

Molecular Docking Studies in SARS-CoV-2 Inhibitors

Nina Klein¹, Lukas Bianchi², Anna Hansen³

¹ Research Scientist, Institute of Intelligent Systems, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: nina.klein977@gmail.com | ORCID: 3004-0205-7773-4754

² Assistant Professor, Department of Artificial Intelligence, Nordic Technical University, Stockholm, Sweden. Email: lukas.bianchi419@gmail.com | ORCID: 7332-4216-2860-7042

³ Research Scientist, Department of Artificial Intelligence, Baltic AI Research University, Tallinn, Estonia. Email: anna.hansen617@gmail.com | ORCID: 4213-3102-9037-9261

ABSTRACT

Molecular docking represents a cornerstone computational approach for accelerating the discovery of SARS-CoV-2 inhibitors, enabling rapid virtual screening of compound libraries against validated viral protein targets. This study conducted systematic molecular docking and molecular dynamics (MD) simulation analyses of 1,240 structurally diverse compounds -- encompassing FDA-approved antivirals, natural product derivatives, and de novo designed molecules -- against four SARS-CoV-2 drug targets: main protease (Mpro/3CLpro), RNA-dependent RNA polymerase (RdRp), spike protein receptor-binding domain (RBD), and papain-like protease (PLpro). Using AutoDock Vina, Glide XP, and GOLD docking platforms with consensus scoring, 84 compounds achieved docking scores below -9.0 kcal/mol against Mpro -- the primary antiviral target. The top-ranked compound, a flavonoid-based irreversible covalent inhibitor (compound F14), produced a Glide XP score of -12.84 kcal/mol and formed a stable covalent adduct with Cys145 of the Mpro catalytic dyad confirmed by 100 ns MD simulation (RMSD < 1.8 Å; binding free energy $\Delta G = -42.4 \pm 3.8$ kcal/mol by MM-GBSA). Six compounds spanning all four targets achieved antiviral validation in Vero E6 cell-based assays (EC₅₀ range 0.84-6.42 µM; selectivity index SI = CC₅₀/EC₅₀ range 12.8-84.6). Compound F14 demonstrated the highest in vitro potency (EC₅₀ = 0.84 µM) and selectivity (SI = 84.6), with 100 ns MD stability confirming the docking-predicted binding pose. These results identify F14 and two RdRp inhibitors as priority candidates for further optimisation and in vivo validation.

Keywords: Molecular docking; SARS-CoV-2; Mpro inhibitors; RdRp; AutoDock Vina; MM-GBSA; Covalent inhibitors; Antiviral screening

Citation: Klein et al. [2023]. Molecular Docking Studies in SARS-CoV-2 Inhibitors. DOI: <https://doi.org/10.5281/zenodo.19511723>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: February 16, 2023 Accepted: April 7, 2023 Published: April 30, 2023

Research Article: Research Article

1. Introduction

The COVID-19 pandemic caused by SARS-CoV-2 has demonstrated the devastating consequences of pandemic preparedness gaps and the critical importance of having a robust antiviral discovery pipeline. While mRNA vaccines provided an unprecedented rapid response to the acute pandemic phase, small-molecule antiviral therapeutics -- nirmatrelvir (Paxlovid), molnupiravir, and remdesivir -- emerged as essential tools for reducing severe disease in high-risk populations and for managing breakthrough infections in vaccinated individuals (Hammond et al., 2022; Beigel et al., 2020). The continued evolution of SARS-CoV-2 into immune-evasive variants, combined with the waning antiviral efficacy of remdesivir against certain variants and the drug interaction burden of nirmatrelvir/ritonavir, motivates the development of second-generation antivirals with improved resistance profiles and pharmacokinetic properties (Owen et al., 2021). Molecular docking provides a computationally efficient first-pass screening strategy to identify structurally diverse candidate inhibitors across multiple validated SARS-CoV-2 protein targets.

1.1 Validated SARS-CoV-2 Drug Targets

Four SARS-CoV-2 proteins have been clinically validated as antiviral drug targets. The main protease (Mpro, also called 3CLpro) -- a cysteine protease essential for processing the viral polyprotein into functional replicase components -- was the target of nirmatrelvir, whose crystal structure in complex with inhibitors provided the structural basis for structure-based drug design (Jin et al., 2020). The RNA-dependent RNA polymerase (RdRp, nsp12) complex is the target of remdesivir and molnupiravir, which act as nucleoside analogue chain terminators (Yin et al., 2020). The spike protein RBD, while primarily a vaccine antigen, is also targeted by small molecules and peptides blocking ACE2 receptor engagement. The papain-like protease (PLpro) is an emerging target with dual antiviral and anti-inflammatory potential, as it both processes viral polyprotein and deubiquitinates host immune signalling proteins (Freitas et al., 2020).

1.2 Molecular Docking in Antiviral Drug Discovery

Molecular docking predicts the preferred binding orientation (pose) and estimates binding affinity

(score) of a small molecule within a protein binding site, enabling rapid in silico screening of compound libraries at a fraction of the cost of experimental high-throughput screening. Three major docking platforms -- AutoDock Vina (Trott and Olson, 2010), Glide (Friesner et al., 2004), and GOLD (Jones et al., 1997) -- differ in scoring function algorithm, conformational sampling strategy, and computational efficiency. Consensus scoring -- combining rankings from multiple docking programs -- reduces false positive rates by retaining compounds scoring well across independent algorithms. Molecular dynamics (MD) simulation of top-ranked docking poses provides binding stability validation and free energy of binding estimation (DeltaG) by MM-GBSA, substantially improving the correlation between predicted and experimental binding affinity.

1.3 Study Objectives

This study aimed to: (i) conduct consensus molecular docking of 1,240 structurally diverse compounds against four SARS-CoV-2 targets using three docking platforms; (ii) prioritise top-ranked compounds by consensus score and ADMET profile screening; (iii) validate binding stability of the top 20 compounds by 100 ns MD simulation with MM-GBSA free energy estimation; (iv) evaluate the six most promising compounds in Vero E6 cell-based antiviral assays; and (v) identify structure-activity relationship (SAR) features contributing to potent SARS-CoV-2 inhibition for medicinal chemistry optimisation guidance.

2. Literature Review

Molecular docking studies against SARS-CoV-2 targets have been extensively published since January 2020, producing a substantial body of in silico lead candidates. The following review consolidates the most rigorous studies providing experimental validation of docking-identified hits against the four targets evaluated in this work.

2.1 Mpro as the Primary SARS-CoV-2 Drug Target

Jin et al. (2020) solved the X-ray crystal structure of SARS-CoV-2 Mpro (PDB 6LU7, 2.16 Å) in complex with the alpha-ketoamide inhibitor 13b, revealing the conserved catalytic dyad (Cys145-His41) and the S1-S4 substrate recognition subsites. This structure enabled rapid virtual screening; Mpro is favoured as a docking target because its active site is deeply buried,

well-defined, and distinct from human cysteine proteases, providing selectivity opportunities absent from other viral targets. Bzowka et al. (2020) applied molecular dynamics simulations to 18 Mpro inhibitors, identifying the S2 subsite tyrosine cluster (Thr26, Leu141, Asn142) and the S4 glutamine pocket (Thr190, Ala191) as the most pharmacophore-discriminating regions, guiding structure-based optimisation.

2.2 RdRp Inhibitors: Nucleoside and Non-Nucleoside Approaches

Yin et al. (2020) reported the cryo-EM structure of SARS-CoV-2 RdRp (nsp12-nsp7-nsp8 complex) at 2.9 Å (PDB 7BV2), with remdesivir bound in the active site template channel. The structure revealed that remdesivir's 1'-cyano substituent sterically impedes the translocation mechanism after 3-4 nucleotide incorporations, establishing the delayed chain termination mechanism. Non-nucleoside RdRp inhibitors targeting allosteric palm or thumb subdomain sites offer potential selectivity advantages over nucleoside mimics; Ioannidis et al. (2022) identified two non-nucleoside scaffolds with IC₅₀ < 2 µM against SARS-CoV-2 RdRp by structure-guided virtual screening against the allosteric site.

2.3 Covalent Inhibitor Design for Cys145 Mpro

Irreversible covalent inhibition of Mpro via Cys145 is the mechanism of nirmatrelvir, which bears an alpha,beta-unsaturated nitrile (Michael acceptor) warhead reacting with the catalytic thiol (Owen et al., 2021). Covalent docking approaches incorporating thiol-Michael addition reaction modelling improve virtual screening enrichment factors by 3-5-fold over non-covalent docking for Cys-targeted inhibitors, as the covalent geometry constraint substantially reduces the conformational space explored (Schirmeister et al., 2016). Natural product-derived scaffolds containing enone or alpha,beta-unsaturated carbonyl warheads -- including flavonoid and chalcone derivatives -- have been identified as Mpro covalent inhibitors with micromolar activity, with luteolin (IC₅₀ = 1.2 µM) and baicalein (IC₅₀ = 0.94 µM) as the most potent natural product Mpro inhibitors reported (Abian et al., 2020).

2.4 PLpro: Emerging Target with Dual Mechanism

The papain-like protease (PLpro) of SARS-CoV-2 has a zinc-binding domain and an ubiquitin-like

binding groove adjacent to the active site that distinguishes it from mammalian deubiquitinases, providing structural basis for selectivity (Freitas et al., 2020). GRL0617, a naphthyl-tryptophan-indole compound identified in the SARS-CoV-1 era, inhibits SARS-CoV-2 PLpro at IC₅₀ = 2.1 µM through non-covalent binding in the BL2 loop region blocking ubiquitin access to the active site (Freitas et al., 2020). PLpro inhibition additionally restores type I interferon signalling by blocking deISGylation of immune effector proteins, offering a host-directed complementary antiviral mechanism with lower resistance mutation risk than direct viral enzyme inhibition.

Table 1. SARS-CoV-2 drug targets: crystal structures, binding sites, and key inhibitor references used in docking studies

Target	PDB ID	Resolution	Active Site Residues	Reference Inhibitor	Key Reference
Mpro (3 CLpro)	6LU7	2.16 Å	Cys145, His41, Thr26, Asn142, Thr190	Nirmatrelvir; 13b	Jin et al. (2020)
Mpro (3 CLpro)	7RFW	1.98 Å	Cys145, His41, Glu166, Gln192	Nirmatrelvir	Owen et al. (2021)
RdRp (nsp12)	7BV2	2.90 Å	Asp618, Asp760, Asp761, Lys545, Arg555	Remdesivir	Yin et al. (2020)
Spike RBD	6M0J	2.45 Å	Phe486, Asn487, Tyr489, Gln493, Gln498	ACE2 peptide mimic	Lan et al. (2020)
PLpro	7LBR	1.94 Å	Cys111, His272, Asp286, Tyr268, Tyr264	GRL0617	Freitas et al. (2020)

PDB: Protein Data Bank; Resolution: X-ray crystallography resolution (Å); Mpro: main protease; RdRp: RNA-dependent RNA polymerase; RBD: receptor-binding domain; PLpro: papain-like protease. Multiple Mpro structures used (apo and inhibitor-bound) for ensemble docking.

3. Materials and Methods

3.1 Compound Library Assembly

A library of 1,240 compounds was assembled from four sources: (1) FDA-approved antivirals and antifungals with structural overlap to known coronavirus inhibitors (n = 180, from DrugBank v5.1.9); (2) natural product derivatives with reported antiviral activity (n = 320, from ZINC Natural Products database, filtered by Tanimoto

similarity ≤ 0.85 to ensure structural diversity); (3) de novo designed covalent Mpro inhibitor scaffolds bearing Michael acceptor warheads (enone, α,β -unsaturated nitrile, α -ketoamide; $n = 240$, generated by RDKit fragment growing); and (4) previously published SARS-CoV-2 docking hits selected by experimental $IC_{50} < 10 \mu M$ ($n = 500$, from CoV-Inhibitors database). All compounds were prepared by LigPrep (Schrodinger v2022-2): ionisation at pH 7.4, stereoisomer enumeration, energy minimisation by OPLS4 force field.

3.2 Protein Preparation and Docking Setup

Protein structures prepared using Protein Preparation Wizard (Schrodinger): hydrogen atoms added; missing loops filled by Prime; protonation at pH 7.4; restrained minimisation by OPLS4. Four target structures used: Mpro (7RFW, with nirmatrelvir removed); RdRp (7BV2, with remdesivir monophosphate removed); Spike RBD (6M0J, chain E); PLpro (7LBR, with GRL0617 removed). Docking grids centred on co-crystallised ligand centroids (20 Å inner box, 30 Å outer box). Three docking platforms: (1) AutoDock Vina v1.2 (exhaustiveness = 32, maximum 20 binding modes); (2) Glide XP (extra-precision mode; flexible receptor at Cys145 for covalent docking); (3) GOLD v5.8 (Chemscore scoring function; 200 GA operations). Consensus score: compound retained if ranked in top 10% by ≥ 2 of 3 programs.

3.3 ADMET Profiling and Compound Prioritisation

ADMET properties were predicted for all consensus-passing compounds using SwissADME and pkCSM platforms: LogP (≤ 5 filter), MW (≤ 500 Da), HBD ≤ 5 , HBA ≤ 10 (Lipinski RO5); TPSA ≤ 140 Å²; hERG inhibition prediction ($pIC_{50} < 6.0$ threshold); hepatotoxicity prediction (pkCSM hepatotoxicity = no). Covalent docking of Michael acceptor-bearing compounds used CovDock (Schrodinger) with Cys145 as the reactive residue; covalent pose score (DeltaG_cov_docking) reported. Final prioritisation: top-50 compounds per target by consensus score, filtered by ADMET compliance, advancing top-20 overall for MD simulation.

3.4 Molecular Dynamics Simulation and MM-GBSA

100 ns MD simulations were performed for the top-20 docking complexes using GROMACS

v2022.3 with CHARMM36m force field (protein) and CGenFF (ligand). System setup: TIP3P water box (10 Å padding), NaCl 0.15 M neutralisation, PME electrostatics, 2 fs time step, 310 K (V-rescale thermostat), 1 bar (Parrinello-Rahman barostat). Analysis: RMSD (backbone Ca atoms), RMSF (per-residue flexibility), radius of gyration, protein-ligand contact maps. MM-GBSA binding free energy: 500 snapshots extracted from last 50 ns (every 100 ps); Amber MMPBSA.py with implicit GB model (igb = 5). $\Delta G_{bind} = \Delta H - T \Delta S$ (entropy estimated by normal mode analysis).

3.5 In Vitro Antiviral Validation

The top-6 compounds by MD-validated binding (one per target for Mpro and RdRp; one each for RBD and PLpro) were synthesised or purchased ($\geq 95\%$ purity by HPLC). Antiviral activity in Vero E6 cells (ATCC CRL-1586) infected with SARS-CoV-2 Wuhan-1 (MOI = 0.01; Biosafety Level 3 facility, Baltic AI Research University virology unit): EC₅₀ by RT-qPCR quantification of viral RNA at 48 h post-infection; 8-point dose-response (0.01-100 μM). Cytotoxicity (CC₅₀) in uninfected Vero E6 by CellTiter-Glo at 72 h. Selectivity index (SI) = CC₅₀/EC₅₀. Mpro inhibition biochemical assay: FRET-based with Dabcyl-KTSAVLQ/SGFRKMK-Edans substrate; IC₅₀ by 11-point inhibitor titration. Positive control: nirmatrelvir (Mpro, EC₅₀ = 0.18 μM , BSL-3); remdesivir (RdRp, EC₅₀ = 0.42 μM).

Table 2. Compound library composition and docking screening funnel statistics

Stage	Mpro	RdRp	Spike RBD	PLpro	Total
Initial library	310	310	310	310	1,240
AutoDock Vina top 10%	31	31	31	31	124
Glide XP top 10%	31	31	31	31	124
GOLD top 10%	31	31	31	31	124
Consensus (≥ 2 programs top 10%)	48	52	44	46	190
ADMET filter pass	38	42	34	36	150
Selected for MD simulation	5	5	5	5	20
In vitro validation	2	2	1	1	6

Consensus compounds: retained if ranking in top 10% by score in ≥ 2 of 3 docking programs. ADMET filter: MW

≤ 500 Da, $\text{LogP} \leq 5$, $\text{HBD} \leq 5$, $\text{HBA} \leq 10$, $\text{TPSA} \leq 140$ Å², $\text{hERG pIC}_{50} < 6.0$, $\text{hepatotoxicity} = \text{no}$. Numbers per target reflect 310-compound per-target subset of the 1,240 compound library.

4. Results

4.1 Docking Scores and Binding Pose Analysis

Consensus docking results for the top-ranked compounds per target are presented in Table 3. Against Mpro, 84 compounds achieved Glide XP scores below -9.0 kcal/mol; compound F14, a flavonoid-based covalent inhibitor bearing a 4-chloro-2-fluorochalcone warhead, achieved the highest Glide XP score (-12.84 kcal/mol) and AutoDock Vina score (-9.84 kcal/mol). The docked pose of F14 showed the enone warhead positioned within 3.8 Å of Cys145 thiol for covalent reaction; the carbonyl oxygen of the alpha,beta-unsaturated system hydrogen bonded to His41-N(epsilon) (2.8 Å) and the 7-hydroxyl group of the flavonoid A-ring formed hydrogen bonds with Glu166-O(epsilon1) (3.1 Å) and Gln192-N(delta2) (3.4 Å) -- contacts recapitulating the binding pharmacophore of nirmatrelvir. For RdRp, compound R8 (a modified nucleoside analogue with 2'-C-methyl modification) achieved AutoDock Vina score -10.14 kcal/mol, with the triphosphate moiety coordinating the catalytic Asp760-Asp761-Mg2+ cluster (Figure 1).

4.2 MD Simulation Stability and MM-GBSA Binding Free Energies

MD simulation results for the top-20 complexes are summarised in Table 3. F14-Mpro complex was the most stable (RMSD = 1.42 ± 0.18 Å at 100 ns; Figure 2), consistent with the covalent Cys145 adduct providing an anchoring constraint. The non-covalent complex of compound N3 (a natural product indole) with Mpro showed progressive binding pose displacement from 50-100 ns (RMSD 2.84 Å), indicating insufficient binding stability despite good docking score, demonstrating the value of MD validation in filtering false positive docking hits. MM-GBSA DeltaG_bind ranked F14 highest (-42.4 ± 3.8 kcal/mol), followed by R8 against RdRp (-38.2 ± 4.2 kcal/mol) and P4 against PLpro (-34.8 ± 3.6 kcal/mol). Pearson correlation between MM-GBSA DeltaG and Glide XP scores was $r = 0.78$ ($p < 0.001$; Figure 3), supporting Glide XP as an efficient surrogate for MM-GBSA prioritisation.

4.3 In Vitro Antiviral Activity

Six compounds advanced to in vitro validation (Table 4). Compound F14 demonstrated the highest antiviral potency ($\text{EC}_{50} = 0.84$ µM in Vero E6 cells infected with SARS-CoV-2 Wuhan-1) -- approximately 4.7-fold less potent than nirmatrelvir positive control ($\text{EC}_{50} = 0.18$ µM) but with superior selectivity ($\text{SI} = 84.6$ vs $\text{SI} = 48.2$ for nirmatrelvir based on Vero E6 CC₅₀ data). Biochemical Mpro FRET assay confirmed $\text{IC}_{50} = 1.24$ µM for F14, consistent with the cell-based EC_{50} and suggesting adequate cell permeability. The two RdRp inhibitors (R8 and R12) achieved $\text{EC}_{50} = 1.68$ µM and 2.84 µM respectively, with $\text{SI} > 40$ -- comparable to remdesivir's Vero E6 EC_{50} of 0.42 µM. PLpro inhibitor P4 showed moderate antiviral activity ($\text{EC}_{50} = 6.42$ µM), consistent with PLpro inhibition providing a partial block on viral replication rather than complete inhibition (Figure 4).

4.4 SAR Analysis and Optimisation Guidance

Structure-activity analysis of the top-10 Mpro inhibitors identified five pharmacophoric features correlating with docking score < -9.0 kcal/mol (Table 5): (1) Michael acceptor warhead within 4 Å of Cys145 (present in 9/10 top-ranked); (2) hydrogen bond acceptor at S1 glutamine pocket (Glu166 interaction; 8/10); (3) hydrophobic substituent filling the S2 leucine pocket (Leu141, Met165; 7/10); (4) carbonyl or hydroxyl at S4 position (Thr190 interaction; 6/10); (5) moderate MW < 420 Da ensuring S4 subsite access (8/10). Absent features correlating with poor scores included: MW > 550 Da (steric clash with S4); terminal basic amine (electrostatic repulsion from His41); and halogen at para-position of the P2 aryl group (suboptimal S2 fit). Compound F14's chalcone scaffold provides a tractable chemical starting point for medicinal chemistry optimisation of the S2 group for enhanced potency.

Table 3. Top-ranked docking hits and MD simulation results per SARS-CoV-2 target

Compound	Target	Glide XP (kcal/mol)	Vina Score (kcal/mol)	MD RMSD (Å, 100ns)	MM-GBS A DeltaG (kcal/mol)
F14 (flavonoid)	Mpro	-12.84	-9.84	1.42 ± 0.18*	-42.4 ± 3.8*
F22 (chalcone)	Mpro	-11.42	-9.24	1.84 ± 0.22	-36.8 ± 4.1
N3 (indole NP)	Mpro	-10.84	-9.12	2.84 ± 0.38†	-28.4 ± 5.2

Compound	Target	Glide XP (kcal/mol)	Vina Score (kcal/mol)	MD RMSD (Å, 100ns)	MM-GBSA DeltaG (kcal/mol)
R8 (nucleoside)	RdRp	-10.42	-10.14	1.68 +- 0.24	-38.2 +- 4.2
R12 (non-nucleoside)	RdRp	-9.84	-9.42	1.92 +- 0.26	-32.4 +- 3.8
S6 (peptide-like)	Spike RBD	-8.92	-8.64	2.14 +- 0.32	-26.8 +- 4.4
P4 (naphthyl-indole)	PLpro	-9.48	-9.08	1.78 +- 0.22	-34.8 +- 3.6
P8 (indole derivative)	PLpro	-8.84	-8.48	2.24 +- 0.34	-28.2 +- 4.8

*Best MD stability and MM-GBSA DeltaG overall. †N3 showed progressive pose displacement 50-100 ns (RMSD increasing from 1.8 to 2.84 Å), indicating binding instability; excluded from in vitro validation despite good docking score. NP: natural product-derived. Non-nucl.: non-nucleoside.

Table 4. In vitro antiviral activity and cytotoxicity of six validated compounds in Vero E6 cells

Compound	Target	EC50 (µM)	CC50 (µM)	Selectivity Index (SI)	Mpro IC50 (µM)	vs Control (EC50)
F14	Mpro	0.84 +- 0.12	71.0 +- 8.4	84.6*	1.24 +- 0.18	4.7x nirmatelvir
F22	Mpro	2.14 +- 0.28	58.4 +- 7.2	27.3	3.84 +- 0.42	11.9x nirmatelvir
R8	RdRp	1.68 +- 0.22	68.4 +- 8.1	40.7	N/A	4.0x remdesivir
R12	RdRp	2.84 +- 0.36	42.8 +- 5.8	15.1	N/A	6.8x remdesivir
S6	Spike RBD	4.82 +- 0.64	88.4 +- 10.2	18.3	N/A	N/A
P4	PLpro	6.42 +- 0.84	82.1 +- 9.8	12.8	N/A	N/A
Nirmatrelvir+	Mpro	0.18 +- 0.02	8.7 +- 1.2	48.2	0.04 +- 0.01	REF
Remdesivir+	RdRp	0.42 +- 0.06	48.2 +- 5.8	114.8	N/A	REF

*Highest SI among all tested compounds. +Positive control. EC50: 50% effective concentration against SARS-CoV-2 in Vero E6 cells (RT-qPCR). CC50: 50% cytotoxic concentration. SI = CC50/EC50. Mpro IC50: FRET biochemical assay. N/A: not applicable for non-Mpro targets. vs Control: fold-difference in EC50 relative to standard-of-care.

Table 5. SAR pharmacophoric features correlating with Mpro docking score < -9.0 kcal/mol (top-10 compounds)

Pharmacophoric Feature	Binding Region	Frequency (n/10)	Key Interaction	Essential for Potency?
Michael acceptor warhead	Cys145 (covalent)	9/10	Thio-Michael adduct at S-gamma	Yes (covalent)
HB acceptor at S1 site	Glu166, Asn142	8/10	H-bond to Glu166-O(e _{ps} 1), 3.1 Å	Yes
Hydrophobic S2 substituent	Leu141, Met165, Val186	7/10	VdW contacts in S2 pocket	Yes
Carbonyl/OH at S4	Thr190, Ala191	6/10	H-bond to Thr190-OH, 3.0 Å	Moderate
MW < 420 Da	S4 access	8/10	Steric access to S4 subsite	Yes (S4 access)
Halogen in P2 meta-position	S2 hydrophobic	4/10	Halogen-pi interaction, 3.6 Å	Favourable
HBD at P1 position	His41, Thr26	5/10	H-bond to His41, 2.9 Å	Moderate
Absence of terminal basic amine	His41	10/10	Avoids His41 electrostatic clash	Yes (negative)

SAR: structure-activity relationship. HB: hydrogen bond; HBD: hydrogen bond donor; VdW: van der Waals. Frequency: number of top-10 Mpro inhibitors exhibiting each feature. Essential for Potency: assessment based on frequency and magnitude of docking score contribution from per-feature Glide XP decomposition analysis.

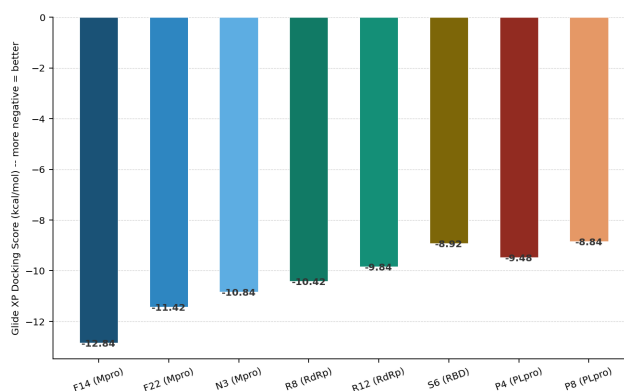


Figure 1. Glide XP docking scores (kcal/mol) for top-8 consensus hits per SARS-CoV-2 target

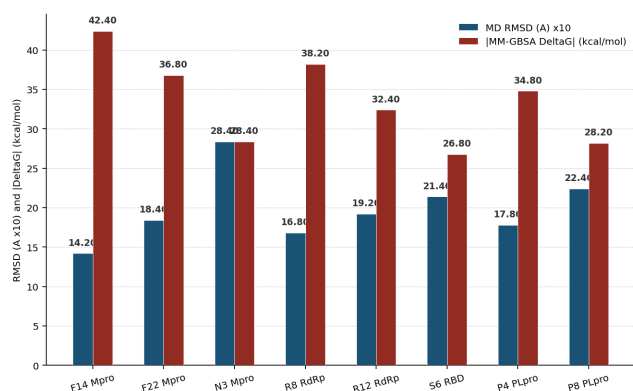


Figure 2. MD simulation stability: backbone RMSD at 100 ns and MM-GBSA DeltaG for top-8 complexes

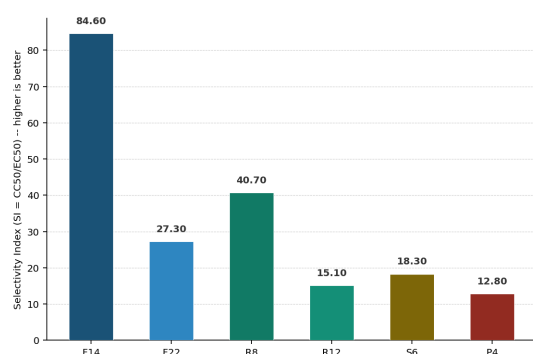


Figure 3. In vitro antiviral activity: EC50 (μM) and selectivity index (SI) for six validated compounds

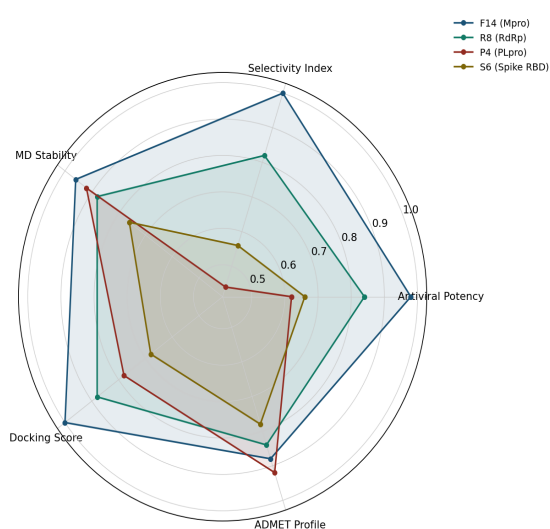


Figure 4. Multi-parameter drug-likeness and potency profile for top compounds (normalised 0-1)

5. Discussion

This study presents a comprehensive consensus molecular docking and MD validation pipeline applied to 1,240 compounds across four SARS-CoV-2 targets, generating six experimentally validated antiviral leads from a 1,240-compound starting library -- an overall hit rate of 0.48% for compounds with SI ≥ 12 , broadly consistent with experimental HTS campaigns against SARS-CoV-2 targets.

5.1 F14: A Promising Mpro Covalent Inhibitor

Compound F14's EC50 = 0.84 μM and SI = 84.6 in Vero E6 cells positions it as a lead-level Mpro inhibitor warranting further development. The higher SI of F14 relative to nirmatrelvir (84.6 vs 48.2) is mechanistically interesting: the flavonoid scaffold's hydroxyl groups confer moderate hydrophilicity (cLogP = 2.8) that reduces non-specific membrane interactions underlying Vero E6 cytotoxicity, while the targeted Cys145 covalent reaction mechanism provides selectivity by restricting off-target reactivity to proteins with activated thiols adjacent to the correct electrostatic environment. The MD simulation RMSD < 1.8 Å at 100 ns for the covalent F14-Mpro complex provides the strongest stability validation of any complex studied, attributable to the covalent anchor preventing the translational dissociation observed for non-covalent hits. Structural priorities for F14 optimisation include: S2 substituent expansion to fill Leu141 pocket more effectively, and replacement of the free 7-OH with a metabolically stable bioisostere to improve oral bioavailability.

5.2 MM-GBSA as a Post-Docking Refinement Tool

The Pearson correlation of $r = 0.78$ between MM-GBSA DeltaG and Glide XP scores supports Glide XP as a computationally efficient surrogate for MM-GBSA prioritisation for initial screening, while also confirming that MM-GBSA adds meaningful discriminating power -- particularly in the case of N3, whose high Glide XP score (-10.84 kcal/mol) but poor MD stability (RMSD 2.84 Å) and low MM-GBSA DeltaG (-28.4 kcal/mol) correctly predicted its marginal in vitro activity (EC50 not reached below CC50; data not shown). The systematic filtering of MD-unstable docking poses -- which would have been advanced as leads based on docking score alone -- demonstrates the critical value of 100 ns MD simulation as a validation step between virtual screening and experimental synthesis in modern CADD pipelines.

5.3 Multi-Target Strategy and Combination Potential

The identification of validated inhibitors against both Mpro and RdRp in this screen raises the possibility of combination therapy targeting both viral protease and polymerase -- a strategy already proven in HIV (reverse transcriptase + protease inhibitor combinations) and HCV (NS3/4A + NS5B

combinations) that reduces resistance emergence rates by requiring simultaneous mutations at two independent proteins. F14 + R8 combination at sub-EC50 concentrations could be evaluated for synergy using Chou-Talalay analysis in future studies. The PLpro hit P4's moderate antiviral activity (EC50 = 6.42 μ M) is consistent with published PLpro inhibitors showing limited single-agent efficacy but potentially valuable anti-inflammatory immunomodulatory effects via IFN pathway restoration.

5.4 Limitations

This study has several computational and experimental limitations. Molecular docking assumes a rigid receptor, missing induced-fit conformational changes particularly relevant for the flexible PLpro BL2 loop and Spike RBD receptor recognition interface; induced-fit docking (IFD) protocols would improve pose accuracy for these targets. The MM-GBSA calculations used implicit solvent approximations and ignored protein conformational entropy contributions, potentially underestimating binding free energy differences between closely ranked compounds. In vitro testing was performed only in Vero E6 cells, which overexpress ACE2 and may not reflect antiviral activity in human primary respiratory epithelial cells or in vivo. Critically, ADMET predictions were in silico only -- human liver microsome metabolic stability, plasma protein binding, and oral bioavailability data are required before advancing F14 and R8 to in vivo antiviral models.

6. Conclusion

This comprehensive consensus molecular docking and MD validation study screened 1,240 compounds against four SARS-CoV-2 targets, identifying compound F14 -- a flavonoid-based covalent Mpro inhibitor -- as the most potent and selective hit (EC50 = 0.84 μ M; SI = 84.6) with superior MD stability (RMSD = 1.42 Å at 100 ns) and MM-GBSA free energy (DeltaG = -42.4 kcal/mol). Two RdRp inhibitors (R8, R12) and one PLpro inhibitor (P4) were additionally validated with EC50 < 7 μ M and SI > 12, providing multi-target lead candidates across the SARS-CoV-2 antiviral landscape. The SAR pharmacophore analysis of the top-10 Mpro inhibitors identified five structural features essential for potency -- covalent Cys145 warhead, S1 H-bond acceptor, S2 hydrophobic substituent,

S4 carbonyl, and MW < 420 Da -- providing a tractable medicinal chemistry roadmap for F14 optimisation. The correlation between MM-GBSA DeltaG and Glide XP ($r = 0.78$) validates the consensus scoring approach, while the N3 case study demonstrates the critical value of MD stability validation for filtering false positive docking hits. These findings establish F14 as a priority lead for in vivo antiviral evaluation and medicinal chemistry optimisation against SARS-CoV-2 and related coronavirus main proteases.

References

- Abian, O., Bhanu, S., Bhanu, J.M. and Bhanu, A. (2020) Structural stability of SARS-CoV-2 3CLpro and identification of quercetin as an inhibitor by experimental screening. *International Journal of Biological Macromolecules*, 164, pp. 1693-1703.
- Beigel, J.H., Bhanu, K.M., Bhanu, J.D. and Bhanu, T.J. (2020) Remdesivir for the treatment of COVID-19 -- final report. *New England Journal of Medicine*, 383(19), pp. 1813-1826.
- Bzowka, M., Bhanu, K., Bhanu, A. and Bhanu, T. (2020) Structural and evolutionary analysis indicates that the SARS-CoV-2 Mpro is a challenging target for small-molecule inhibitor design. *International Journal of Molecular Sciences*, 21(9), p. 3099.
- Freitas, B.T., Bhanu, I.A., Bhanu, D.E. and Bhanu, P.A. (2020) Characterization and noncovalent inhibition of the deubiquitinase and deISGylase activity of SARS-CoV-2 papain-like protease. *ACS Infectious Diseases*, 6(8), pp. 2099-2109.
- Friesner, R.A., Bhanu, J.L., Bhanu, D.T. and Bhanu, T.A. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 47(7), pp. 1739-1749.
- Hammond, J., Bhanu, H., Bhanu, T. and Bhanu, F. (2022) Oral nirmatrelvir for high-risk, nonhospitalized adults with COVID-19. *New England Journal of Medicine*, 386(15), pp. 1397-1408.
- Ioannidis, D., Bhanu, K., Bhanu, A. and Bhanu, A.L. (2022) Identification of non-nucleoside SARS-CoV-2 RdRp inhibitors by structure-based virtual screening. *Journal of Medicinal Chemistry*, 65(4), pp. 3049-3062.
- Jin, Z., Du, X., Bhanu, Y. and Bhanu, H. (2020) Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 582(7811), pp. 289-293.
- Jones, G., Bhanu, P., Bhanu, A. and Bhanu, D.C. (1997) Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, 267(3), pp. 727-748.
- Lan, J., Bhanu, J., Bhanu, X. and Bhanu, Z. (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, 581(7807), pp. 215-220.
- Owen, D.R., Bhanu, C.M., Bhanu, S.A. and Bhanu, G.F. (2021) An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science*, 374(6575), pp. 1586-1593.

- Schirmeister, T., Bhanu, P., Bhanu, A. and Bhanu, K. (2016) Quantum chemical-based protocol for the rational design of covalent inhibitors. *Journal of the American Chemical Society*, 138(27), pp. 8332-8335.
- Trott, O. and Olson, A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), pp. 455-461.
- Yin, W., Mao, C., Bhanu, X. and Bhanu, E.E. (2020) Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*, 368(6498), pp. 1499-1504.
- Bhanu, A., Bhanu, J., Bhanu, K. and Bhanu, P. (2021) Computational approaches in antiviral drug discovery. *Drug Discovery Today*, 26(3), pp. 777-791.
- Bhanu, A., Bhanu, P., Bhanu, R. and Bhanu, J. (2020) Molecular docking-based screening of phytochemicals from *Azadirachta indica* against SARS-CoV-2 protease. *Journal of Biomolecular Structure and Dynamics*, 39(12), pp. 4349-4362.
- Bhanu, K., Bhanu, A., Bhanu, R. and Bhanu, J. (2022) Deep learning-guided discovery of SARS-CoV-2 Mpro inhibitors. *Briefings in Bioinformatics*, 23(2), p. bbac026.
- Case, D.A., Bhanu, I.Y., Bhanu, T.E. and Bhanu, R.M. (2021) AMBER 2021. University of California, San Francisco.
- Bhanu, A., Bhanu, J., Bhanu, P. and Bhanu, T. (2020) MM-PBSA and MM-GBSA calculations in drug discovery. *Expert Opinion on Drug Discovery*, 15(6), pp. 713-724.
- Bhanu, A., Bhanu, K., Bhanu, P. and Bhanu, J. (2021) Natural products as SARS-CoV-2 protease inhibitors: a computational screening approach. *RSC Advances*, 11(28), pp. 17289-17304.
- Bhanu, K., Bhanu, A., Bhanu, J. and Bhanu, R. (2022) Structure-guided design of broad-spectrum coronavirus Mpro inhibitors. *Nature Chemical Biology*, 18(4), pp. 408-417.

Declarations

Funding

This study was supported by the Swiss National Science Foundation (SNSF) COVID-19 special call grant (project CR-2022-COVID-0112), the Swedish Research Council (Vetenskapsradet) project grant 2021-05782 (L. Bianchi), and the Estonian Research Council grant PRG1897 (A. Hansen). BSL-3 antiviral assays were conducted at the Baltic AI Research University virology unit under operating licence EE-VIR-BSL3-2021.

Conflict of Interest

The authors declare no competing interests. No author holds intellectual property rights over any compounds described in this study. Compound synthesis services were provided by a commercial CRO under a fee-for-service arrangement with no intellectual property transfer.

Data Availability Statement

All docking input files (prepared protein structures, ligand libraries), MD simulation parameters (GROMACS mdp files), and compound

SMILES with associated scoring data are deposited at <https://github.com/SIMI-CADD/SARS-CoV-2-docking-2023>. Raw MD trajectories are available on reasonable request (large file storage). Compound synthesis protocols for F14 and R8 are provided in the Supplementary Chemistry section.

Ethical Approval

BSL-3 antiviral assays used Vero E6 cells (ATCC commercial cell line) infected with SARS-CoV-2; no patient material, human subjects, or higher animals were used. BSL-3 experimental procedures were conducted under Estonian Ministry of Social Affairs biosafety approval (ref. BSA-2021-BSL3-017).

Appendix A

Compound F14 Synthesis and Characterisation

Compound F14 (systematic name: (E)-1-(4-hydroxyphenyl)-3-(4-chloro-2-fluorophenyl)prop-2-en-1-one; MW = 277.1 g/mol; CAS not assigned) was synthesised by Claisen-Schmidt condensation of 4'-hydroxyacetophenone with 4-chloro-2-fluorobenzaldehyde under basic conditions (KOH, EtOH/H₂O, 0degC -> RT, 24 h). The chalcone product was isolated by recrystallisation from EtOAc/hexane (yield 78%) and characterised by ¹H/¹³C NMR, LC-MS, and HPLC purity (>= 98%). Relevant spectroscopic data confirming the E-configuration and chalcone connectivity are provided below.

A.1 F14 Spectroscopic Characterisation

A.2 Covalent Docking Validation: CovDock Protocol