, while the temperature was adjusted to 310 K in NPT and NVT ensembles. In the equilibration stage, the temperature and pressure were controlled using the V-rescale thermostat and Parinello-Rahman barostat. Van der Waals and electrostatic contacts were also computed within a 1 nm cutoff. The simulation also utilized the PME (Particle Mesh Ewald) algorithm to calculate long-range electrostatic contacts involving Coulomb potential energies. Bond constraints for heavy atoms were set using the LINCS algorithm⁴¹, with 2 fs time step. Ultimately, 200 ns MD simulations were performed through the leap frog algorithm⁴², with frames saved every 10 ps for analysis. 2D and 3D molecular interactions were visualized using LigPlot⁺ (V 2.2.9) and Discovery Studio (V 4.5). The MD indicator analyses were conducted after the system RMSD (root mean square deviation) computed with the GROMACS gmx_rms module reached a point where the system was in a thermodynamically stable state. The MD indicators were used to compute the amino acid fluctuations and changes in proteins compression by measuring the radius of gyration throughout the equilibrated time of the trajectory. A key use of MD simulations is to estimate the binding free energy of receptor-ligand complexes⁴³. Here, the binding free energy of receptor-ligand complexes are calculated through MM/PBSA (Molecular Mechanics/Poisson-Boltzmann Surface Area) approach.

Results & discussion

Library mining and pharmacophore-based virtual screening

A starting set of 200,583 natural compounds was sourced from the ZINC webserver. This library underwent filtration via the FAF-Drugs4 to remove duplicates and evaluate PAINS, drug-likeness, and RO5 properties, thereby excluding false-positive and interfering molecules. PAINS-filtered molecules often show activity from reactivity rather than specific binding, leading to high reactivity with proteins in many bioassays. This filtration process reduced the library to 104,047 compounds. Subsequently, a structure-based pharmacophore model was developed using the Pharmit server.

Moreover, the rationale for our feature selection was focused on the critical molecular recognition patterns required for polymerase inhibition¹⁰. As established in the literature, the hydroxyl (OH) moiety at the 2' position of the substrate ribose is indispensable for substrate recognition and catalytic incorporation in viral polymerases, including SARS-CoV-2 RdRp¹⁰, HCV NS5B²⁵, and HIV-1 RT²⁹. Consequently, Ribavirin, a potent broad-spectrum antiviral that mimics this essential ribose feature was utilized as the template for pharmacophore construction. The pharmacophore model was defined with a tolerance radius of 1 Å, capturing four hydrogen bond acceptors, one hydrogen bond donor, and a hydrophobic moiety (Fig. 2). These features were specifically selected to map the spatial orientation of residues involved in the catalytic mechanism, such as the coordination with divalent magnesium ions and conserved aspartate residues within the active site. By applying this pharmacophore filter to the 104,047 drug-like compounds, the library was narrowed down to 1,801 high-priority hits. To validate the reliability of this pharmacophore model and the subsequent docking protocol, an enrichment analysis was performed using validated active inhibitors and property-matched decoys.

Binding affinity assessment via docking

In the present research, AutoDock Vina experiments were conducted to provide binding affinity insights useful for prioritizing compounds based on their binding strengths. The docking results offer insights that can guide rational lead generation and reduce the number of library members when applying appropriate criteria. Docking examinations of all 1801 compounds were conducted with the three viral polymerases in AutoDock Vina to determine relative binding affinities, which can be used to further prioritize and minimize the number of lead compounds. After the docking process, the obtained binding affinities ranged from −4.6 to −9.6 kcal/mol for SARS-CoV-2 RdRp, −4.8 to −12.2 kcal/mol for HIV RT, and −4.6 to −10.1 kcal/mol for HCV NS5B. The top ~25% ligands based on these affinities were then selected for further analysis. In total, 117 compounds were identified for SARS-CoV-2 RdRp with binding energies ≥ −8.3 kcal/mol, 50 for HIV RT with binding energies ≥ −10.4 kcal/mol, and 104 for HCV NS5B with binding energies ≥ −8.7 kcal/mol were pre-selected (Fig. 3). Subsequently, a consensus of shared hits across the viral receptors was identified, and 21 molecules were chosen for subsequent toxicity and ADME evaluation.

Validation of the virtual screening protocol

To ensure the reliability and discriminative power of the docking protocol, a comprehensive enrichment analysis was performed across three viral targets: SARS-CoV-2 RdRp, HCV NS5B, and HIV-1 RT. The benchmarking sets were designed with a high decoy-to-active ratio to mimic the "needle-in-a-haystack" scenario inherent in real-world virtual screening.

Dataset composition and benchmarking quality

The robustness of any virtual screening validation depends heavily on the diversity and size of the decoy set. In this study, we utilized; for SARS-CoV-2 RdRp: 3 actives vs. 1085 decoys (Ratio ~1:360), for HCV NS5B: 3 actives vs. 1022 decoys (Ratio ~1:340), and HIV-1 RT: 4 actives vs. 1235 decoys (Ratio ~1:300). According to the standards established in the literature, a decoy-to-active ratio exceeding 1:50 is considered robust for calculating Enrichment Factors (EF). Our average ratio of ~1:330 significantly exceeds this threshold, providing a more rigorous test for the scoring function's ability to minimize false positives. Furthermore, the property-matched nature of the ZINC15 decoys ensures that the results are not biased by simple physicochemical differences, but rather reflect true molecular recognition.

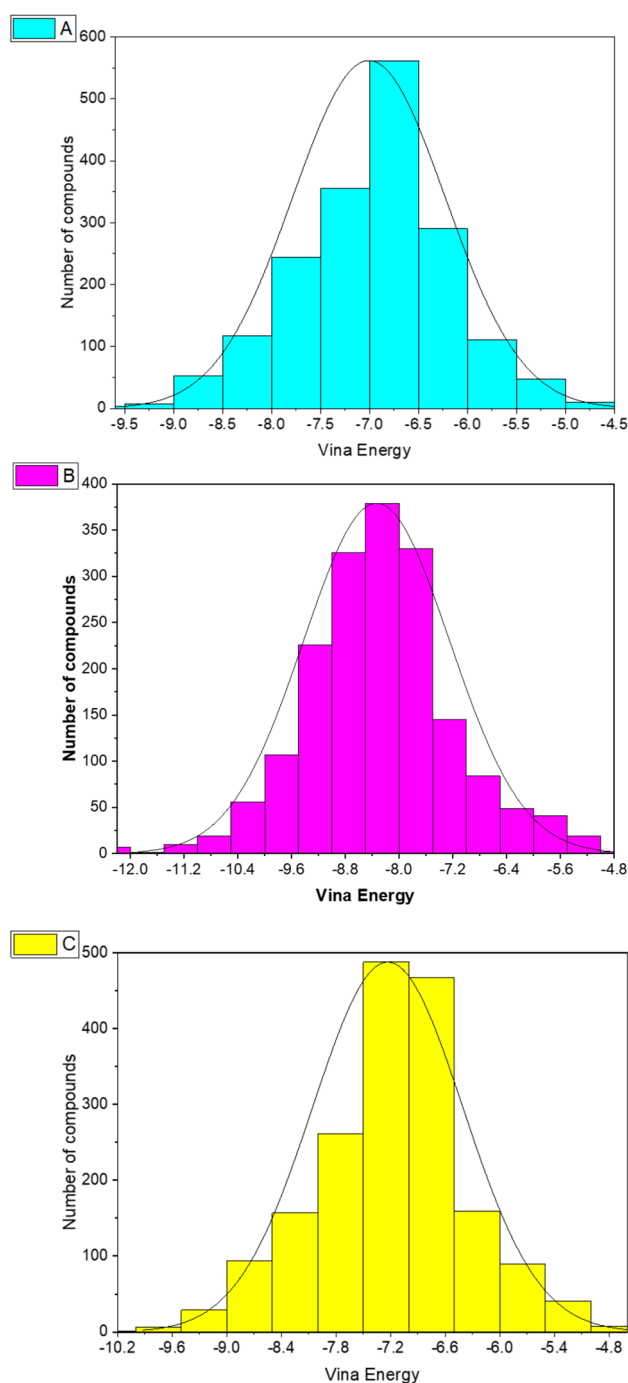


Fig. 3. Frequency distributions and desirability function histograms of docking affinities from AutoDock Vina. (A) SARS-CoV-2 RdRp, (B) HIV-1 RT, and (C) HCV NS5B.

Evaluation of discriminative power (AUC-ROC and Log-AUC)

The Area Under the Receiver Operating Characteristic curve (AUC-ROC) was used to assess the overall performance of the docking protocol. As summarized in Table 1 and Fig. 4, the protocol demonstrated exceptional performance across all targets; SARS-CoV-2 RdRp achieved the highest AUC-ROC of 0.9745, indicating a 97.4% probability that a randomly chosen active will rank higher than a randomly chosen decoy. HCV NS5B and HIV-1 RT also showed significant discriminative power with AUC-ROC magnitudes of 0.9669 and 0.9668, respectively. The high AUC values (all > 0.95) suggest that the AutoDock Vina scoring function, combined with our specific grid box parameters and preprocessing steps, is highly effective at identifying viral polymerase inhibitors. The consistency of these results across three different viral families (Coronaviridae, Flaviviridae, and Retroviridae) highlights the universal applicability of the established protocol.

Target	Decoys	Actives	AUC-ROC (95% CI)	BEDROC ($\alpha=20$)	EF 1%	EF 5%
SARS-CoV-2 RdRp	1088	3	0.9745 (0.961–0.988)	4.6276	72.5333	13.4321
HCV NS5B	1025	3	0.9669 (0.948–0.985)	3.8308	34.1667	13.3987
HIV-1 RT	1239	4	0.9668 (0.952–0.981)	2.4980	0.0	20.3115

Table 1. Detailed validation metrics for the three targeted viral polymerases.

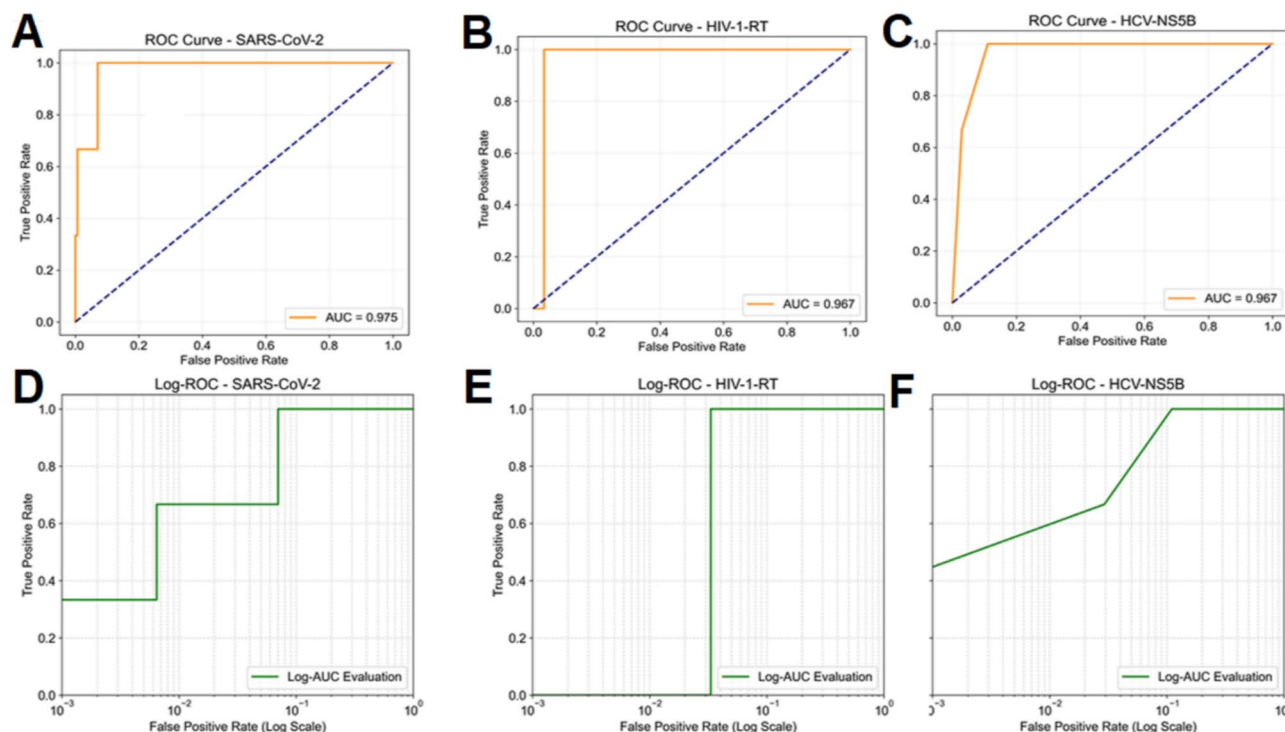


Fig. 4. Validation of the virtual screening protocol across three viral polymerase targets. (A–C) Standard ROC curves and (D–F) Logarithmic ROC curves for SARS-CoV-2 RdRp, HIV-1 RT, and HCV NS5B, respectively. High AUC-ROC values (>0.95) across all targets confirm the robustness and consistency of the docking parameters. The steep initial slope in the Log-ROC plots illustrates superior early recognition to ensure known inhibitors are prioritized at very low False Positive.

Early enrichment and recognition performance While AUC-ROC measures global performance, the early recognition problem is more critical for virtual screening where only the top 1% or 5% of a library can typically be transitioned to in vitro assays.

Enrichment factors (EF) The Enrichment Factor at 1% and 5% (EF_{1%} and EF_{5%}) provides a quantitative measure of how many more actives are found in the top-ranked results compared to a random selection. For SARS-CoV-2 RdRp, a staggering EF_{1%} of 72.53 was observed, meaning that the top 1% of the ranked list is more than 70 times richer in active compounds than the overall dataset. For HIV-1 RT, while the EF_{1%} was 0.00 (indicating that actives were ranked slightly below the first 12 compounds), the EF_{5%} of 20.31 confirms that all actives were successfully prioritized within the top 5% of the library. This shift is a known phenomenon when dealing with specific nucleoside analogs that may require more flexible docking considerations but are still effectively captured early in the screening process.

BEDROC and Log-ROC analysis

The Boltzmann-Enhanced Discrimination of Receiver Operating Characteristic (BEDROC) with $\alpha=20$ was calculated to further penalize late-ranking actives. The BEDROC scores (ranging from 1.96 to 4.62) confirmed that the protocol excels at early enrichment. Log-ROC analysis (visualized in Fig. 4) further illustrates this by showing a steep vertical increase at the beginning of the logarithmic scale (FPR $<10^{-2}$). This steepness confirms that the protocol is highly efficient at prioritizing true inhibitors even when they are outnumbered by thousands of decoys. The integration of high AUC-ROC values, significant EF_{5%} scores, and positive BEDROC metrics provides a solid foundation for the subsequent virtual screening of the targeted library. The results demonstrate that the docking parameters utilized herein are not only statistically significant, but also practically viable for identifying novel polymerase inhibitors with high precision.

ADME & toxicity investigation

The ADME/T peculiarities of the 21 natural compounds identified above were introduced to the SwissADME, pkCSM, and ProTox-II online web servers, and compared against standard parameters. Relevant criteria for the four most promising candidates are listed in Table 2. A crucial criterion for effective drugs is their bioavailability. The four selected natural compounds (*i.e.*, ZINC85875904 (Lig-1), ZINC4081776 (Lig-2), ZINC8789971 (Lig-3), and ZINC4096280 (Lig-4)) have a bioavailability score of 0.55, which satisfies the Lipinski rule of five. The four ligands are predicted to have high gastrointestinal absorption, at least 30% above the intestinal absorption threshold. Typically, low renal clearance correlates with a long half-life in the body⁴⁴, and our calculations indicate that the four ligands possess renal clearance magnitudes of between 0.1 to 1.0, suggesting medium half-lives. They also show an acceptable volume of distribution, but their blood brain barrier ($-1.199 < \log BB < -0.784$) and central nervous system ($-3.834 < \log PS < -3.481$) permeabilities are poor. However, since the viral targets of these compounds are not present in the nervous system, crossing the blood–brain barrier may not be essential for these inhibitors. Significantly, none of the candidates are predicted to interact with other pharmaceuticals, and are also unlikely to act as substrates or inhibitors for Cytochrome P450s.

Compounds Name	Compounds				
	Lig-1 (ZINC85875904)	Lig-2 (ZINC4081776)	Lig-3 (ZINC8789971)	Lig-4 (ZINC4096280)	Ribavirin (Control Drug)
Physicochemical Properties					
MW ^a	456.49	400.38	473.56	402.39	244.20
HBA ^b	8	8	7	8	7
HBD ^c	4	4	4	4	4
RB ^d	5	4	5	4	3
tPSA (Å ²) ^e	125.68	129.59	125.32	125.68	143.72
Bioavailability Score	0.55	0.55	0.55	0.55	0.55
Log P ^f	2.44	1.48	1.32	1.75	-1.85
Log S ^g	-4.72	-2.77	-3.55	-2.97	-0.65
GI Absorption	High	High	High	High	Low
Human intestinal absorption (HIA%) ^h	57.935	50.785	58.11	53.985	43.158
P-glycoprotein inhibition	Non	Non	Non	Non	Non
Distribution					
VD _{ss} (Human) ⁱ	-0.609	-0.743	-0.364	-0.828	-0.16
barrier Blood–brain permeability ^j	-1.143	-1.197	-0.784	-1.199	-1.159
Central nervous system permeability ^k	-3.596	-3.813	-3.481	-3.834	-4.134
Metabolism					
Cytochrome P450 1A2 inhibition	Non	Non	Non	Non	Non
Cytochrome P450 2C19 inhibition	Non	Non	Non	Non	Non
Cytochrome P450 2C9 inhibition	Non	Non	Non	Non	Non
Cytochrome P450 2D6 inhibition	Non	Non	Non	Non	Non
Cytochrome P450 2D6 substrate	Non	Non	Non	Non	Non
Cytochrome P450 3A4 inhibition	Non	Non	Non	Non	Non
Cytochrome P450 3A4 substrate	Non	Non	Non	Non	Non
Excretion					
CL _T ^l	0.568	0.499	0.607	0.729	0.692
Renal OCT2 substrate	Non	Non	Non	Non	Non
Toxicity					
AMES toxicity	Non	Non	Non	Non	Non
HERG_inhibition ^m	Non	Non	Non	Non	Non
LD ₅₀ ⁿ	2190	5000	3000	2300	2700
Hepatotoxicity	Non	Non	Non	Non	Non
Skin sensitization	Non	Non	Non	Non	Non

Table 2. Predicted pharmacokinetic and toxicity characteristics of the four selected compounds (prior to MD simulation). ^a Molecular weight (≤ 500 Da), ^b Number of hydrogen bond acceptors (≤ 10), ^c Number of hydrogen bond donors (≤ 5), ^d Number of rotational bonds (≤ 10), ^e topological polar surface area (≤ 150 Å²), ^f Partition coefficient in oil to water (≤ 5), ^g Aqueous solubility descriptor (≤ 0), ^h absorption (human, % > 30), ⁱ volume of distribution (human, log L/kg) (low if < -0.15 and high if > 0.45), ^j permeability ($\log BB > 0.3$), poorly ($\log BB < -1$), ^k permeability ($\log PS > -2$), non-penetrability ($\log PS < -3$), ^l total renal clearance; high (> 1 mL/min/kg), medium (> 0.1 to < 1 mL/min/kg) or low (≤ 0.1 mL/min/kg), ^m HERG: Human ether related gene. ⁿ Lethal dose 50 (mg/kg).

No mutagenicity was monitored for these natural ligands, and their toxicities, predicted using the ProTox-II online webserver, are low. Each compound has acceptable lethal dose levels, with Lig-2 having the highest LD₅₀ (5000 mg/kg). In summary, these natural ligands are predicted to display promising pharmacological characteristics, and possess effective absorption, distribution, and diffusion through the human body. Subsequently, they were introduced to 200 ns MD simulations to explore their impacts on the viral target polymerases (see below).

Off-target and toxicity investigations

It has been previously reported that nucleoside analogs might often hit human polymerases. Especially, inhibition of DNA polymerase γ or human mitochondrial RNA polymerase (POLRMT) is a well-documented toxicity mechanism for antiviral ribonucleotides. Recent studies have also demonstrated that off-target toxicity caused by inhibition of POLRMT is a major liability for RdRp inhibitors^{45,46}. We have explicitly acknowledged this, and use ligand-based target prediction tools to flag potential human targets. To do this, SwissTargetPrediction webserver (<https://www.swisstargetprediction.ch>) has been utilized to identify any likely human polymerase targets. According to our results, none of the four selected ligands have shown any off-target toxicity against DNA polymerase γ or human mitochondrial RNA polymerase (POLRMT).

Docking studies

Molecular docking is a powerful technique that offers crucial insights into how small molecules bind to macromolecular targets, including proteins^{14,47,48}. The results of the molecular docking investigation with Ligands 1–4 are summarized in Table 3, and their 3D and 2D interactions are shown in Figs. 5 and 6, respectively.

Docking studies with SARS-CoV-2 RdRp

Initially, to validate the molecular docking, the crystallographic inhibitor of each receptor was extracted and redocked into its active site. Then, the RMSD between the crystallographic ligand and redocked ones was calculated. Accordingly, the results showed that the redocked conformers in the structures of SARS CoV-2 RdRp, HCV NS5B, and HIV-1 RT have shown low RMSDs of 0.745 Å, 1.123 Å, and 1.153 Å, with their crystallographic counterparts, which can be suitable since the RMSD below 2.0 Å can propose the suitability of the docking process⁴⁹. As for the interactions between the studied ligands and the receptor, Ribavirin, the control drug, exhibited a binding affinity of -6.1 kcal/mol at the SARS-CoV-2 RdRp binding site, interacting through nine H-bonds with residues located at the finger subdomain; five of which by OH groups at 2' and 3' position of the ribose ring with Gly683, Asp684 (motif B), Val560 (motif F), and Thr565, respectively. These H-bond formations highlight the vital role of OH groups at 2' and 3' positions of the Ribavirin's ribose ring to get engaged in H-bonding with crucial residues in the catalytic pocket. The other four H-bonds are seen between 5' OH and Ser501, Ala502 (motif G), along with H-bonds between triazole-3-carboxamide moiety and Asp684, and Thr687 (motif B; Figs. 5, 6 and 7). The four ligands are predicted to bind more strongly to SARS RdRp than Ribavirin, with binding energies ranging from -8.3 to -9.3 kcal/mol. Interestingly, hydroxyl groups at the 2' and 3' positions of sugar rings in our selected ligands play a vital role in making interactions with critical residues in the catalytic pocket, e.g., Lig-1 forms two hydrogen bonds through the 2' and 3' OH groups on sugar ring with the catalytic residues Asp618 (motif A) and Asp760 (motif C), and two with the metallic cofactor Mg²⁺ (A), and one with Arg553 (motif F). This compound forms additional interactions with residues located in motifs A (Pro620, Lys621, and Cys622) and D (Phe793, Ser794, and Lys798) in the catalytic pocket of the enzyme. It was previously shown that closure of motif A stabilizes substrate binding via the coordination of the metallic cofactors by Asp618 and Tyr619, and by forming a hydrogen-bonding network via Asp623, Ser682, Thr687 and Asn691 from motif B¹⁰. Thus, the interactions of Lig-1 with motif A residues may destabilize disrupt the coordination of Mg²⁺ (A) by Asp618, which may prevent H-bonding between Asp623 and the substrate. Lig-2 is bound to the catalytic pocket of the viral target involving seven hydrogen bonds (Figs. 5, 6 and 7); three with Lys621 (motif A) by ribose ring's 2' OH and two with Arg553 (motif F), as well as two other H-bonds with the catalytic Asp618 and Cys622 (motif A). Additional residues in motif A (Pro620, Asp623, Arg624) and motif B (Asn691) form hydrophobic contacts with Lig-2. Three residues from motif F, Lys545, Arg553 and Arg555, form the nucleotide triphosphates (NTPs) entry channel^{10,12}, while the RNA-template enters the active site through a groove that contains the catalytic motifs A and C¹². It is thus noteworthy that Lig-2 may block the NTP entry channel by H-bonding with Arg553, and can also prevent the correct binding of incoming nucleotides to the catalytic pocket by blocking the crucial catalytic Asp618.

Lig-3 forms five potent hydrogen bonds; two H-bonds through 3' OH with Arg553 (motif F), Cys622 (motif A), and by 4' and 5' OH moieties of the ribose ring with catalytic Asp618 (Figs. 5, 6 and 7) in SARS-CoV-2 RdRp. Therefore, this ligand is likely to have a strong potential inhibitory effect against the viral polymerase by targeting the NTP entry channel and causing important disruptions in substrate recognition. Other contacts with residues located in the palm subdomain, such as Asp618, Lys621, Arg624 (catalytic motif A), Ala550 and Lys551, further support this hypothesis. Lig-4 is predicted to form stronger interactions with the viral polymerase compared to the other three candidates as it forms eight strong hydrogen bonds with amino acid residues in the catalytic motifs A and F of the palm subdomain (Figs. 5, 6 and 7). Surprisingly, the role of OH groups at 2' and 3' positions of sugar ring in polar interactions with the catalytic residues of the polymerase is still remarkable since OH at 2' position forms one H-bond with Cys622 (motif A), while OH at 3' position establishes three H-bonds with Thr556 (NTP entry channel, motif F), Cys622 and catalytic Asp623 (motif A). This ligand also makes two hydrogen bonds through its sugar ring with Lys545 and Thr556 (Motif F) (Figs. 5 and 6). The other two H-bonds with Lys621 (motif A) can also promote its conformer in the catalytic site of SARS CoV-2 RdRp. It should be noted that hydrogen bonding with Lys545, and hydrophobic interactions with three other residues of motif F (Lys551, Arg553 and Arg555) can significantly interfere and disrupt the NTP entry channel formation.

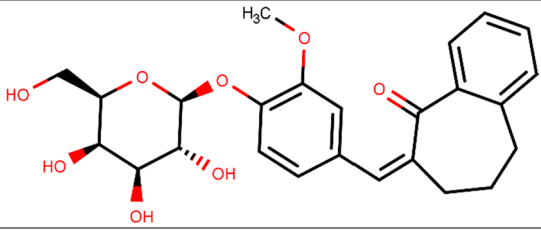
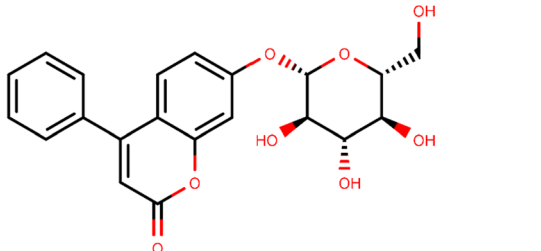
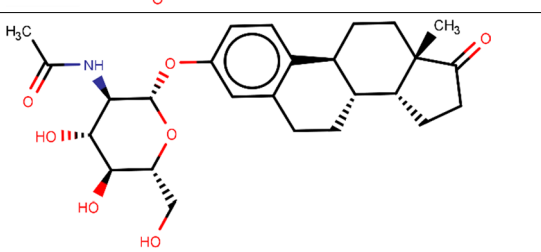
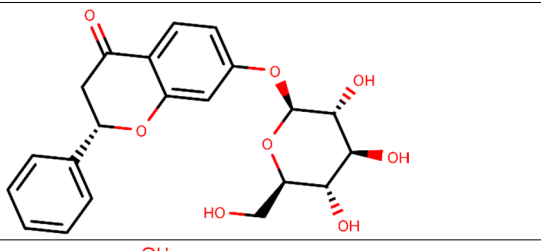
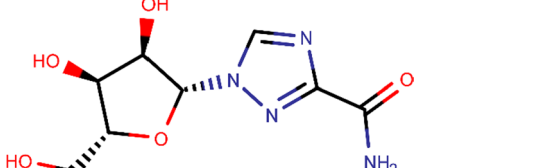
ZINC ID	Structure	Binding Energy (kcal/mol)		
		SARS-CoV-2 RdRp	HIV-1 RT	HCV NS5B
ZINC85875904 (Lig-1)		-8.5	-11.1	-9.0
ZINC4081776 (Lig-2)		-9.1	-10.6	-9.3
ZINC8789971 (Lig-3)		-8.9	-11.3	-10.1
ZINC4096280 (Lig-4)		-9.3	-10.9	-9.6
ZINC1035331 (Ribavirin)		-6.1	-7.1	-6.5

Table 3. 2D structures and binding affinities of the four selected natural ligands and the drug Ribavirin to the viral receptors (*i.e.*, SARS-CoV-2 RdRp, HCV NS5B, and HIV-1RT).

Moreover, multiple, strong hydrogen bonds with catalytic motif A (Asp621, Asp623) may impose remarkable negative effects on the substrate recognition process thus hindering the catalytic activities of the viral polymerase.

To investigate the antiviral potential of the four selected ligands with the known antivirals, their binding energies are compared against benchmark drugs (such as remdesivir, sofosbuvir, and favipiravir) (Table S2 Supporting Information) since it has been studied in previous research⁵⁰. Our molecular docking results shows that three benchmark drugs remdesivir, sofosbuvir, and favipiravir revealed binding affinities of -8.5 kcal/mol, -7.3 kcal/mol, and -5.8 kcal/mol, respectively, which seems weaker to the four selected hits in this study. These results may indicate that our hits perform comparably to, or even better than, the known antiviral drugs since our four ligands reveal comparable binding energies between -8.5 and -9.3 kcal/mol against SARS CoV-2 RdRp.

In summary, the docking analysis indicates that all four ligands bind to the palm subdomain of SARS-CoV-2 RdRp, which forms the catalytic pocket of the polymerase. In particular, they interact with the NTP entry channel and establish multiple hydrogen bonds with key catalytic aspartates (618, 623, 760, 761) located in motif A and C. It is interesting to note that nearly 90% of the polar interactions between ligands and polymerase are governed by the sugar rings of the ligands, especially OH groups at their 2' and 3' positions, which are involved in H-bonds with catalytic motifs A and F (NTP entry channel) (Figs. 5, 6 and 7). Thus, it is plausible that these interactions alter the conformation of the catalytic pocket, block nucleotide entry, and inhibit the recognition and optimal binding of the substrate to the catalytic pocket. The net effect of these interactions is anticipated to

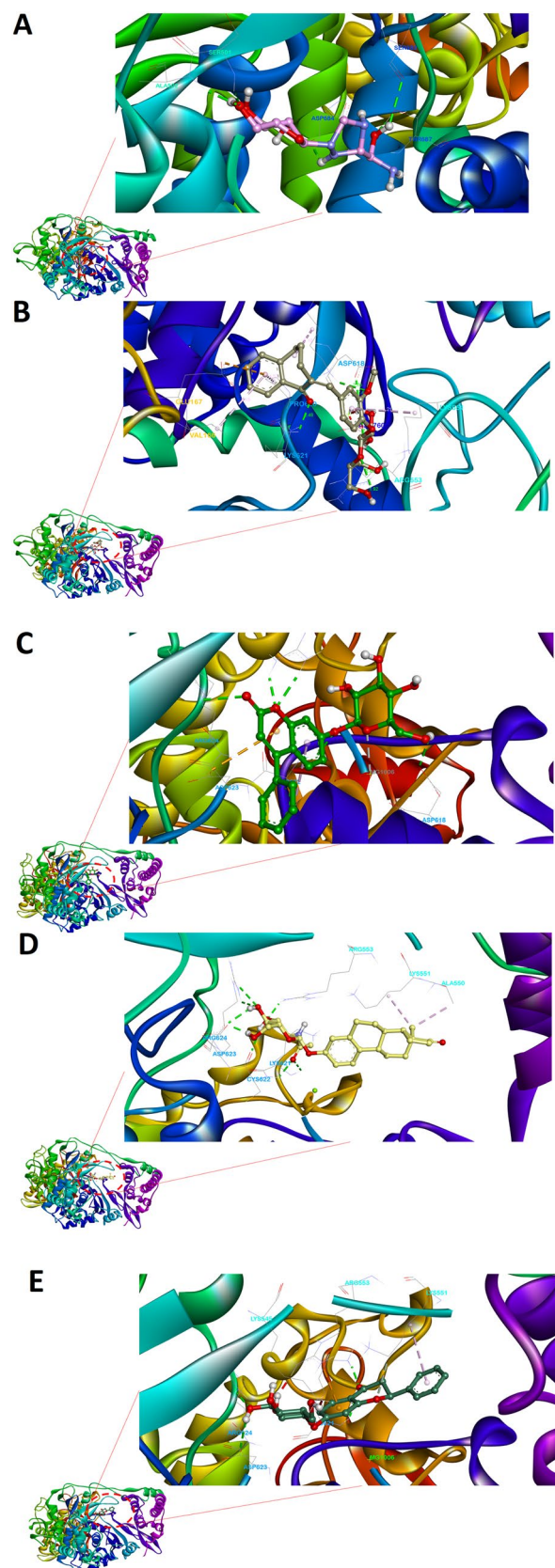


Fig. 5. 3D structural representations illustrating the binding interactions of SARS-CoV-2 RdRp with (A) Ribavirin, (B) Lig-1, (C) Lig-2, (D) Lig-3, and (E) Lig-4.

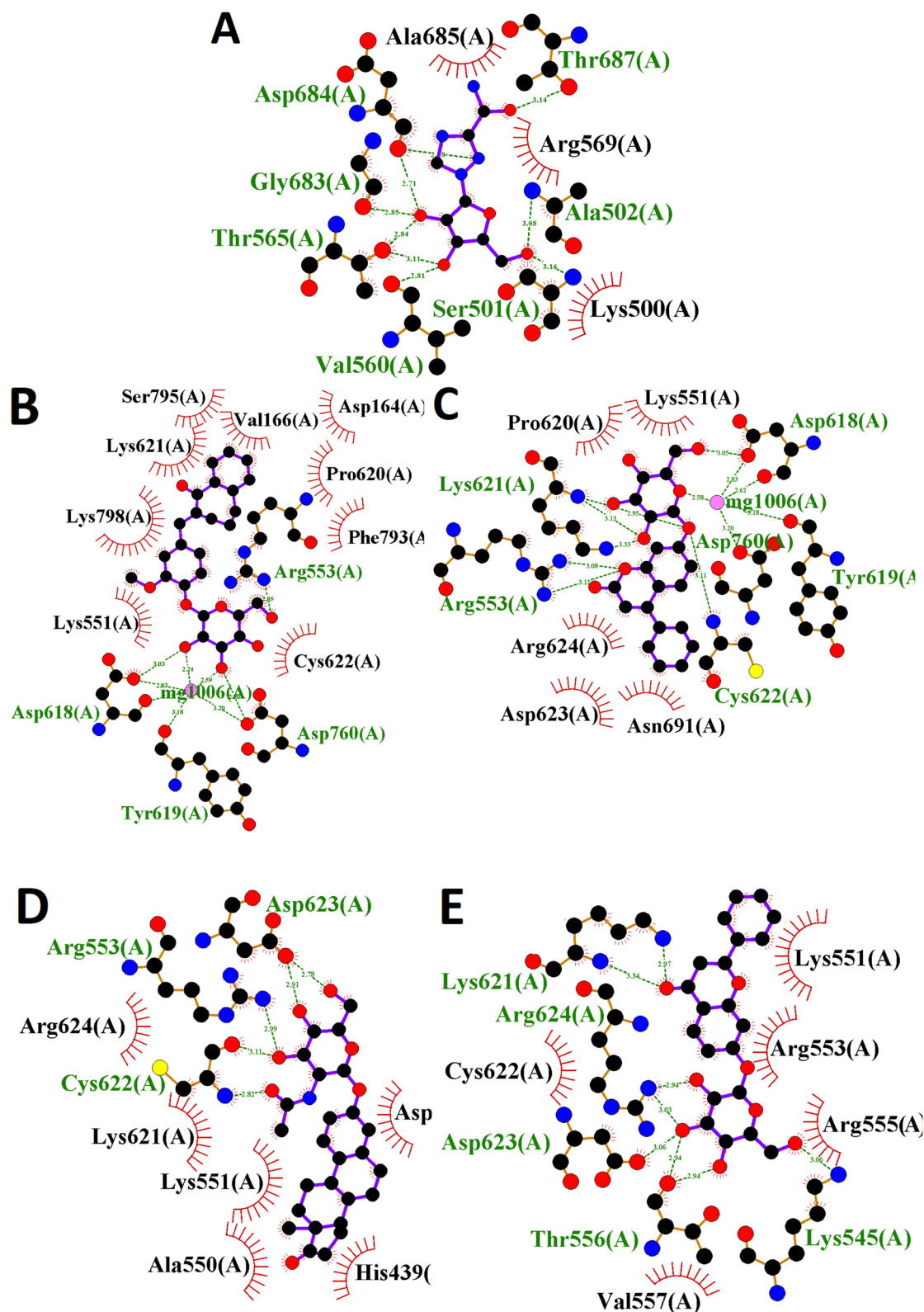
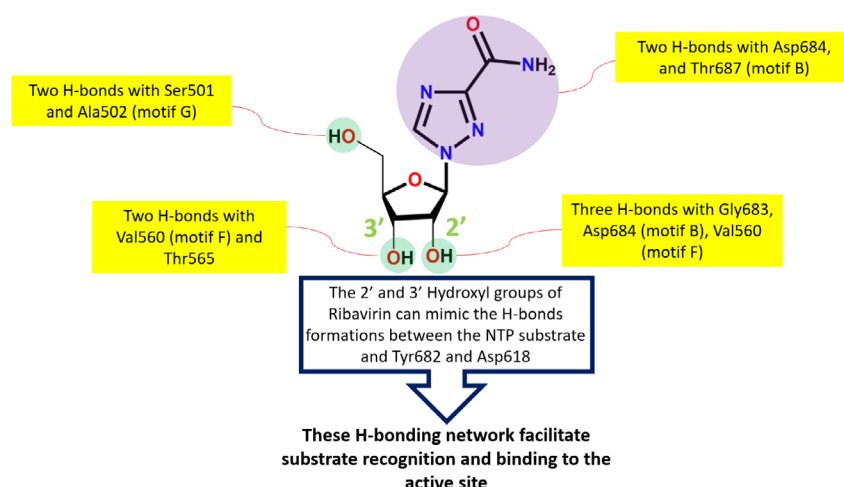
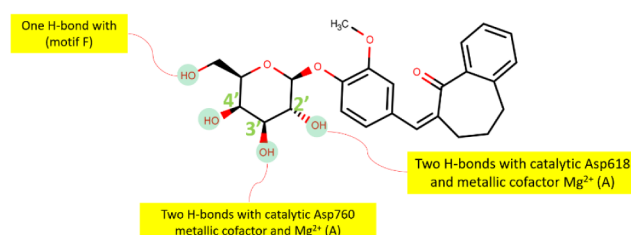


Fig. 6. 2D structural representations illustrating the binding interactions of SARS-CoV-2 RdRp with (A) Ribavirin, (B) Lig1, (C) Lig2, (D) Lig3, and (E) Lig-4. The name of amino acids that interact with the ligands are shown in green, as are the bonds.

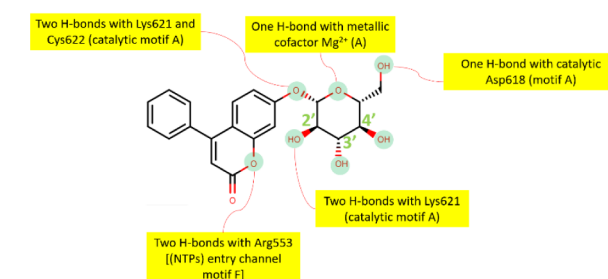
A



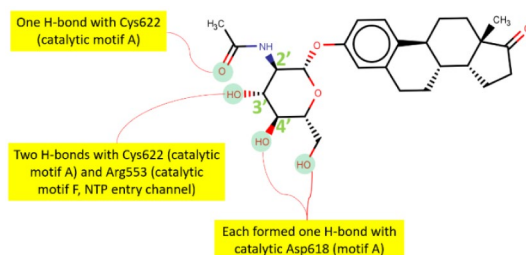
B



C



D



E

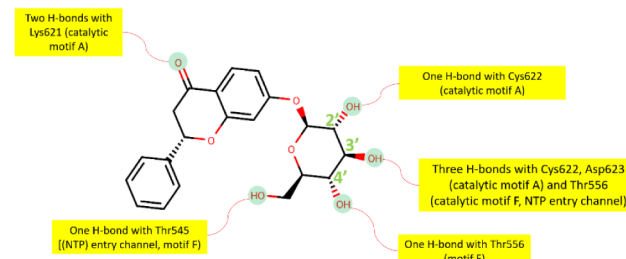


Fig. 7. Pharmacophoric investigations of (A) Ribavirin, (B) Lig-1, (C) Lig-2, (D) Lig-3, and (E) Lig-4 with the SARS-CoV-2 RdRp.

be a significant disruption in the catalytic function of the polymerase. Similar results were also obtained with docking studies involving HIV RT and HCV NS5B (Fig. S2, Supporting Information).

MD simulations

MD simulations of the SARS-CoV-2 RdRp

Benchmarking Strategy: For the subsequent dynamic and energetic analyses (MD simulations and MM/PBSA), Ribavirin was employed as the primary control compound across all three viral polymerase systems. This approach allows for a direct and consistent comparative benchmark to evaluate whether the newly identified natural ligands can match or surpass the binding stability and affinity of a proven broad-spectrum nucleoside analog inhibitor within the identical simulation framework for each target.

Molecular dynamics simulation is a highly effective method for exploring the dynamic nature of proteins at a molecular and atomic level¹⁰. This approach offers new opportunities for drug discovery and formulation. To evaluate the dynamics of each ligand-receptor complex and understand the effects of these interactions on the target viral structures, 200 ns MD simulations were performed for each ligand-enzyme pair. Initially, changes in the RMSD of all protein atoms were measured to determine the system's equilibration state (Fig. 8). Relevant data for SARS-CoV-2 RdRp are discussed below and equivalent information for HIV-1 RT and HCV NS5B are provided in Figs. 8, 9, 10, 11, 12 and 13, along with in the Supplementary Information (Figs S5–S6). For free SARS-CoV-2 RdRp, the RMSD increased sharply to above 0.22 nm within the first ~4 ns of the simulation but

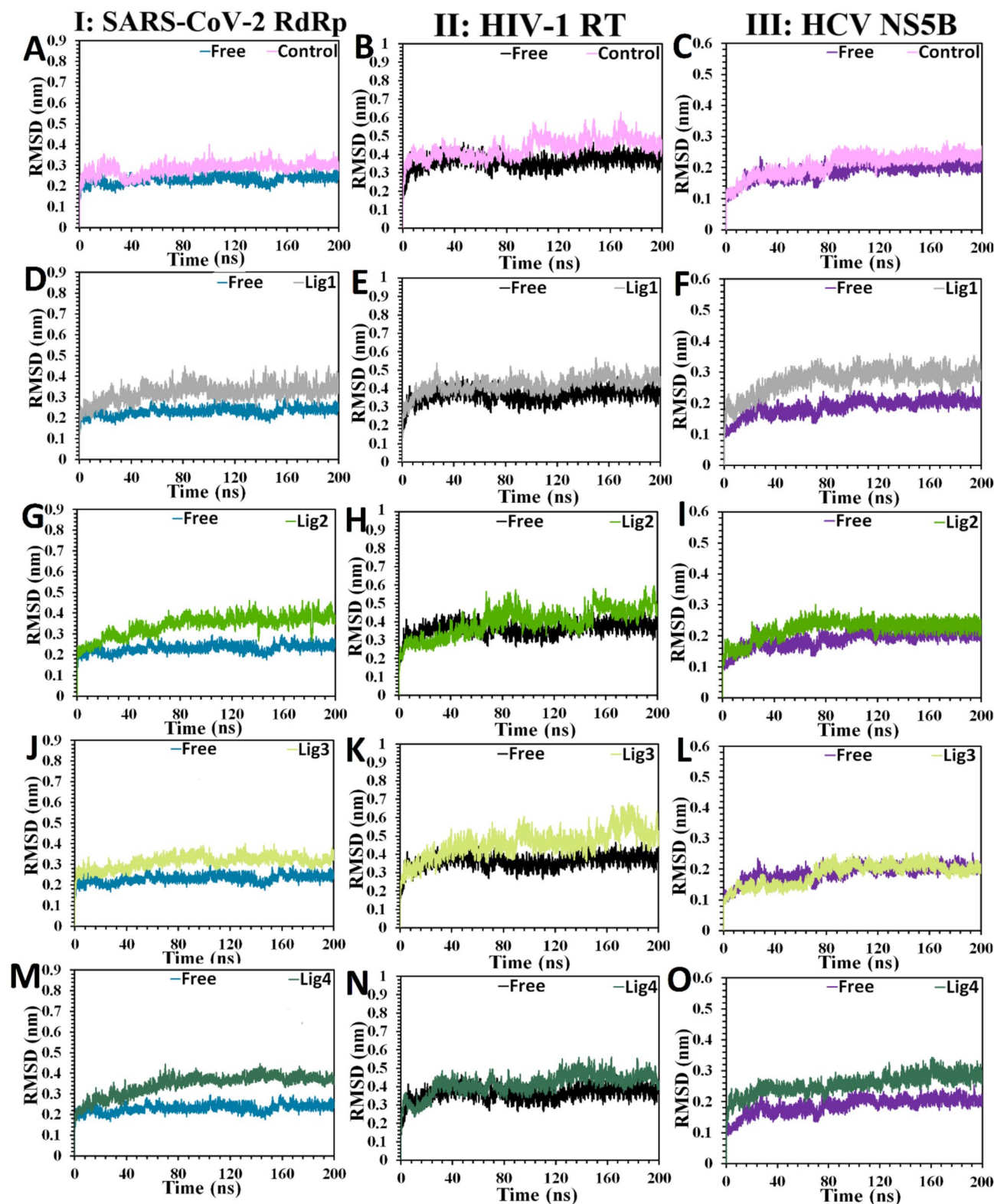


Fig. 8. Changes in RMSD for (I) SARS-CoV-2 RdRp, (II) HIV RT, (III) HCV NS5B in free state and in complex with (A–C) Ribavirin (control), (D–F) Lig-1, (G–I) Lig-2, (J–L) Lig-3, and (M–O) Lig-4, respectively.

then displays only moderate fluctuations at the same level until 100 ns, followed by a modest decrease between 130–150 ns to 0.23 nm until the end of simulation period (Fig. 8A). RMSD values stabilized > 0.23–0.25 nm after 80 ns, indicating a well-equilibrated complex. These findings indicate that the protein reaches its equilibrium state rapidly, and hence 200 ns is sufficiently long for an effective simulation of this system. In presence of

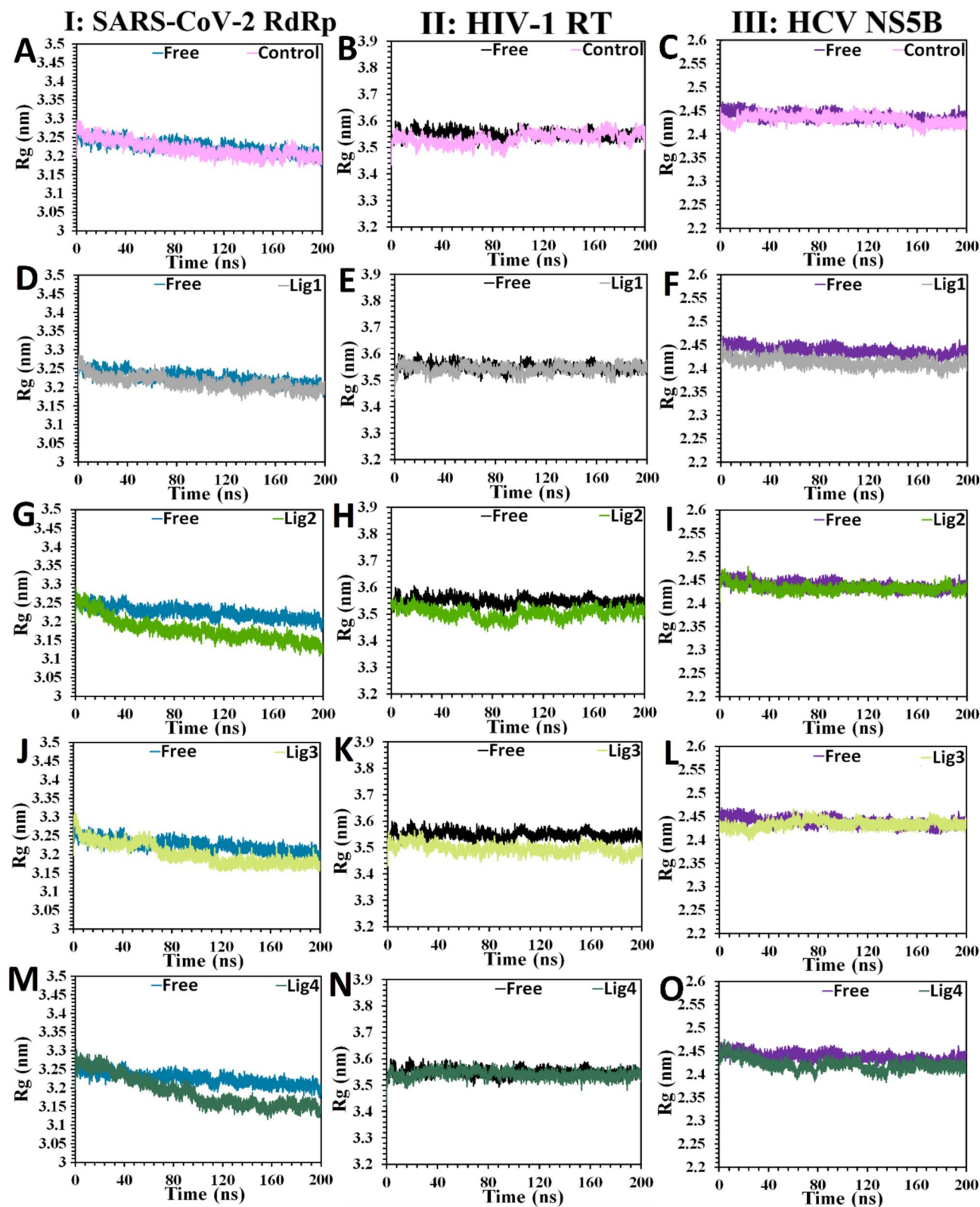


Fig. 9. Changes in R_g for (I) SARS-CoV-2 RdRp, (II) HIV RT, (III) HCV NS5B in free state and in complex with (A–C) Ribavirin (control), (D–F) Lig-1, (G–I) Lig-2, (J–L) Lig-3, and (M–O) Lig-4, respectively.

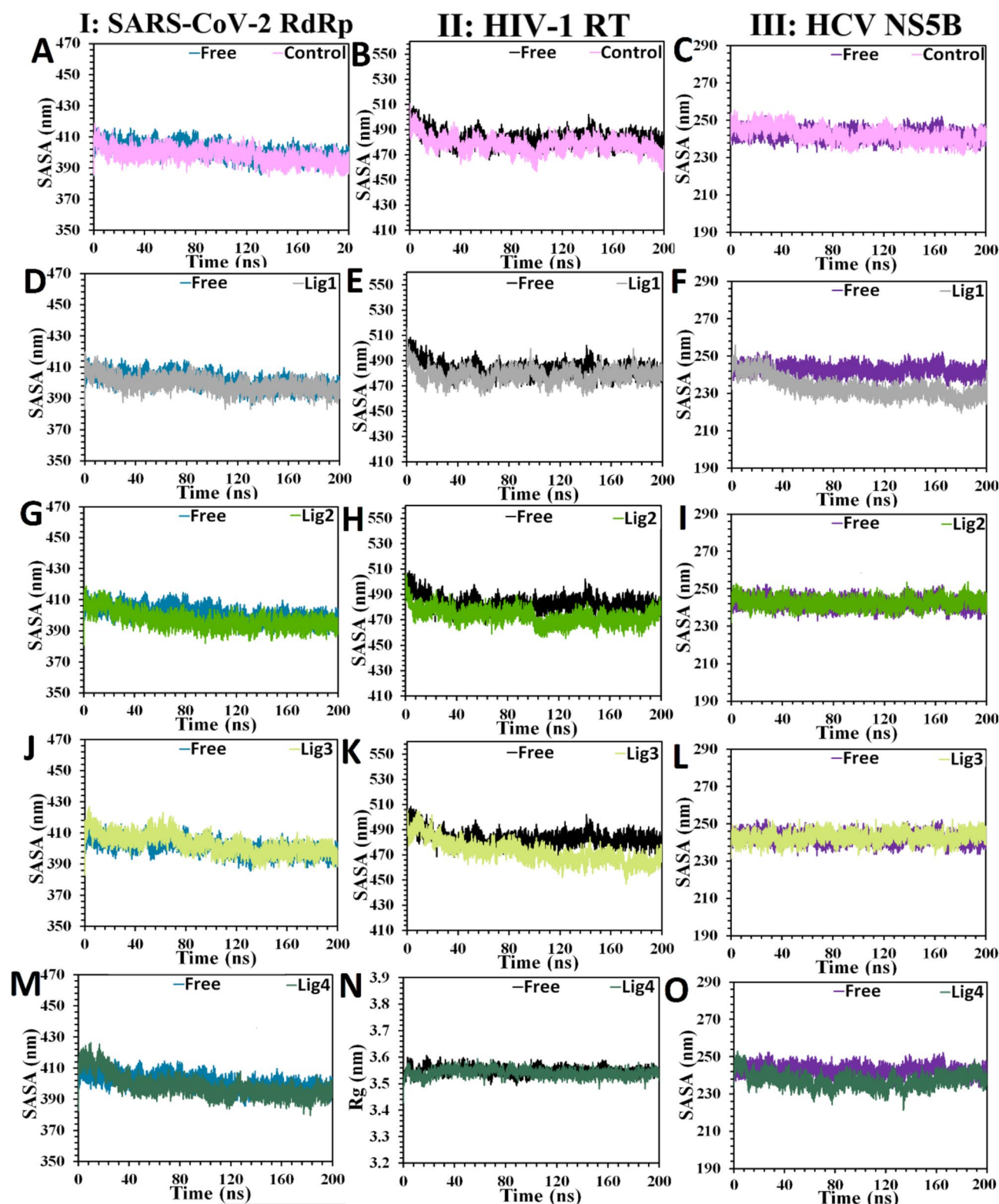


Fig. 10. Changes in SASA for (I) SARS-CoV-2 RdRp, (II) HIV RT, (III) HCV NS5B in free state and in complex with A–C) Ribavirin (control), (D–F) Lig-1, (G–I) Lig-2, (J–L) Lig-3, and (M–O) Lig-4, respectively.

either Ribavirin or the studied ligands, the RMSD values also rapidly increase to above 0.25 nm and then more gradually further increase up to 0.42 nm (Fig. 8D, G, J, M). Interestingly, the presence of the four novel ligands leads to higher RMSDs than Ribavirin, with the Lig-2-RdRp complex displaying the highest average RMSD values of all the systems tested. On average, the overall RMSD values for the free protein are 1.21 Å, and 1.45

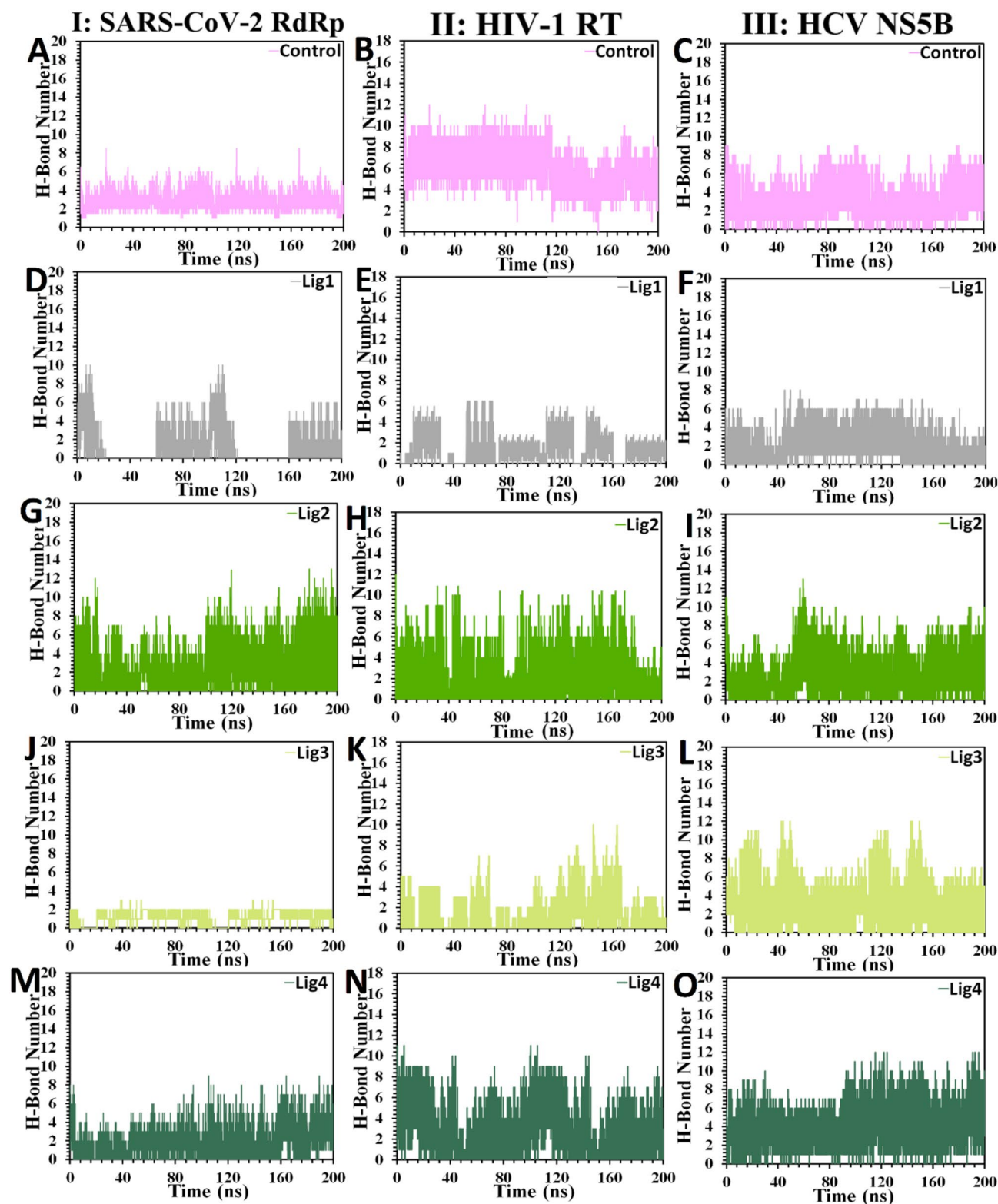


Fig. 11. Changes in H-bond for (I) SARS-CoV-2 RdRp, (II) HIV RT, (III) HCV NS5B in complex with (A-C) Ribavirin (control), (D-F) Lig-1, (G-I) Lig-2, (J-L) Lig-3, and (M-O) Lig-4, respectively.

Å, 1.69 Å, 1.80 Å, 1.74 Å and 1.76 Å for the complexes containing Ribavirin, Lig-1, Lig-2, Lig-3 and Lig-4, respectively, a trend that is in reasonable agreement with the binding energies of these ligands.

The Rg (radius of gyration) measures a protein's average size. Therefore, changes in Rg, either increasing or decreasing, reflect variations in a macromolecule's molecular compactness, suggesting structural compression or

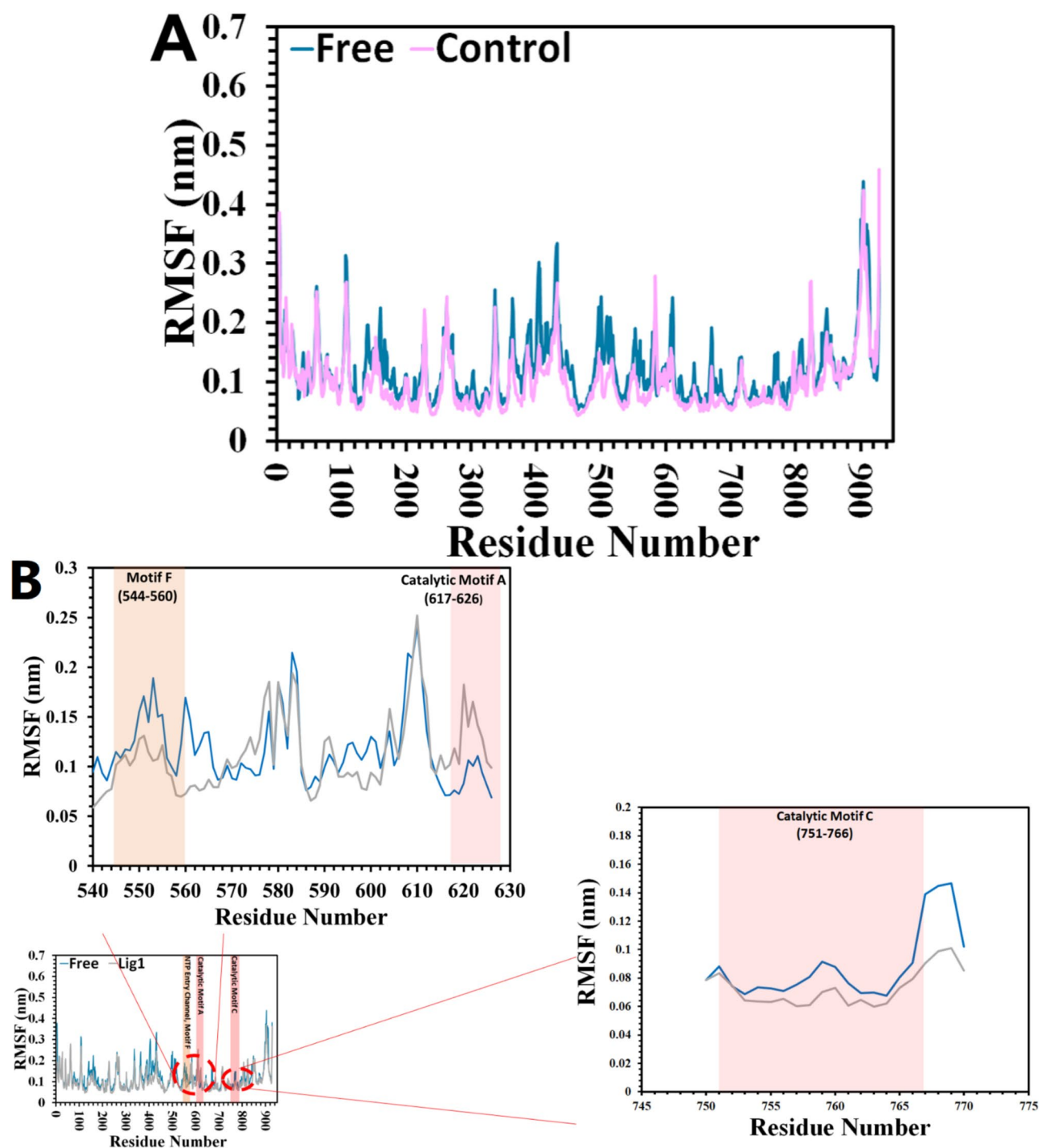


Fig. 12. Changes in the RMSF of SARS-CoV-2 RdRp in its free state and when bound to A) Ribavirin, B) Lig-1, (C) Lig-2, (D) Lig-3, and (E) Lig-4.

expansion¹⁴. The compactness of an enzyme is expected to influence the binding efficiency of a substrate to the binding site^{14,51}. The Rg values for the free protein remains nearly constant throughout the 200 ns simulation, and its fluctuations remained within acceptable range of ± 0.1 – 0.2 nm, suggesting preserved compactness, with a mean value of 3.19 nm (31.99 Å) (Fig. 9A). However, for all the ligand-bound systems compressions are observed. For the Ribavirin-bound system, the average Rg decreases to 31.94 Å (Fig. 9A). For the ligands-bound systems, the corresponding values are 31.90 Å, 31.70 Å, 31.86 Å and 31.77 Å, for Lig-1, Lig-2, Lig-3, and Lig-4, respectively, a trend that is comparable with that observed for the RMSD values; again, the effect is greatest for Lig-2 (Fig. 9D, G, J, M). Interestingly, these declining in the Rg magnitudes can highlight the compression in protein size upon ligand binding, which reduce the size of the active site, thus preventing substrate molecules from binding. Importantly, the viral protein in complex with Lig-2, Lig-4, and Lig-3 showed the higher

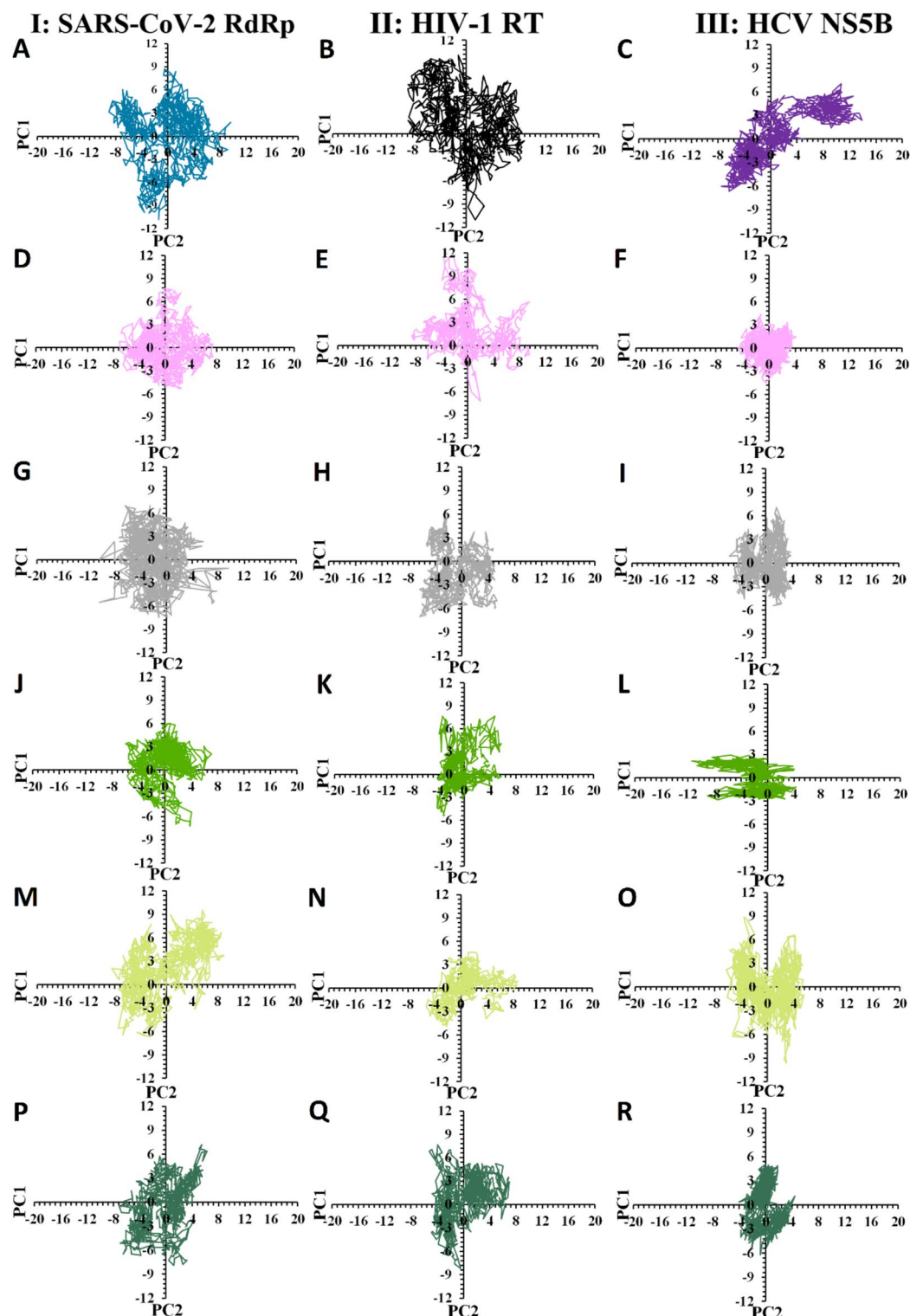


Fig. 13. Changes in the Principal Component Analysis (PCA) for (I) SARS-CoV-2 RdRp, (II) HIV RT, (III) HCV NS5B in free state (A–C), and in complex with (D–F) Ribavirin (control), (G–I) Lig-1, (J–L) Lig-2, (M–O) Lig-3, and (P–R) Lig-4, respectively.

fluctuations for Rg magnitudes compared to that of the Ribavirin, and such significant fluctuations in the viral enzyme can destabilize the structure of protein, leading to disruptive effect on its catalytic activities.

This interpretation is supported by an analysis of the solvent accessible surface area (SASA) for these systems (Fig. 10D, G, J, M). In general, SASA values fluctuated within a narrow range without systematic increase, indicating no unfolding, and these values are smaller for the protein–ligand complexes in agreement with their

more compact structures. Notably, the complexes containing Lig-2 and Lig-4 exhibit the highest magnitudes of reduce in the Rg and SASA average values. These values are even lower than those of Ribavirin, as standard drug, which demonstrates the higher compactness in the enzyme structure upon binding to the two compounds. Notably, this compression can decline solvent exposure, affecting internal water channels and blocking diffusion, which can destabilize the protein structure and impair its enzymatic function.

To enable quantitative interpretation of MD simulations, explicit stability thresholds were defined. Structural stability was investigated by RMSD analysis, almost RMSD values of all selected ligand–polymerase complexes rapidly converged and fluctuated within the acceptable range of 0.20–0.30 nm, showing structural equilibration. Compression was also assessed by Rg data, showing acceptable fluctuations limited to ± 0.1 –0.2 nm and no long-term drift was observed. SASA analysis was checked out to ensure absence of progressive unfolding, demonstrating preserved slight compactness without signs of unfolding. Residue-level flexibility was analyzed by RMSF data, where structured regions were expected to remain > 0.15 –0.20 nm, while higher fluctuations in loop regions were considered normal. Binding free energies were calculated using MM/PBSA over the equilibrated trajectory window, and convergence was confirmed by stable energy plateaus.

The docking analysis (Figs. 5, 6 and 7) indicated significant differences in binding interactions between the enzyme and the five ligands. In particular, H-bond interactions are important in maintaining ligand stability in the active site of a protein⁵². We thus analyzed the number of H-bonds engaged in molecular interactions between the ligand and protein over 200 ns (Fig. 11D, G, J, M). For each ligand, the number of H bonds fluctuates throughout the simulation time, and in general our four selected ligands are stabilized by at least three H-bonds on average, except for Lig-3 which formed, on average, two H-bonds within the viral polymerase binding site.

An analysis of root-mean-square fluctuations (RMSFs) was applied to assess the extent to which the position of each amino acid in a protein fluctuates. It measures the standard deviation of the displacements of individual residues in relation to the total protein displacement. Higher RMSF values indicate more unstable amino acids, while lower values suggest greater stability¹⁴. These magnitudes can vary when the protein interacts with a ligand. Binding of the Ligands 1–4 and the Ribavirin to SARS-CoV-2 RdRp altered the RMSF of protein amino acids in several regions of the enzyme. In terms of the control drug, remarkable RMSF fluctuations were experienced by residues 11–30, 34, and 227–230 situated at the N-terminal domain, residues 259–263 in the interface domain, residues 582–585 (palm subdomain), 821–826, and 829 situated at the thumb subdomain. However, RMSF fluctuations decreased in some parts of the protein after binding of Ribavirin, for example at residues 41, 62, 107–109, 125–139, 163–182, 203 located at the N-terminal domain, 252, 260, 304–312, 337, 363–365, 386–399 in the interface domain, 404–413, 431, 444–461, 494–512 (motif G), 514–516, 543, and 554–565 (motif F) in the finger subdomain, residues 607–611 (motif A in the palm subdomain), 643, 673–685 (motif B in the finger subdomain), 720–725, 767–771, 780–782, 795 (motif D), and 807–812 (motif D) in the palm subdomain, and finally residues 845–848, 903–905 in the thumb subdomain (Fig. 12A). Compared to Ribavirin, significant differences in RMSF values were observed after binding of the four ligands, particularly around those residues responsible for the catalytic functions and substrate binding. Binding of Lig-1 to the viral polymerase caused increases in the RMSF magnitudes of residues 14–27, 32–49, 51–53, 59–66, 224–230 (N-terminal domain), 300–304, 325–331 (interface domain), 617–626 (catalytic motif A, palm subdomain), 684–696 (motif B, palm subdomain), 711–720, 738–747 (palm subdomain), and 803–806 (palm subdomain). However, remarkable decreases can be observed in the RMSF values of residues 7, 68–78, 87–92, 107, 113–118, 123–135, 139–147, 165–182, 197, 203, 236–249 (N-terminal domain), 251–254, 262, 277, 285, 293, 336–338, 364–366, 369–379, 389–392 (Interface domain), 395–401, 407–421, 431–433, 447–456, 494, 497–511 (motif G, fingers subdomain), 516–519, 552–555 (NTP entry channel, finger subdomain), 557–562 (motif F, finger subdomain), 643, 654 (finger subdomain), 670–685 (motif B, finger subdomain), 751–766 (catalytic motif C, palm subdomain), 791–797 (palm subdomain), 848, 903, and 929 (thumb subdomain) (Fig. 12B).

Upon binding of Lig-2, increased RMSF values can be seen in residues 6–13, 151–156, 169–172, 191–199, 200–211, 221–229 of the N-terminal domain, residues 271, 278–289, 291–293 (Interface domain), 458, 462, 650–663 in the finger subdomain, residues 682–685, 689–702 and 706–709 (motif B, palm subdomain), 711–732, 742–749, (palm subdomain), 750–764 (catalytic motif C, palm subdomain), 771–780, 788–795 (motif D, palm subdomain), 811–818 (motif E, palm subdomain), 825, 850–852, 866, 882, and 903–910 (thumb subdomain). Nonetheless, decreased RMSF values were observed at amino acids 17, 40, 51, 77–85, 92, 109, 121, 135, 139–151, 177 of the N-terminal domain, residues 262, 304, 362–366, 381–393 of Interface domain, residues 411–416, 421–427, 430, 447 of finger subdomain, residues 496–508, 511 in motif G, finger subdomain, residues 518, 539–541, 551–554 in finger subdomain, amino acids 552–555 at NTP entry channel, finger subdomain, residues 557–560 (motif F, finger subdomain), 576–581 (finger subdomain), 601–607 and 612–626 (motif A, palm subdomain), residues 643 and 670 of palm subdomain (Fig. 12C).

In the Lig-3-SARS-CoV-2 RdRp complex, RMSF fluctuations were seen to rise at residues 8–19, 25, 31–35, 44–58, 60, 67, 80–87, 95–98, 107, 147–164, 166, 184–189, 203, 225–229, 236, and 243–249 of the N-terminal domain, positions 254–257, 262–270, 286–290, 300–304, 307–319, 325–331, 350, 386–397 of the interface domain, residues 403–409, 443–458, 470–490, 497–499, 523–529, and 694–698 of the finger subdomain, residues 693–710 (motif B) 713–720, 726–739, 740–750 (palm subdomain), 758–765 (catalytic motif C), 823, 847–854, 901–907, and 929 (thumb subdomain). However, the RMSF magnitudes diminished for amino acid positions 107–108, 140, 166, 178, and 181 of the N-terminal domain, 336 and 363–365 of the interface domain, residues 410–415, 429–432, 452, 455, 549–556 (NTP entry channel, motif F), 557–560 (motif F), 562–564, and 580–584 in the finger subdomain, residues 609–622 (motif A), 644, 668–670, 672–685 (motif B), 797, 804, 806–809 located in the palm subdomain, and position 873 in the thumb subdomain (Fig. 12D).

As for the Lig-4 complex, the RMSF magnitudes increased in the majority of the regions of the viral enzyme, in particular in residues 8, 18–27, 29–48, 56–67, 77–92, 114–122, 147–156, 169, 198–200, 213, 227, 243–245 in the N-terminal domain, positions 262, 265, 293, 302, 319–323, 382–389 located in the interface domain,

Compounds	Elec (kJ/mol)	VdW (kJ/mol)	Total	H-bond
SARS-CoV-2 RdRp				
Ribavirin	-53.00 ± 5.89	-50.39 ± 4.96	-103.39 ± 6.7	4.5 ± 0.3
Lig-1	-64.48 ± 3.9	-67.67 ± 5.2	-132.15 ± 6.7	2.5 ± 0.5
Lig-2	-65.28 ± 7.9	-83.88 ± 5.3	-149.17 ± 8.3	3.5 ± 0.3
Lig-3	-36.61 ± 5.7	-89.43 ± 2.9	-126.04 ± 5.6	2.3 ± 0.4
Lig-4	-41.42 ± 2.39	-92.01 ± 6.8	-133.43 ± 2.20	3.9 ± 0.5
HIV-1 RT				
Ribavirin	-52.85 ± 4.79	-72.73 ± 2.64	-125.58 ± 2.15	3.9 ± 0.5
Lig-1	-34.12 ± 5.26	-129.57 ± 3.31	-163.69 ± 8.58	0.6 ± 0.3
Lig-2	-41.86 ± 9.81	-112.41 ± 3.00	-154.26 ± 6.81	3.8 ± 0.5
Lig-3	-23.76 ± 7.57	-97.38 ± 5.44	-121.13 ± 2.13	1.6 ± 0.4
Lig-4	-43.05 ± 3.18	-86.30 ± 3.81	-129.36 ± 0.63	3.4 ± 0.5
HCV NS5B				
Ribavirin	-64.77 ± 5.93	-164.11 ± 5.57	-228.88 ± 1.36	4.1 ± 0.6
Lig-1	-43.18 ± 3.14	-144.59 ± 0.98	-187.77 ± 4.12	4.0 ± 0.6
Lig-2	-39.63 ± 3.83	-127.37 ± 5.18	-167.00 ± 4.98	3.9 ± 0.5
Lig-3	-30.29 ± 2.52	-101.47 ± 2.59	-131.75 ± 0.07	2.6 ± 0.5
Lig-4	-26.52 ± 5.58	-144.67 ± 3.54	-171.19 ± 8.12	3.1 ± 0.5

Table 4. Average binding energies of Ribavirin and Ligands 1–4 toward SARS-CoV-2 RdRp, HIV-1 RT, and HCV NS5B were computed under dynamic conditions.

residues 402–409, 494–496 (motif G), 712–714, 734–738, 757–764 (catalytic motif C), 797, and 823–825 in the palm subdomain, and finally positions 901–910 in the thumb subdomain. Conversely, RMSF values were seen to decrease for residues 107, 147, 155, 172–174, and 180 in the N-terminal domain, residues 265, 337, and 363 in the interface domain, residues 425, 450–456, 510–514, 551–564 (NTP entry channel, motif F), 568–572, and 580–583 in the finger subdomain, residues 611–626 (motif A), 643–645, 669–672 in the palm subdomain, and positions 847–849 located in the thumb subdomain (Fig. 12E). Taken together, the analysis of RMSF values suggests that, compared to Ribavirin, the binding of the four compounds L1–L4 to SARS-CoV-2 RdRp induces significant changes in the fluctuation levels of amino acid residues, especially those situated in the catalytic motifs A, C, and F, as well as in the NTP entry channel. These residues are crucial for either RNA binding within the catalytic pocket or the polymerase's catalytic activity. Therefore, binding of the Ligands 1–4 is expected to impair the catalytic ability of the polymerase, either indirectly by obstructing the NTP entry channel, which leads to significant disturbances in RNA binding to the active site, or directly by affecting catalytically-critical residues, thus ultimately hindering the replication of the virus. Similar observations were also made when analyzing the interactions between these ligands and HIV-1 RT and HCV NS5B (Fig. S5, Supporting Information).

To examine the stability of the ligands within the catalytic pocket, the time-dependent changes in the average distance between the active site and the ligands and were monitored (Fig. S6, Supporting Information). The four newly-discovered ligands are located at mean distances of between 0.27 nm to 0.38 nm to the catalytic pocket of the three viral polymerases, with minimal fluctuations. These average distances from the center of the receptors' catalytic site is favorable for strong interactions, and consequently, this study provides strong evidence that the natural ligands discovered here are spatially well-aligned within the binding pockets of the three viral polymerases to form stable complexes throughout the duration of the MD simulations.

PCA can be used to assess the primary movements of proteins. These essential movements are crucial to the structure and function of macromolecules, and any alterations in these patterns can result in protein dysfunction¹⁴. The 2D paradigm of viral polymerase movements in the free state and when bound to the ligands has been retrieved via PCA (Fig. 13). Free SARS-CoV-2 RdRp demonstrated a somehow large movement paradigm, varying from -9.0 to 10 nm (Fig. 13A). Interestingly, all of ligands changed the paradigm and movement range of the protein (Fig. 13D, G, J, M, P), suggesting that these ligands may strongly destabilize and inhibit the SARS-CoV-2 RdRp structure and function. In comparison to Ribavirin, Lig-2 and Lig-4 showed stronger potential to alter the movement pattern and motion range of the protein, consistent with the Rg findings reported above. Together, these analyses indicate that the structural compression of the enzyme, caused by binding the ligands, might cause serious disruptions in the catalytic function.

Binding free energies for interactions between Ribavirin and ligands (Lig-1 to Lig-4) with the viral polymerases were estimated via the MM/PBSA approach (Table 4). In SARS-CoV-2 RdRp complexes, compared to the Ribavirin, the complexes containing Lig-2 and Lig-4 have established stronger complexes with the target with the total binding energies of -149.17 ± 8.3 kJ/mol and -132.15 ± 6.7 kJ/mol, correspondingly. The other ligands, Lig-1 and Lig-3, also exhibited more favorable total binding energies, of -126.04 ± 5.6 kJ/mol and -132.15 ± 6.7 kJ/mol, than the standard drug. Additionally, nearly most of the studied natural compounds, in particular Lig-2 and Lig-4, established superior van der Waals contacts, and comparable-superior electrostatic energies in comparison to the Ribavirin, which might be the main cause for their stronger total binding energies.

Majority of binding energies arise from van der Waals forces, while the largest contributions from van der Waals and electrostatic contacts are associated with Lig-2, and to a lesser extent, Lig-4 and Lig-1.

As for the complex of HIV-1 RT with the studied ligands, Lig-1 and Lig-2, with the binding energies of -163.69 ± 8.58 kJ/mol and -154.26 ± 6.81 kJ/mol, account for the strongest binding energies, which are higher than that of the control drug (-125.58 ± 2.15 kJ/mol). Astonishingly, Lig-2 also made a comparable complex with the HIV-1 RT in comparison to the Ribavirin. The largest share of interaction energies in all complexes comes from van der Waals interactions, which seem to be the main driving forces in all other complexes, while the share of electrostatic energies in the complexes containing Ribavirin, Lig-2, and Lig-4 are highlighted.

Among the three viral polymerases, HCV NS5B exhibited the strongest binding with Ribavirin -228.88 ± 1.36 kJ/mol, followed by Lig-1 (-187.77 ± 4.12 kJ/mol) and Lig-4 (-171.19 ± 8.12 kJ/mol). Van der Waals forces also dominated across all complexes. Almost below one-third of the total binding energy for the complex with Ribavirin is electrostatic forces, while the share of electrostatic forces in the complexes with the other ligands seems to be one-fourth of the total energy, which is why Ribavirin has shown the strongest complex with the HCV NS5B. But, compared to the Ribavirin, the share of van der Waals forces in systems containing Lig-1 and Lig-4 is pronounced in comparison to that of the Ribavirin.

Conclusion

In this computational study, a rational pharmacophore-based discovery process was employed using Ribavirin as a control drug to discover novel multi-target inhibitory agents for aspartyl-based viral polymerases (*i.e.*, SARS-CoV-2 RdRp, HIV-1 RT, and HCV NS5B) from the virtual screening of a library of 200,583 natural compounds. The screening process began by evaluating Lipinski's rule of five (RO5), drug-likeness, and pan-assay interference compounds (PAINS), which narrowed down the number of natural compounds to 104,047 compounds. The choice of Ribavirin as the core pharmacophore then allowed filtering of this natural compounds library down to 1801 potential targets. These natural compounds were virtually screened through AutoDock Vina, and four final compounds (Ligands 1–4) were selected based on their high binding affinities to the catalytic pockets of the target polymerases, as determined by consensus docking and ADME/T properties studies. Based on molecular docking results, the four hits bound to the SARS-CoV-2 RdRp via interactions within the NTP entry channel and catalytic aspartates (Asp618, Asp623, Asp760, and Asp761). They also bind to the HIV-1 RT catalytic pocket via polar and hydrophobic contacts, thereby likely disrupting RT function. All of ligands also target catalytic motifs A, B and C in the HCV NS5B enzyme.

MD simulations revealed significant structural changes in the three viral polymerases upon binding of the ligands. The largest RMSDs were observed upon binding of Lig-2 and Lig-4 to SARS-CoV-2 RdRp, Lig-2 and Lig-3 to HIV-1 RT, and Lig-1, Lig-2, and Lig-4 to HCV NS5B. Importantly, these RMSDs were at least comparable if not higher than those observed upon the binding of the benchmark drug Ribavirin. Overall, the most significant structural changes (as measured by average RMSDs) upon binding of these ligands were observed for SARS-CoV-2 RdRp, followed by HCV NS5B and HIV-1 RT. These changes in RMSDs are in good agreement with simulated Rg values. The most significant structural compressions were observed upon binding of Lig-2 and Lig-4 to SARS-CoV-2 RdRp, Lig-2 and Lig-3 to HIV-1 RT and Lig1, Lig-2 and Lig-4 to HCV NS5B. The increased compactness of the three polymerases upon binding of the four ligands was also demonstrated by an analysis of their respective SASA values, indicating that solvent access to the interior of these enzymes is restricted in presence of the ligands. The greatest effects are observed upon binding of Lig-1, Lig-2 and Lig-4, followed by Lig-3 and Ribavirin. Furthermore, the PCA analysis indicated that the mobility of the viral proteins was substantially altered upon ligand binding, especially by Lig-2 and Lig-4 for SARS-CoV-2 RdRp, Lig-2 and Lig-3 for HIV-1 RT, and Lig-1, Lig-2, and Lig-4 for HCV NS5B. All of the ligands kept consistent distances from the catalytic pocket during the simulation period, demonstrating stable binding interactions.

The analysis of RMSF values demonstrated that in particular residues located within the catalytic motifs A, C and E, and in the NTP entry channel in the case of SARS-CoV-2 RdRp, the catalytic YMDD motif in HIV-1 RT, and the NTP entry tunnel and catalytic motifs A, B and C in HCV NS5B are affected by ligand binding. These changes may directly affect the binding of substrates to their active sites, thereby hindering polymerase function and disrupting viral replication. This study thus provides strong evidence that the natural ligands identified here bind efficiently to the active sites of the three viral polymerases, forming stable complexes. The binding energies of the ligands and viral polymerases, computed using the MM/PBSA method, show that Lig-2 and Lig-1 surpass Ribavirin in SARS-CoV-2 RdRp and HIV-1 RT complexes. In the HCV NS5B complex, Ribavirin exhibits the strongest binding energy, followed by Lig-1 and Lig-4. Van der Waals interactions are the main driving forces across all systems, with electrostatic interactions most prominent in Ribavirin and Lig-1. Lig-1 and Lig-2 show comparable H-bonding to the control drug. Computational analyses and pharmacological predictions also suggest that in particular Lig-2 has great potential as effective drug lead against each of the targets tested here, but Lig-4 (against SARS-CoV-2 RdRp and HCV NS5B), Lig-1 (against HCV NS5B) and Lig-3 (against HIV-1 RT) are also promising leads for the development of multi-target antiviral compounds.

Data availability

All the associated with this work are presented here (and its Supplementary Information file) and further will be available from the corresponding author on reasonable request.

Received: 28 June 2025; Accepted: 16 March 2026

Published online: 26 May 2026

References

- Chu, J. T. S. & Lamers, M. M. Organoids in virology. *npj Viruses* **2**, 5 (2024).
- Manns, M. P. et al. Hepatitis C virus infection. *Nat. Rev. Dis. Primers* **3**, 17006 (2017).
- WHO. Hepatitis C, <<https://www.who.int/news-room/fact-sheets/detail/hepatitis-c>> ((2021)).
- Casey, J. L., Feld, J. J. & MacParland, S. A. Restoration of HCV-specific immune responses with antiviral therapy: A case for DAA treatment in acute HCV infection. *Cells* **8**, 317 (2019).
- Bekker, L.-G. et al. HIV infection. *Nat. Rev. Dis. Primers* **9**, 42 (2023).
- Wang, C., Horby, P. W., Hayden, F. G. & Gao, G. F. A novel coronavirus outbreak of global health concern. *The Lancet* **395**, 470–473 (2020).
- Lee, D. et al. Bioengineered amyloid peptide for rapid screening of inhibitors against main protease of SARS-CoV-2. *Nat. Commun.* **15**, 2108 (2024).
- Tran, V.-T., Porcher, R., Pane, I. & Ravaut, P. Course of post COVID-19 disease symptoms over time in the ComPaRe long COVID prospective e-cohort. *Nat. Commun.* **13**, 1812 (2022).
- Mizrahi, B. et al. Longitudinal symptom dynamics of COVID-19 infection. *Nat. Commun.* **11**, 6208 (2020).
- Malone, B. F. et al. Structural basis for substrate selection by the SARS-CoV-2 replicase. *Nature* **614**, 781–787 (2023).
- Christy, M. P., Uekusa, Y., Gerwick, L. & Gerwick, W. H. Natural products with potential to treat RNA virus pathogens including SARS-CoV-2. *J. Nat. Prod.* **84**, 161–182 (2021).
- Gao, Y. et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **368**, 779–782 (2020).
- Aranda, J., Wiczcór, M., Terrazas, M., Brun-Heath, I. & Orozco, M. Mechanism of reaction of RNA-dependent RNA polymerase from SARS-CoV-2. *Chem. Catal.* **2**, 1084–1099 (2022).
- Ansari, M. et al. Computational investigation on the potential inhibitory effect of fungal metabolites against viral RdRp of SARS-CoV-2 and HCV. *J. Mol. Liq.* **390**, 123153 (2023).
- Malone, B., Urakova, N., Snijder, E. J. & Campbell, E. A. Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* **23**, 21–39 (2022).
- O’Farrell, D., Trowbridge, R., Rowlands, D. & Jäger, J. Substrate complexes of Hepatitis C virus RNA polymerase (HC-J4): Structural evidence for nucleotide import and de-novo initiation. *J. Mol. Biol.* **326**, 1025–1035 (2003).
- Sarafianos, S. G. et al. Structure and function of HIV-1 reverse transcriptase: Molecular mechanisms of polymerization and inhibition. *J. Mol. Biol.* **385**, 693–713 (2009).
- Padhi, A. K., Shukla, R., Saudagar, P. & Tripathi, T. High-throughput rational design of the remdesivir binding site in the RdRp of SARS-CoV-2: Implications for potential resistance. *Iscience* <https://doi.org/10.1016/j.isci.2020.101992> (2021).
- Beigel, J. H. et al. Remdesivir for the treatment of Covid-19. *N. Engl. J. Med.* **383**, 1813–1826 (2020).
- Jayk Bernal, A. et al. Molnupiravir for oral treatment of Covid-19 in nonhospitalized patients. *N. Engl. J. Med.* **386**, 509–520 (2022).
- Aw, D. Z. H., Zhang, D. X. & Vignuzzi, M. Strategies and efforts in circumventing the emergence of antiviral resistance against conventional antivirals. *npj Antimicrob. Resist.* **3**, 54 (2025).
- Palazzotti, D. et al. Small molecule drugs targeting viral polymerases. *Pharmaceuticals* **17**, 661 (2024).
- Sterling, T. & Irwin, J. J. ZINC 15 – Ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
- Lagorce, D., Bouslama, L., Becot, J., Miteva, M. A. & Villoutreix, B. O. FAF-Drugs4: Free ADME-tox filtering computations for chemical biology and early stages drug discovery. *Bioinformatics* **33**, 3658–3660 (2017).
- Mathur, P., Kottitil, S. & Wilson, E. Use of ribavirin for Hepatitis C treatment in the modern direct-acting antiviral era. *J. Clin. Transl. Hepatol.* **6**, 431 (2018).
- Kamzееva, P. N., Aralov, A. V., Alferova, V. A. & Korshun, V. A. Recent advances in molecular mechanisms of nucleoside antivirals. *Curr. Issues Mol. Biol.* **45**, 6851–6879 (2023).
- Wahyuni, T. S. et al. Inhibition of hepatitis C virus replication by chalepin and pseudane IX isolated from *Ruta angustifolia* leaves. *Fitoterapia* **99**, 276–283 (2014).
- Markland, W., McQuaid, T., Jain, J. & Kwong, A. Broad-spectrum antiviral activity of the IMP dehydrogenase inhibitor VX-497: A comparison with ribavirin and demonstration of antiviral additivity with alpha interferon. *Antimicrob. Agents Chemother.* **44**, 859–866 (2000).
- Fernandez-Larsson, R. & Patterson, J. L. Ribavirin is an inhibitor of Human Immunodeficiency Virus reverse transcriptase. *Mol. Pharmacol.* **38**, 766–770 (1990).
- Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
- Pettersen, E. F. et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
- Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Chem. Inf. Model.* **31**, 455–461 (2010).
- Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng. Des. Sel.* **8**, 127–134 (1995).
- Khalidan, A. et al. 3D-QSAR modeling, molecular docking and ADMET properties of benzothiazole derivatives as α -glucosidase inhibitors. *Mater. Today Proc.* **45**, 7643–7652 (2021).
- Nagar, P. R., Gajjar, N. D. & Dhameliya, T. M. In search of SARS CoV-2 replication inhibitors: Virtual screening, molecular dynamics simulations and ADMET analysis. *J. Mol. Struct.* **1246**, 131190 (2021).
- Banerjee, P., Eckert, A. O., Schrey, A. K. & Preissner, R. ProTox-II: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* **46**, W257–W263 (2018).
- Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**, 6474–6487 (2001).
- Dodda, L. S., Cabeza de Vaca, I., Tirado-Rives, J. & Jorgensen, W. L. LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **45**, W331–W336 (2017).
- Hirshman, S. P. & Whitson, J. Steepest descent moment method for three-dimensional magnetohydrodynamic equilibria. (Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 1983).
- Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
- Van Gunsteren, W. F. & Berendsen, H. J. A leap-frog algorithm for stochastic dynamics. *Mol. Simul.* **1**, 173–185 (1988).
- Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461 (2015).
- Aarjane, M., Slassi, S., Ghaleb, A., Tazi, B. & Amine, A. Synthesis, biological evaluation, molecular docking and in silico ADMET screening studies of novel isoxazoline derivatives from acridone. *Arab. J. Chem.* **14**, 103057 (2021).
- Freedman, H., Winter, P., Tuszyński, J., Tyrrell, D. L. & Houghton, M. A computational approach for predicting off-target toxicity of antiviral ribonucleoside analogues to mitochondrial RNA polymerase. *J. Biol. Chem.* **293**, 9696–9705 (2018).
- Yu, H.-J., Xiao, G.-L., Zhao, Y.-Y., Wang, X.-X. & Lan, R. Targeting mitochondrial metabolism and RNA polymerase POLRMT to overcome multidrug resistance in cancer. *Front. Chem.* **9**, 775226 (2021).

47. Zarei, A., Ramazani, A., Rezaei, A. & Moradi, S. Screening of honey bee pollen constituents against COVID-19: An emerging hot spot in targeting SARS-CoV-2-ACE-2 interaction. *Nat. Prod. Res.* **37**, 974–980 (2023).
48. Amirkhani, R., Zarei, A., Gholampour, M., Tavakoli, H. & Ramazani, A. Honey and royal jelly natural products as possible antiviral nominations to combat SARS-CoV-2 main protease. *ECC*, 567–579, (2022).
49. Shamsian, S., Sokouti, B. & Dastmalchi, S. Benchmarking different docking protocols for predicting the binding poses of ligands complexed with cyclooxygenase enzymes and screening chemical libraries. **14**, 29955, (2024)
50. Elfiky, A. A. Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir against SARS-CoV-2 RNA dependent RNA polymerase (RdRp): A molecular docking study. *Life Sci.* **253**, 117592 (2020).
51. Ebrahimi, K. S. et al. In silico investigation on the inhibitory effect of fungal secondary metabolites on RNA dependent RNA polymerase of SARS-CoV-II: A docking and molecular dynamic simulation study. *Comput. Biol. Med.* **135**, 104613 (2021).
52. Mishra, C. B. et al. Identifying the natural polyphenol catechin as a multi-targeted agent against SARS-CoV-2 for the plausible therapy of COVID-19: an integrated computational approach. *Brief. Bioinform.* **22**, 1346–1360 (2021).

Acknowledgements

This work was supported by the Kermanshah University of Medical Sciences Research Council (Grant No. 4020771).

Author contributions

A.Z. Formal analysis, Investigation, Methodology, Project administration, Software, Resources, Visualization, Writing – original draft, and Writing – review & editing M.E. Formal analysis, Investigation, and Methodology M.T. M. Writing – review & editing G.S. Data curation, Validation, and Writing – review & editing H.A. Conceptualization, Project administration, Funding acquisition, Supervision, and Writing – review & editing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-44924-4>.

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

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2026