



Rational identification of small molecules derived from 9,10-dihydrophenanthrene as potential inhibitors of 3CL^{pro} enzyme for COVID-19 therapy: a computer-aided drug design approach

Ossama Daoui¹ · Souad Elkhattabi¹ · Samir Chtita²

Received: 8 May 2022 / Accepted: 23 June 2022 / Published online: 7 July 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Small molecules such as 9,10-dihydrophenanthrene derivatives have remarkable activity toward inhibition of SARS-CoV-2 3CL^{pro} and COVID-19 proliferation, which show a strong correlation between their structures and bioactivity. Therefore, these small compounds could be suitable for clinical pharmaceutical use against COVID-19. The objective of this study was to remodel the structures of 9,10-dihydrophenanthrene derivatives to achieve a powerful biological activity against 3CL^{pro} and favorable pharmacokinetic properties for drug design and discovery. Therefore, by the use of bioinformatics techniques, we developed robust 3D-QSAR models that are capable of describing the structure–activity relationship for 46 molecules based on 9,10-dihydrophenanthrene derivatives using CoMFA/SE ($R^2=0.97$, $Q^2=0.81$, $R^2_{\text{pred}}=0.95$, $^cR^2_p=0.71$) and CoMSIA/SEHDA ($R^2=0.94$, $Q^2=0.76$, $R^2_{\text{pred}}=0.91$, $^cR^2_p=0.65$) techniques. Accordingly, 96 lead compounds were generated based on a template molecule that showed the highest observed activity in vitro (T40, $\text{pIC}_{50}=5.81$) and predicted their activities and bioavailability in silico. The rational screening outputs of 3D-QSAR, Molecular docking, ADMET, and MM-GBSA led to the identification of 9 novel modeled molecules as potent noncovalent drugs against SARS-CoV-2-3CL^{pro}. Finally, by molecular dynamics simulations, the stability and structural dynamics of 3CL^{pro} free and complex (PDB code: 6LU7) were discussed in the presence of samples of 9,10-dihydrophenanthrene derivative in an aqueous environment. Overall, the retrosynthesis of the proposed drug compounds in this study and the evaluation of their bioactivity in vitro and in vivo may be interesting for designing and discovering a new drug effective against COVID-19.

Keywords SARS-CoV-2-3CL^{pro} · Non-covalent inhibitors · 9,10-dihydrophenanthrene · Potential drug

Introduction

The world is experiencing an unstable situation at all poles due to the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), coronavirus disease 2019 (COVID-19), which

has spread throughout the world and affected the majority of the community [1, 2]. Simultaneously, various strict precautionary measures have been taken to prevent the transmission and rapid spread of COVID-19 among people. Likewise, many efforts have been made worldwide to develop an effective drug to treat patients who are potentially infected with COVID-19, leading to the development of various vaccines that have been widely applied extensively to the public [3, 4]. However, SARS-CoV-2 is known for viral progression leading to the production of new mutants cloned from it. For example, alpha variant (B.1.1.7), beta variant (B.1.351), gamma variant (P.1), delta variant (B.1.617.2) [5, 6]. Moreover, there are several other mutations of this virus that have been monitored around the world. In this regard, several reports published by the World Health Organization have pointed out that the known SARS-CoV-2 mutations have little effect on the characteristics of COVID-19 represented in the propagation speed and the level of seriousness [7, 8]. Although

✉ Ossama Daoui
ossama.daoui@usmba.ac.ma

✉ Souad Elkhattabi
souad.elkhattabi@usmba.ac.ma

✉ Samir Chtita
samirchtita@gmail.com

¹ Laboratory of Engineering, Systems and Applications, National School of Applied Sciences, Sidi Mohamed Ben Abdellah-Fez University, BP Box 72, Fez, Morocco

² Laboratory of Analytical and Molecular Chemistry, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, B.P 7955 Casablanca, Morocco

the vaccine proved successful against COVID-19 mutations, people who have been vaccinated are still at risk of reinfection with this virus [9]. There is a great need to find effective antiviral drugs related to coronavirus, such as COVID-19 as a model. Meanwhile, researchers are eager to discover and design suitable drugs or vaccine solutions that can be used successfully against diseases related to SARS-CoV-2 and its variants [10]. This is done by targeting different protein pathways such as coronavirus 3C-like protease (3CL^{pro}) [11–13], papain-like protease (PL^{pro}) [14], RNA-dependent RNA polymerase (RdRp) [15], viral spike glycoprotein (S protein) [16], transmembrane protease serine 2 (TMPRSS2) [17], angiotensin-converting enzyme 2 (ACE2) [18], angiotensin AT2 receptor [19], etc. In this work, we will focus our research efforts on identifying novel agents that can be able to inhibit and eliminate coronavirus life cycle processes by targeting the enzymatic activity of 3-chymotrypsin-like cysteine protease (3CL^{pro}) [20]. This is due to the enzymatic environment of 3CL^{pro} is highly suitable for coronavirus regeneration, it can cleave viral proteins into distinct functional fractions [21]. The critical importance of the 3CL^{pro} enzyme in the coronavirus life cycle makes it an ideal therapeutic target for the development of antiviral drugs with optimal activity against coronavirus [22]. The successful inhibition of 3CL^{pro} enzymatic activity may reduce the drug resistance potential associated with COVID-19 variants [23]. To date, there are two classes of 3CL^{pro} inhibitors for SARS-CoV-2, covalent and non-covalent inhibitors [24]. Their pattern of inhibition varies based on interaction mechanisms between the drug ligands and active amino acid residue sites in the 3CL^{pro} receptor pocket [25, 26]. Among many covalent 3CL^{pro} inhibitors, only PF-07321332 [27] and PF-00835231 [28] have reached the clinical trial stage [29–32]. On the other hand, several small molecules have been reported as non-covalent inhibitors of 3CL^{pro} such as sciadopitysin [33], 23R [34], CCF981 [35], and pyridine [36, 37]. Despite all these efforts, most of the proposed covalent and non-covalent SARS-CoV-2-3CL^{pro} inhibitors have drug-like properties and pharmacokinetics that are not compelling for appropriate drug use; the use of covalent peptide inhibitors exacerbates the adverse side effects of clinical use of these drugs [38, 39]. Therefore, it has become necessary to explore non-covalent 3CL^{pro} inhibitors with favorable drug properties, good pharmacokinetics, and high inhibitory activity against COVID-19. In this context, a recent study conducted by Zhang et al. the *in vitro* discovery of several derivatives of 9,10-dihydrophenanthrene was reported as potential non-covalent, non-peptide inhibitors of the SARS-CoV-2-3CL^{pro} [2]. Based on *in vitro* results on the suitability of 9,10-dihydrophenanthrene derivatives for the inhibition of 3CL^{pro}, in this work, we perform a large-scale computational study aided by bioinformatics techniques to validate *in vitro* results and also to rationalize the design of new non-covalent inhibitors of 3CL^{pro} as more convenient drugs against COVID-19. For this purpose, this study includes the following principal objectives: the study of the quantitative structure–activity relationship for

9,10-dihydrophenanthrene derivatives based on 3D-QSAR techniques [8, 38], improvement of the biological inhibitory activity of these molecules based on the obtained pharmacological hypotheses, screening the most active molecular structures that are able to inhibit SARS-CoV-2 3CL^{pro}, predicting the drug-like and drug-pharmacokinetic ADMET profiles of the candidate drug molecules [41, 42], exploring the potential interaction patterns between active amino acids in the protein pocket of 3CL^{pro} (PDB code: 6LU7) (Fig. 1) [29] and the drug molecules using molecular docking and MM-GBSA techniques, and performing molecular dynamics simulations to check the stability of candidate SARS-CoV-2 3CL^{pro} inhibitors in the aquatic environment.

The rest of this work is organized as follows: the second section presents the materials and methods. The third section includes the simulation results and discussions. The last section concludes the performed work.

Material and methods

Investigated database

Through a previous study conducted by Zhang et al. [2], we collected 46 small molecules derived from 9,10-dihydrophenanthrene as non-covalent inhibitors of SARS-CoV-2 3CL^{pro}. Table 1 presents the structures of all 46 compounds as well as their corresponding half-maximal inhibitory concentration ($pIC_{50} = -\log IC_{50}$) data in the presence of disulfiram (DSF) drug substance used as a reference in the *in vitro* positive control [43, 44].

3D quantitative structure–activity relationship (3D-QSAR) analysis

Preparation of the database

Using SYBYL-X 2.1.1 software [45], we sketched the 3D structures of the 9,10-dihydrophenanthrene derivatives,

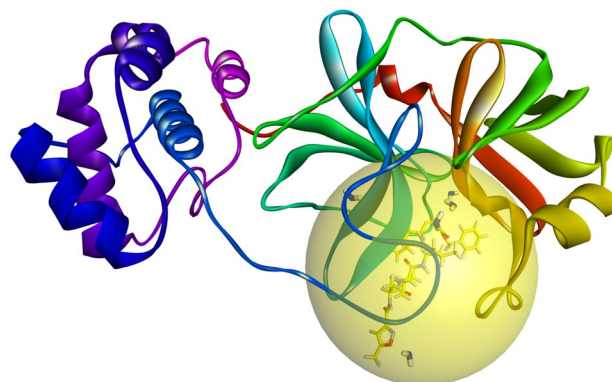
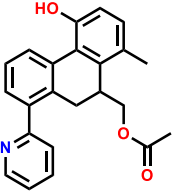
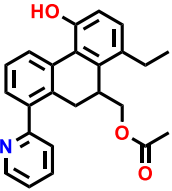
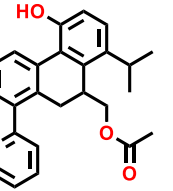
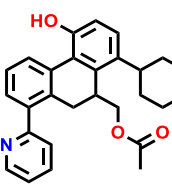
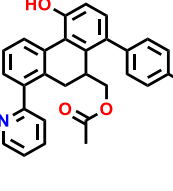
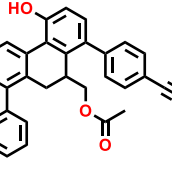
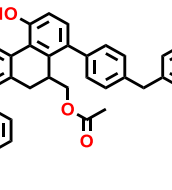
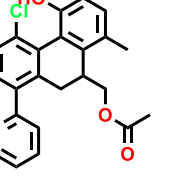
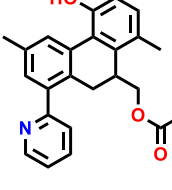
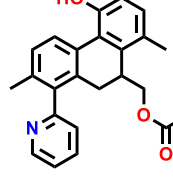
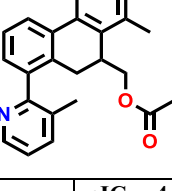
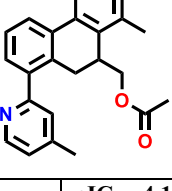
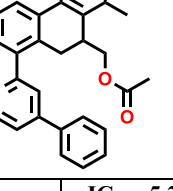
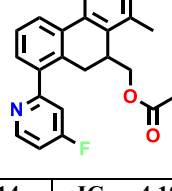
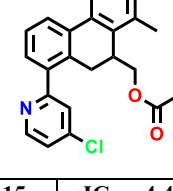
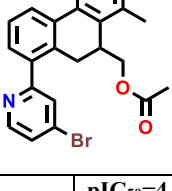
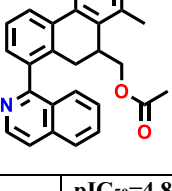
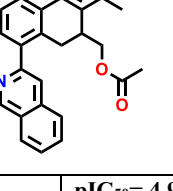
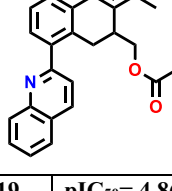
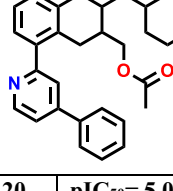
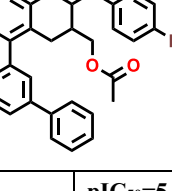
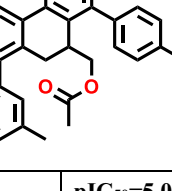
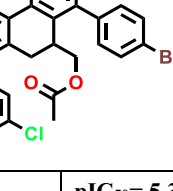
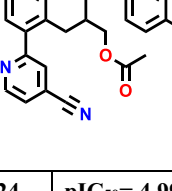
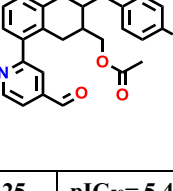
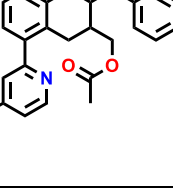
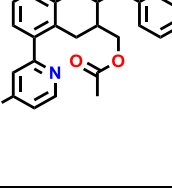
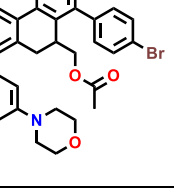
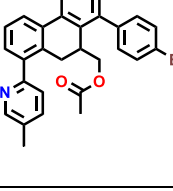
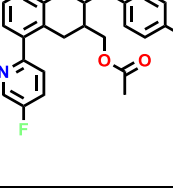


Fig. 1 Model crystal structure of SARS-CoV2-3CL^{pro} enzyme complexed with inhibitor N3 (PDB code: 6LU7)

Table 1 In vitro inhibitory activities of 9,10-dihydrophenanthrene derivatives against SARS-CoV-2 3CL^{pro}

				
01 pIC ₅₀ = 4.21	02 pIC ₅₀ = 4.48	03 pIC ₅₀ = 4.53	04 pIC ₅₀ = 5.04	05 pIC ₅₀ = 5.19
				
06 pIC ₅₀ = 4.71	07 pIC ₅₀ = 4.94	08 pIC ₅₀ = 4.07	09 pIC ₅₀ = 4.24	10 pIC ₅₀ = 4.27
				
11 pIC ₅₀ =4.23	12 pIC ₅₀ =4.17	13 pIC ₅₀ = 5.25	14 pIC ₅₀ = 4.18	15 pIC ₅₀ = 4.45
				
16 pIC ₅₀ =4.73	17 pIC ₅₀ =4.85	18 pIC ₅₀ = 4.92	19 pIC ₅₀ = 4.86	20 pIC ₅₀ = 5.07
				
21 pIC ₅₀ =5.61	22 pIC ₅₀ =5.01	23 pIC ₅₀ = 5.32	24 pIC ₅₀ = 4.99	25 pIC ₅₀ = 5.48
				

optimizing their energy structure to the minimum via the standard Tripos Powell force field (Gasteiger-Huckel partial atomic charge computation, spatial isolation function,

RMSD of 0.01 kcal/mol, input grid spacing at 2 Å, 200 iterations). Then, due to the importance of molecular alignment in generating the pharmacological hypotheses

Table 1 (continued)

26	$pIC_{50}=5.00$	27	$pIC_{50}=4.95$	28	$pIC_{50}= 5.13$	29	$pIC_{50}= 5.28$	30	$pIC_{50}= 5.37$
31	$pIC_{50}=5.57$	32	$pIC_{50}=5.17$	33	$pIC_{50}= 5.56$	34	$pIC_{50}= 5.05$	35	$pIC_{50}= 5.18$
36	$pIC_{50}=5.23$	37	$pIC_{50}=5.44$	38	$pIC_{50}= 5.08$	39	$pIC_{50}= 5.48$	40	$pIC_{50}= 5.81$
41	$pIC_{50}=5.74$	42	$pIC_{50}=4.99$	43	$pIC_{50}= 5.27$	44	$pIC_{50}= 5.33$	45	$pIC_{50}= 4.97$
46	$pIC_{50}=5.26$	DSF	$pIC_{50}=5.98$						

underlying the 3D molecular structures in the 3D-QSAR modeling, a molecular alignment of the database items was performed (Fig. 2) [40, 46].

To develop 3D-QSAR pharmacophore models, the database of 46 molecules was divided into two sets (80% for training and 20% for testing) based on the structural diversity method and pIC_{50} scoring ranges. Table S1 presents the obtained division according to the pIC_{50} ranges of the investigated compounds (low active to high active). The training set included 37 molecules with inhibitory activity values (pIC_{50}) against 3CL^{Pro} ranging from 4.07 to 5.81. These molecules were used as input to develop the 3D-QSAR pharmacophore models. Although the test set included nine items with activity values (pIC_{50}) against 3CL^{Pro} ranging from 4.21 to 5.74,

these molecules were used as samples to test the performance and predictive power of the developed 3D-QSAR pharmacophore models.

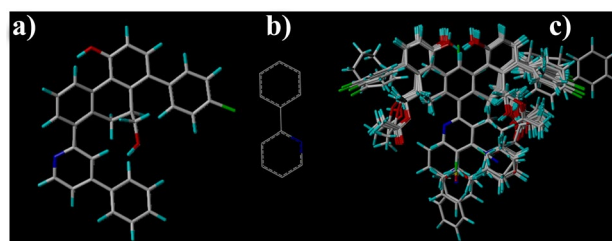


Fig. 2 a template molecule T40, b common core, and c database aligned

3D-QSAR modeling

Approaches include comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA), which are the most recent and popular statistical methods used in 3D-QSAR modeling in drug design [46]. The main importance of these approaches is to provide valuable information on key molecular structural properties related to a biological response [47]. CoMFA analysis is used to characterize the effects of steric (S) and electrostatic (E) descriptors on the biological inhibitory activity, while CoMSIA is used to characterize the effects of steric (S), electrostatic (E), hydrophobic (H), hydrogen bond donor (D), and hydrogen bond acceptor (A) descriptors on biological inhibitory activity of the compounds. In this work, we perform 3D-QSAR modeling based on CoMFA and CoMSIA techniques to describe the relationship between the biological inhibitory activity of 9,10-dihydrophenanthrene derivatives and their 3D field descriptors (S, E, H, D, and A). During this process, the linear structure–activity relationship yielded by CoMFA and CoMSIA analysis is mapped out into contour maps, which are generated by the PLS algorithm (partial least squares) and the Tripos force field (2 Å spatial reference input grid in all Cartesian dimensions, sp³ hybrid carbon atom with a net charge of + 1, default correction factor fixed at 0.3, cutoff force constant setting at 30 kcal/mol). The optimal 3D-QSAR models were selected based on the Organization for Economic Co-operation and Development (OECD) principles, the Golbraikh–Tropsha criteria, and the *Y*-randomization test related to the performance of QSAR statistical models [48–51]. Table S2 presents the significant statistical parameters adopted in this work to validate the developed 3D-QSAR pharmacophore models.

In the initial validation protocol, the internal statistical significance of the 3D-QSAR pharmacophore models was evaluated by the leave-one-out cross-validation (Loocv) approach to determine the optimal component number (*N*) and the correlation coefficient of cross-validation (Q^2). The more $Q^2 > 0.5$ and $1 < N < 6$, the more robust the 3D-QSAR model is. After selecting the ideal *N* and Q^2 for the 3D-QSAR modeling design, we evaluate the overall significance of the developed model by computing statistical parameters such as the coefficient of determination (R^2), the standard error of estimates (SEE), and *F*-value (Fischer's test). If the values of $R^2 > 0.6$ and $F_{\text{test}} > 0.3$ and low SEE, the model 3D-QSAR will be able to provide an ideal description of pIC₅₀ values based on the structural properties of investigated molecules [52].

Then, by using a *Y*-randomization test, we evaluated the stability of the correlation between pIC₅₀ values and molecular descriptors (S, E, H, D, and A) in the developed 3D-QSAR pharmacophore model. In this test, the probability of random correlation between pIC₅₀ values and molecular

descriptors is examined. We perform this test through the random distribution of pIC₅₀ data to training set items in multiple iterations (5 iterations in this work); as a result, new random QSAR models are generated. Upon successful generation of new random models, their statistical significance indices are compared to those of the original QSAR model. Generally, the original QSAR model is considered stable when its original R^2 and Q^2 parameter values are greater than the R^2 and Q^2 of the random models, and that $r^2_p > 0.5$.

After evaluating the stability of QSAR models through internal validation procedures, the external predictive power of QSAR models is examined based on items in the test set using the R^2_{pred} coefficient. As a result, when $R^2_{\text{pred}} > 0.6$, the QSAR model is robust and able to predict the biological activity of compounds reliably.

After completing the validation procedures of the 3D-QSAR pharmacophore models generated by CoMFA/SE and CoMSIA/ESHDA techniques, the structure–activity relationship can be decoded and visualized as contour maps. Contour maps were generated based on the computation of field energies at each point in the grid as standard results of the standard deviation and the coefficient associated with a specific column in the spreadsheet (stdev * coeff), plotted as a fraction in CoMFA and CoMSIA modeling equations [53]. Through the use of contour maps, which are generated by CoMFA and CoMSIA models, the favorable structural field properties of 9,10-dihydrophenanthrene derivatives can be rationalized to achieve a high level of inhibition against SARS-CoV-2-3CL^{pro}. As a result, we can derive new molecular structures from the template molecule (**T40**, pIC₅₀ = 5.81). The utilization of these models also allows us to predict the biological inhibitory activity of the designed molecules against SARS-CoV-2-3CL^{pro}, which allowed us to select molecules with a higher activity compared to synthesized molecules as well as to disulfiram used in the positive control in vitro (pIC₅₀ = 5.98).

Generation of novel compounds, drug-like screening, and ADME-Tox

We use the molecule 1-(4-bromophenyl)-10-(hydroxymethyl)-8-(4-phenylpyridin-2-yl)-9,10-dihydrophenanthren-4-ol (**T40**) as a template to generate 100 new compounds via the LigDream toolkit (<https://playmolecule.org/LigDream/>). The new compound generation is powered through generative shape-based neural network decoding [54]. Novel compounds exhibit new scaffolds and multiple chemical moieties that cover a new site in the sample chemistry space that supports drug-like properties. The pkCSM online tool was used to analyze ADME (absorption, distribution, metabolism, and excretion) pharmacokinetics of proposed drug molecules that achieved high predicted inhibitory activity and binding affinity energies toward SARS-CoV-2-3CL^{pro}, Osiris computations

were used to predict the toxicity risks of the examined molecules [55, 56].

Molecular docking

We perform molecular docking simulations to predict potential non-covalent interaction profiles between active amino acid sites in the 3CL^{pro} receptor pocket (PDB code: 6LU7) to the proposed drug molecules. In this procedure, we consider active amino acid residues that interacted with the inhibitor N3 in 3CL^{pro} (PDB code: 6LU7) as references in inhibiting the enzymatic activity of SARS-CoV-2-3CL^{pro} by the generated small molecules possessing high pIC₅₀ activity and favorable drug-like properties. The crystal structure of the main protease COVID-19 in a complex with an inhibitor N3 (resolution: 2.16 Å, *R*-value free: 0.235, *R*-value work: 0.202, *R*-value observed: 0.204) [57] was adopted as a model to perform molecular docking simulation. Initially, the backbone structure of 3CL^{pro} and the ligands were prepared. Protein is prepared by the elimination of water molecules and nonprotein elements residing in the protein, the addition of polar hydrogen atoms and Gasteiger charges, and the setting up of the flexible molecular docking grid box ($x = -10.75$ Å, $y = 12.46$ Å, $z = 68.92$ Å, grid spacing 0.375 Å, size 40 × 60 × 40 Å³), which contains amino acid residues (His41, Met49, Gln189, Cys145, Thr190, Ala191, Pro168, Leu167, Asn142, Glu166, His172, Met165, His164, Phe140, Glu143, Gln192, Thr26, Thr25, Thr24, Ser144, Gly143, Asn142, Leu141, His172, His164, and Asp187).

On the other hand, the ligands were prepared by optimizing the structures of the designed molecules as well as the inhibitors N3 and disulfiram via the three-parameter Becke hybrid method based on the Lee–Yang–Parr function (B3LYP) with 6-31G+(d,p) base in the solvent water based on the density functional theory (DFT), which is provided by the Gaussian 09 software [58]. This was done to ensure proper equilibration of the system and to check the proton and polar hydrogen atom states of the ligands in the aqueous environment [59]. In the present study, we performed molecular docking using AutoDockTools-1.5.7 and AutoDock vina [60, 61]. The molecular docking visualizations were displayed by Discovery Studio 2016 [62]. The molecular docking protocol is validated by the re-docking strategy based on the evaluation of the root mean square deviation (RMSD) between the original and re-docked N3 inhibitor [40]. After that, we evaluated the backbone stability of 3CL^{pro} protein complexed with N3 in an aqueous system by molecular dynamics simulations.

Molecular mechanics-generalized Born surface area (MM-GBSA)

MM-GBSA computations were used to re-examine novel binding modes in protein–ligand systems (screened by

molecular docking and ADME-Tox predictions). This procedure aimed at selecting the most free-bound ligands to the 3CL^{pro} active pocket site, based on MM-GBSA, which generates parameters such as binding free energy (ΔG_{bind}), hydrogen bond energy (ΔG_{bind} H-bond), van der Waals energy (ΔG_{bind} vdW), covalent energy (ΔG_{bind} Covalent), coulomb energy (ΔG_{bind} Coulomb), lipophilic energy (ΔG_{bind} Lipo), generalized Born electrostatic solvation energy (ΔG_{bind} Solv_GB), and packing energy (ΔG_{bind} Packing). MM-GBSA computations performed using MM-GBSA Prime tool available in Schrödinger 2020–3 [63, 64]. Systems (protein–ligand) were prepared by minimizing their energy using the OPLS3e force field and V5GB solvent model at pH 7 ± 2. Consequently, the energy of the protein ligands (E complex) and ligands (E ligands) is minimized, leading to the generation of binding free energy (ΔG_{bind}) for complexes (Eq. 1) [65].

$$\Delta G_{\text{bind}} = E_{\text{complex}}(\text{minimized}) + E_{\text{ligand}}(\text{minimized}) + E_{\text{receptor}}(\text{minimized}) \quad (1)$$

Molecular dynamics (MD)

Using the simulated protein environment, we perform MD simulations to evaluate the stability level of the affinity of the 3CL^{pro} (PDB code: 6LU7) enzyme to candidate drug ligands. The MD computer simulations allow us to analyze the properties of molecular systems made up of many molecules, as well as to track the trajectories of atoms and molecules based on numerically solving Newton's equations of motion of a system of interacting particles during a time frame [66, 67]. In this work, MD simulations were implemented using the Desmond package available in Schrödinger 2020–3 academic software [68]. The OPLS3e force field was used for modeling the complexes obtained by molecular docking. The docked protein–ligand systems (6LU7 uncomplexed and complexed to ligands) were solvated using the orthorhombic single point charge (SPC) explicit water model [69, 70]. Using the water model (SPC), an orthorhombic simulation box was prepared with the minimum distance between the edge surface of the protein and the protein boundary surface of 10 Å. Then, the charge of the solvated systems was neutralized to zero by adding Na⁺ and Cl[−] counter ions, adjusting the physiological salt concentration to 0.15 M, optimizing the energies of the systems to a minimum of 2000 steps, using a Columbian reaction with 9 Å cutoffs and grid phase of 0.8 Å, as well as smooth particle mesh Ewald method with a tolerance of 1E^{−09} that were used to resolve long-range electrostatic interactions [71]. The Nose–Hoover thermal algorithm and the Martina-Tobias-Klein method were employed to generate slow heating of the systems under 300 K temperature and 1.01325 pressure bar, using 10,000 frames of each simulated

trajectory, 1.2 ps recording intervals for energy, and 10 ps for trajectory under an isothermal-isobaric ensemble (NPT), 1 ps and 2 fs for system relaxation [72, 73], and finally setting up the MD simulation time set to 100 ns for each separate system.

The Maestro software GUI and the Simulation Interaction Diagram tool in the Desmond package were used to extract and analyze the dynamics of the interactions between the target protein 6LU7 and its proposed inhibitors. The Simulation Quality Analysis tool implemented in Desmond was used to analyze the dynamics of certain thermodynamic properties such as total and potential energies (ET and EP), temperature (T), pressure (P), and volume (V) depending on the time frame. This is due to the perfect compatibility of this tool's algorithms with force field computational models such as CHARMM, AMBER, and OPLS used in quantum and biochemical research [74]. Aquatic molecular dynamics simulation was the final step in this study, which is used to test the molecular dynamics stability of 3CL^{Pro} in the presence of a designed 9,10-dihydrophenanthrene derivative.

Results and discussion

3D-QSAR modeling

Upon the preparation and alignment of the database, we use the PLS algorithm to determine the linear relationship between the observed biological activity (pIC₅₀) and the field descriptors of 9,10-dihydrophenanthrene derivatives computed by the CoMFA and CoMSIA techniques. Table 2 presents the statistical significance scores obtained for selected CoMFA and CoMSIA models.

CoMFA and CoMSIA statistical models

The results of the PLS analysis presented in Table 2 indicate the high power of both CoMFA and CoMSIA models to describe the quantitative relationship between the field descriptors (S, E, H, D, and A) and the inhibitory activity of 9,10-dihydrophenanthrene derivatives against SARS-CoV-2-3CL^{Pro}. This can be confirmed by the high values of the

determination coefficients ($R^2=0.97$ and $R^2=0.94$), internal validation by (LOOCV) ($Q^2_{\text{loocv}}=0.81$ and $Q^2_{\text{loocv}}=0.76$), Fischer test ($F=107.50$ and $F=67.45$), Y -randomization test (${}^cR^2_p=0.71$ and ${}^cR^2_p=0.65$, see Table S3), external test ($R^2_{\text{pred}}=0.95$ and $R^2_{\text{pred}}=0.91$) on the one hand, and the low SEE values ($SEE=0.105$ and $SEE=0.131$) on the other hand for CoMFA and CoMSIA models, respectively. The high correlation coefficients (R^2) for the CoMFA and CoMSIA models can be confirmed by the low residual values between the observed and predicted activities through the proposed 3D-QSAR models (Table S4). Also, we can notice that electrostatic (E) and steric (S) fields have a significant effect on the predictive performance of the inhibitory activity of SARS-CoV-2-3CL^{Pro} with fractions of 58% and 30% and 41.3% and 13.5% for the CoMFA and CoMSIA models, respectively. Furthermore, through the CoMSIA model, we can see the strong dependence of hydrophobic (H) and donor hydrogen bond (D) fields (27.5% and 21.2%) on the predictive power compared to hydrogen bond acceptor (A) fields (7.9%). By proportioning the contributions of the field descriptors (fractions), we can conclude that the biological inhibitory activity of 9,10-dihydrophenanthrene derivatives is influenced principally by their electrostatic (E), steric (S), hydrophobic (H), and hydrogen bond acceptor (A) properties. All statistical parameters obtained from CoMFA and CoMSIA analyses indicate the excellent performance of the CoMFA/SE and CoMSIA/SEHDA models in predicting inhibitory activity against 3CL^{Pro} based on the structure of 9,10-dihydrophenanthrene derivatives. Thus, these models can be reliably exploited to predict the bioactivity of new small molecules that can be modeled.

Analysis of CoMFA and CoMSIA contour maps

The 3D QSAR models can be interpreted at the molecular scale based on contour surface plots of the 3D structure–activity relationships generated by the proposed CoMFA and CoMSIA models (Figs. 3, 4). From Figs. 3a, b and 4a, b, we can notice a very similar spatial arrangement of the steric and electrostatic contours of CoMFA and CoMSIA on the structure of the template molecule (T40), which confirms the significant impact of the descriptors (S and E)

Table 2 Summary of statistical significance results for CoMFA and CoMSIA models

	Q^2	R^2	SEE	F	N	R^2_{pred}	${}^cR^2_p$	S	E	H	D	A
Threshold	> 0.5	> 0.6	Small	High	1–6	> 0.6	> 0.5	0 < Fractions < 1				
CoMFA	0.81	0.97	0.105	107.50	3	0.95	0.71	0.413	0.587	-	-	-
CoMSIA	0.76	0.94	0.131	67.45	4	0.91	0.65	0.135	0.299	0.275	0.079	0.212

Q^2 coefficient of cross-validation correlation, N optimal number of components identified by leave-one-out cross-validation (loocv), SEE standard error of estimate, R^2 conventional coefficient of determination, R^2_{pred} coefficient of determination according to the external test, ${}^cR^2_p$: Y -randomization test; fractions: contributions of steric (S), electrostatic (E), hydrophobic (H), donor (D), and acceptor (A) hydrogen bonds

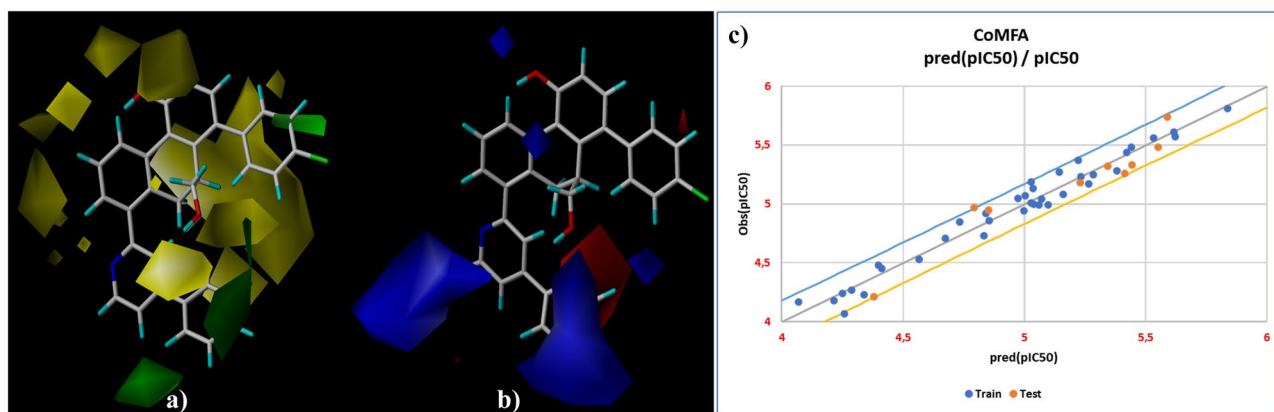


Fig. 3 Contour maps generated by CoMFA model, **a** steric field interactions (green = 80% favorable/yellow = 20% unfavorable), **b** electrostatic field interactions (blue = 80% favorable/red = 20% unfavorable). **c** pIC_{50} observed vs. pIC_{50} predicted

on the inhibitory biological activity (pIC_{50}). The predominance of yellow contour maps over green contour maps in the molecule **T40** backbone indicates that large ring radicals negatively affected the bioactivity of 9,10-dihydrophenanthrene derivatives against SARS-CoV-2 3CL^{pro}. Thus, the biological activity of the new molecules derived from molecule **T40** can be improved by introducing smaller moieties into the pharmacological sites covered by the yellow contours. Moreover, the predominance of blue contours over red contours along with the 4-phenylpyridin-2-yl moiety indicates that the enhancement of this site with electron-donating groups is favorable for the improved biological activity of (1-(4-bromophenyl)-10-(hydroxymethyl)-8-(4-phenylpyridin-2-yl)-9,10-dihydrophenanthren-4-ol) against SARS-CoV-2 3CL^{pro}.

On the other hand, the positioning of the white contours on the carbon sites C22 and C30 in the bromobenzene and

phenyl rings indicates that these sites are unfavorable for hydrophobic moieties; so, the upgrade of C22 and C30 sites with hydrophilic moieties can be favorable to improve pIC_{50} biological inhibitory activity (Fig. 4c). The dominance of the cyan over purple contour near the hydroxyl group indicates that this site is suitable for 80% of the hydrogen bond donor radicals (Fig. 4e), while the 4-phenylpyridin-2-yl region covered by 80% of the magenta contours is suitable to enhance the biological activity (pIC_{50}) through the insertion of hydrogen bond donor radicals at this site (Fig. 4e). Furthermore, the linear distribution of observed pIC_{50} values versus those predicted by the CoMFA and CoMSIA models (Fig. 3c and Fig. 4f) confirms the hypotheses of a high pharmacological correlation between pIC_{50} and the field properties (S, E) and (H, D, A) of 9,10-dihydrophenanthrene derivatives. As a result, the developed 3D-QSAR pharmacophore models (CoMSIA and CoMFA) can be exploited reliably

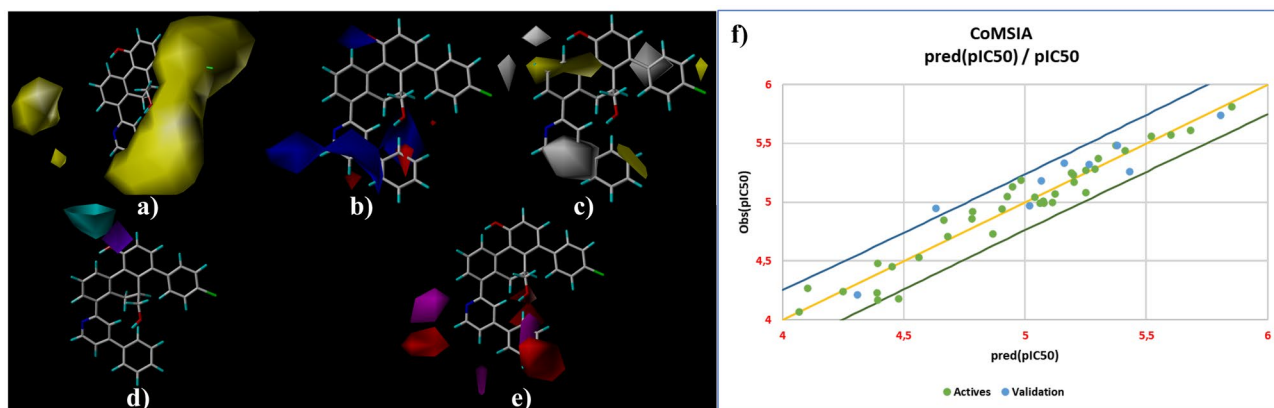


Fig. 4 Contour maps generated by the CoMSIA model, **a** steric field interactions (green = favorable/yellow = unfavorable), **b** electrostatic field interactions (blue = 80% favorable/red = 20% unfavorable), **c** hydrophobic field interactions (yellow = 80% favorable/white = 20%

unfavorable), **d** hydrogen bond-donor field interactions (cyan = 80% favorable / purple = 20% unfavorable), **e** hydrogen bond-acceptor field interactions (magenta = 80% favorable/red = 20% unfavorable). **f** pIC_{50} observed vs. pIC_{50} predicted

to rationalize the generation of new molecules based on reference molecule **T40** and predict their activities against SARS-CoV-2 3CL^{pro}.

Generation of novel compounds from 1-(4-bromophenyl)-10-(hydroxymethyl)-8-(4-phenylpyridin-2-yl)-9,10-dihydrophenanthren-4-ol (**T40**)

Based on the pharmacological hypotheses obtained by the CoMFA and CoMSIA models (Fig. 5), new sets of (1-(4-bromophenyl)-10-(hydroxymethyl)-8-(4-phenylpyridin-2-yl)-9,10-dihydrophenanthren-4-ol) derivatives can be generated and their biological activity and pharmacokinetic properties can be predicted.

Through the use of LigDream and the artificial neural network tools in the PlayMolecule platform (<https://playmolecule.org/LigDream/>), 96 new molecules (D1–D96) were generated from the template compound (**T40**) based on auto-encoders and captioning networks approaches [75]. Once the new molecules are generated, we filter them based on the drug-likeness criteria (e.g., Lipinski RO5, Ghose violations, bioavailability, and synthetic accessibility) [76–79]. Table S5 shows the drug-like profile of the 96 new molecules generated from the structure of the reference molecule (**T40**), as well as the **N3** inhibitor properties. The inflexible drug-like screening predicted that 47 new molecules did not violate any of Lipinski's and Ghose's rules (Table S5). Also, we can notice that the bioavailability and synthetic accessibility of the 47 screened small molecules were 0.55 and ($2 < SA < 4$), respectively. Template molecule (**T40**) and inhibitor **N3** were observed to violate Lipinski's (2 violations) and Ghose's (3 violations), with poor bioavailability (0.17) and synthetic accessibility (4.07 and 6.43). This indicates that the 47 new molecules can act as an effective

alternative oral drug against 3CL^{pro} compared to the template molecule (**T40**) and inhibitor **N3**. To further validate these hypotheses, we apply the CoMFA and CoMSIA models to the new 47 candidate molecules and predict their inhibitory activity against SARS-CoV-2 3CL^{pro}. The pIC₅₀ of the novel designed molecules were predicted after preparing their 3D structures using the same strategy employed for 3D-QSAR modeling in this work (Fig. 6).

Thanks to the perfect alignment between the 47 filtered molecules and the template structure shown in Fig. 6, we can conclude that the structural improvements made to the structure of compound **T40** were coherent with the pharmacological hypotheses extracted from this structure via the proposed CoMFA and CoMSIA models. Hence, the predictions of the biological inhibitory activity of new molecules against SARS-CoV-2 3CL^{pro} by the CoMFA and CoMSIA models will be more stable and precise. Table 3 presents the predicted biological activities generated by CoMFA and CoMSIA models for all 47 candidate non-covalent SARS-CoV-2 3CL^{pro} inhibitors.

From Table 3, we can notice that twenty new derived molecules (**D02**, **D06**, **D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, **D28**, **D30**, **D38**, **D43**, **D44**, **D76**, **D87**, **D88**, **D92**, **D94**) showed significantly better inhibitory activity ($pIC_{50} > 6$) than the reference inhibitors, which are molecule **T40** ($pIC_{50} = 5.8$), **DSF** ($pIC_{50} = 5.98$) and **N3** ($pIC_{50} = 3.90$). This means that the modular changes made to the structure of the template molecule (**T40**) were favorable to improving the biological activity of the proposed new class of non-covalent inhibitors of SARS-CoV-2 3CL^{pro}. Thus, the 20 computer-based non-covalent inhibitors proposed in the current study (Fig. S1) can reach an excellent inhibitory activity against SARS-CoV-2 3CL^{pro} and reach the therapeutic goal against COVID-19. To further confirm these hypotheses,

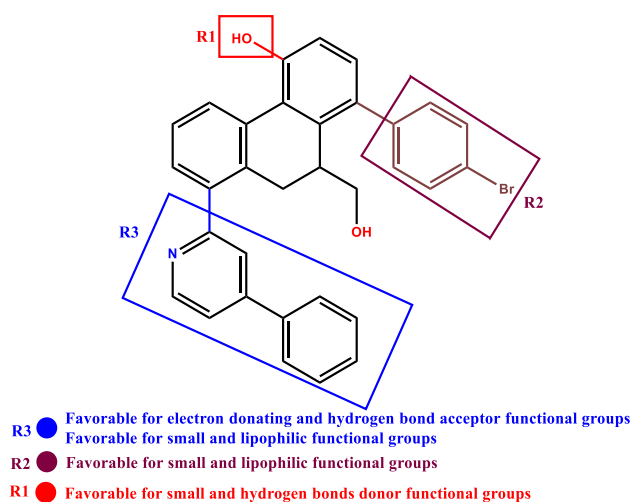


Fig. 5 Characterization of structural properties favorable to design novel SARS-CoV-2 3CL^{pro} inhibitors based on molecular structure (**T40**)

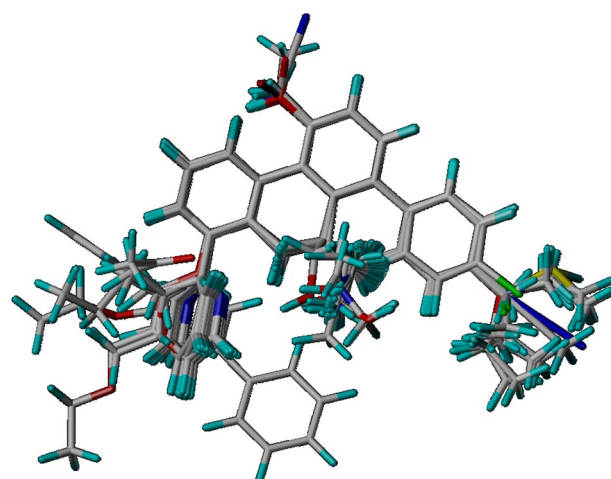


Fig. 6 3D visualization of molecular alignment poses between the template (**T40**) and newly generated compounds

Table 3 The predicted pIC₅₀ activities of 47 screened molecules against SARS-CoV-2 3CL^{pro}

Comp	pIC ₅₀		Comp	pIC ₅₀		Comp	pIC ₅₀	
	CoMFA/SE	CoMSIA/SEHD		CoMFA/SE	CoMSIA/SEHD		CoMFA/SE	CoMSIA/SEHD
D02	6.302	6.938	D25	6.144	6.878	D76	6.300	6.986
D03	6.146	5.918	D26	6.290	6.913	D78	5.180	4.789
D04	6.080	5.929	D27	6.149	6.988	D79	5.167	4.781
D05	6.123	5.877	D28	6.238	6.858	D80	5.160	4.827
D06	6.261	6.967	D30	6.105	6.847	D81	5.163	5.013
D07	6.187	6.893	D38	6.165	6.736	D82	5.179	4.994
D08	6.259	6.852	D41	5.125	5.007	D84	5.172	5.997
D12	6.240	6.826	D42	5.195	4.918	D87	6.221	6.005
D13	6.161	5.931	D43	6.246	6.860	D88	6.268	6.015
D14	6.107	5.950	D44	6.159	6.724	D92	6.239	6.969
D17	6.143	5.963	D45	5.809	4.924	D94	6.226	6.850
D18	6.218	6.857	D46	5.278	4.842	-	-	-
D19	6.190	5.947	D49	3.930	4.675	-	-	-
D20	5.891	4.937	D54	5.227	4.965	-	-	-
D21	6.106	5.983	D60	5.250	5.029	-	-	-
D22	6.369	5.994	D62	5.060	4.893	-	-	-
D23	6.104	6.231	D70	4.997	4.885	-	-	-
D24	6.145	5.830	D73	4.861	4.767	-	-	-

Templates (compound T40: pIC₅₀=5.81), inhibitors (DSF: pIC₅₀=5.98), and (N3:pIC₅₀=3.90 [29])

Compounds with boldface values showed better inhibitory activity than the reference inhibitors (T40, DSF and N3)

we perform molecular docking simulations and evaluate the level of binding energies of proposed molecules to the 3CL^{pro} enzyme.

Molecular docking simulation

Molecular docking protocol validation

First, we perform a re-docking of the covalent inhibitor **N3** with the 3CL^{pro} receptor; this is done to validate the molecular docking protocol and also to get insight into the reference active amino acid residues involved in non-covalent interactions inside the 3CL^{pro} protein pocket (PDB code: 6LU7). Figure 7 displays 3D and 2D visualizations of the re-docking pathways of ligand **N3** inside the 3CL^{pro} receptor pocket (6LU7).

The perfect superposition mode (RMSD = 0.131 < 2 Å) between the original and re-docked **N3** ligand inside the 3CL^{pro} pocket (Fig. 7a) indicates the high precision of molecular docking predictions performed by AutoDock vina software. Also, Fig. 7b indicates that the **N3** ligand (original and re-docked) interacted with the same active amino acid residues that are His41, Met49, Gln189, Thr190, Ala191, Pro168, Leu167, Cys145, Glu166, His172, Met165, His164, Phe140, and Glu143. Furthermore, we can see that most non-covalent interactions between the ligand **N3** and the protein 6LU7 resulted from hydrophobic (Alkyl, Pi-Alkyl, Pi-Sigma), hydrogen bonds (conventional and carbon),

and van der Waals interactions. Therefore, we consider the interactions of molecule **T40** and its novel derivatives with certain of these residues as a reference mechanism for inhibiting 3CL^{pro} via non-covalent interactions. To complete the molecular docking validation protocol, we performed a molecular dynamics simulation on the crystalline complex (PDB:6LU7), this was done to evaluate the stability of the 3D structure of the 6LU7 receptor in the aqueous environment, as well as the stability of **N3** in the active pocket of 3CL^{pro} (Fig. 8).

Figure 8a indicates that the 6LU7 backbone RMSD reached a remarkable equilibrium around the 1.8 Å range with some fluctuations that did not exceed approximately 2.4 Å throughout the MD simulation. The **N3** ligand showed fluctuations in RMSD ranging from 2.4 to 4.8 Å from the onset of the simulation to around 35 ns, before stabilizing in the range of 1.2 to 3 Å until the end of the MD simulation time. From Fig. 8b and c, we can notice that **N3** contacted the same active amino acid residues in the 6LU7 pocket predicted by the molecular re-docking protocol. The majority of the interactions were of the hydrogen bond and hydrophobic and water-bridge type. Most of the interactions formed with **N3** inside the active pocket of 3CL^{pro} spanned from 25 to more than 100% of the MD simulation time, this is due to the large size of the **N3** structure, which does not allow it to perform more flexible interactions inside the active pocket. Therefore, small molecules may be more flexible in terms of their interactions and stability in the 3CL^{pro} pocket

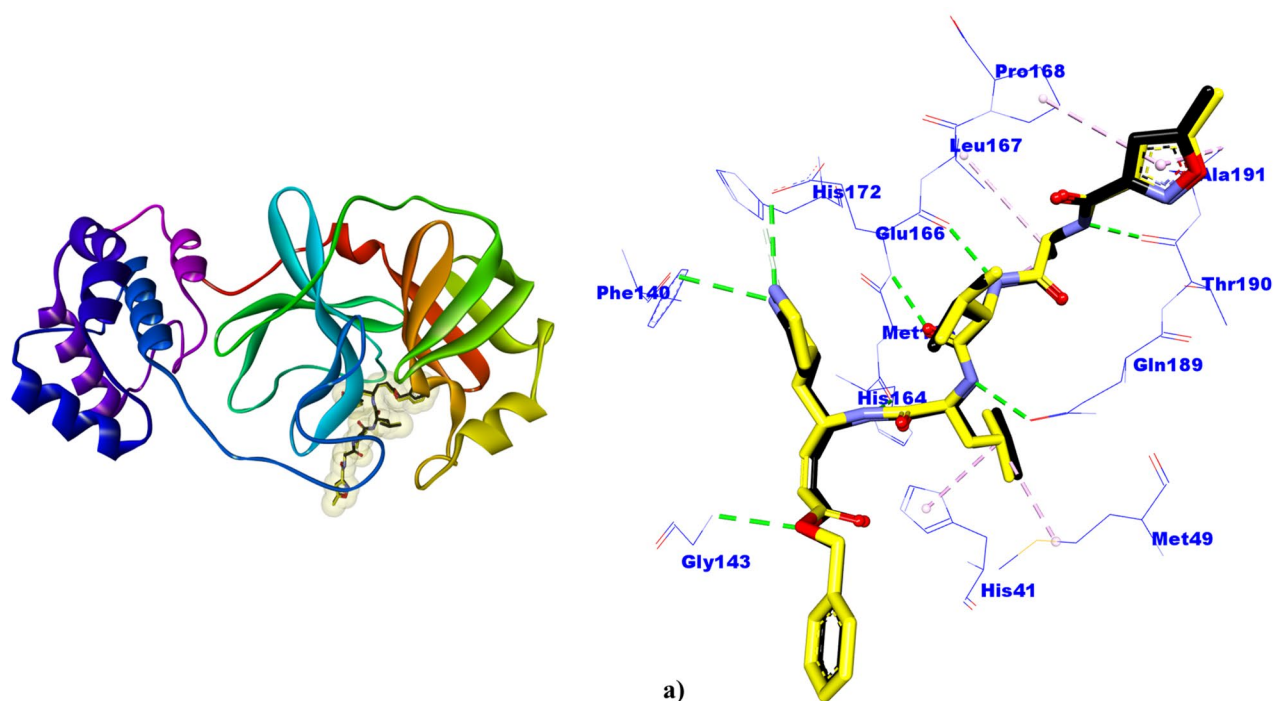


Fig. 7 **a** 3D superposition of original (black) and re-docked (yellow) N3 ligands in the 6LU7 active pocket of 6LU7 (RMSD=0.121 Å). **b** N3 (native and re-docked) interaction patterns with active residues in the 3CL^{pro} pocket

compared to large molecules such as the peptide inhibitor **N3** as a model.

Molecular docking test

Molecular docking was performed to find potential active residue sites in the 3CL^{pro} enzyme pocket with which the designed and synthesized 9,10-dihydrophenanthrene derivatives can interact. Furthermore, we evaluated the binding affinity energies of the examined ligands and screened the best potential non-covalent inhibitors of 3CL^{pro}. This is done by selecting ligands that reach the lowest energy level when binding to the active pocket of the 3CL^{pro} receptor.

Figure 9 shows a 3D visualization of the optimal conformation positions of the new 20 derivatives of 9,10-dihydrophenanthrene, **T40**, inhibitors **DSF** and **N3** in the active pocket of SARS-CoV-2 3CL^{pro} (PDB code:6LU7).

Figure 9 shows the optimal docking position of the investigated molecules in the active pocket of the 3CL^{pro}. This means that the small molecules of 9,10-dihydrophenanthrene derivatives can target the enzymatic activity of SARS-CoV-2 3CL^{pro} and achieve a potential inhibition of 3CL^{pro} through non-covalent interactions. Table S6 provides a detailed summary of the most important molecular docking results regarding the binding affinity energies and interactions of the studied ligands in the active pocket of 3CL^{pro}. From the molecular

docking results presented in Table S6, we can notice that all new modeled molecules, as well as **T40** and the **DSF** inhibitor, interacted with the majority of active residues in the 3CL^{pro} pocket with which inhibitor **N3** interacted. Most non-covalent interactions that were formed between the investigated molecules and 3CL^{pro} were hydrophobic, hydrogen bonds (Pi-donor, carbon, and conventional hydrogen bonds), electrostatic, and van der Waals. This further confirms that the 9,10-dihydrophenanthrene-derivatives are flexible in terms of inhibiting the enzymatic activity of 3CL^{pro} by non-covalent mechanisms. Thus, active sites (His164, Cys145, Glu166, His41, His163, Gly143, Ser144, Asn142, Phe140, Leu141, Met165, Pro168, Leu167) in the 3CL^{pro} pocket, which interact with template molecule (**T40**) can consider those active sites as new potential keys involved in inhibition of SARS-CoV-2 3CL^{pro} enzymatic activity through non-covalent interactions.

Analysis of binding affinity energies of ligands From Table S6, we can notice that among twenty proposed molecules 11 molecules **D06** (−10.5 kcal/mol), **D07** (−10.7 kcal/mol), **D08** (−10.8 kcal/mol), **D12** (−10.6 kcal/mol), **D18** (−10.4 kcal/mol), **D23** (−10.9 kcal/mol), **D25** (−10.6 kcal/mol), **D26** (−10.5 kcal/mol), **D27** (−10.5 kcal/mol), **D30** (−10.5 kcal/mol), **D76** (−11.1 kcal/mol) appeared to be more stable in terms of binding affinity energies compared to template molecules **T40** (−10.2 kcal/mol) and **DSF** (−4.7 kcal/mol) inside the active site of the main protease

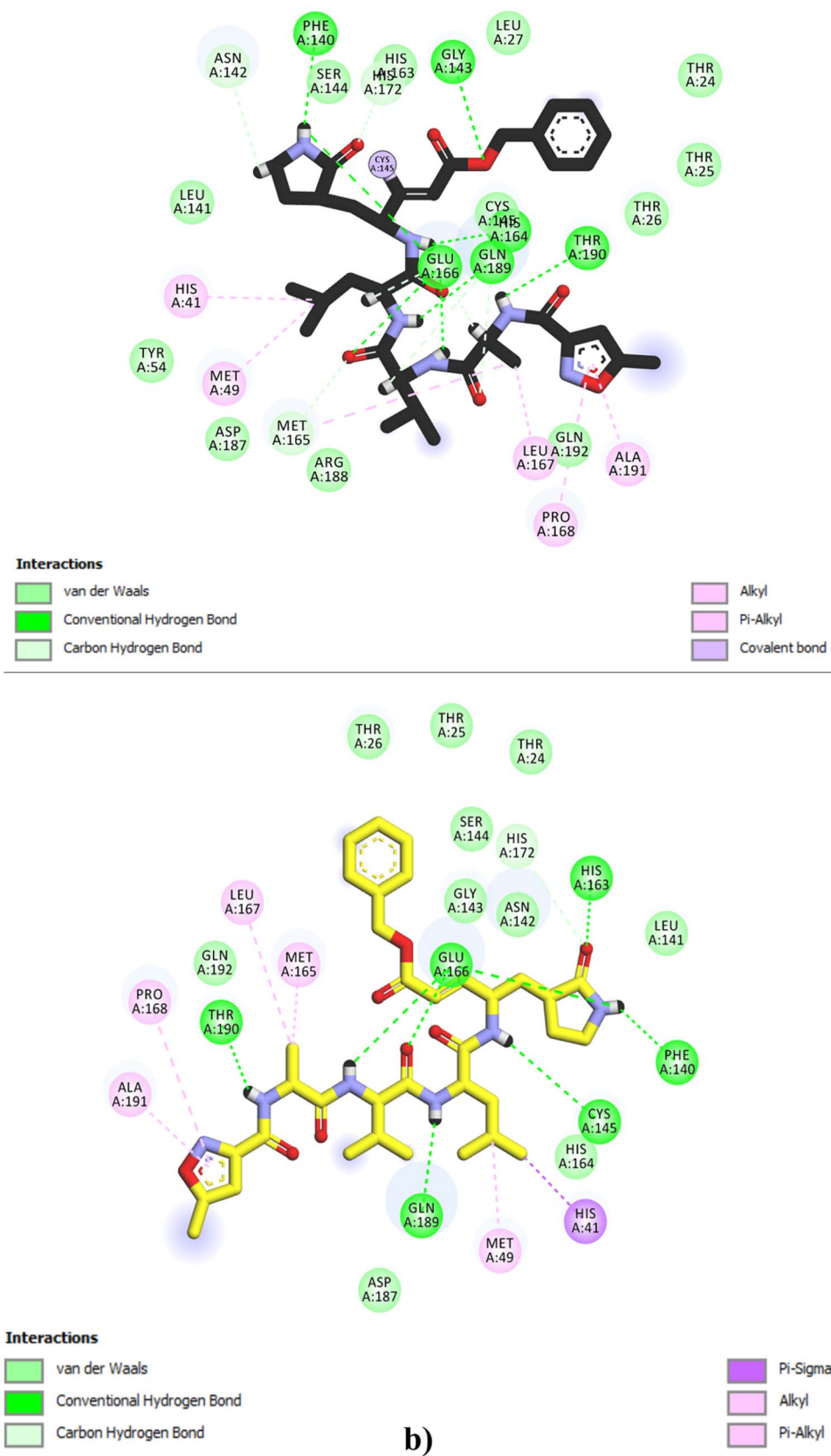


Fig. 7 (continued)

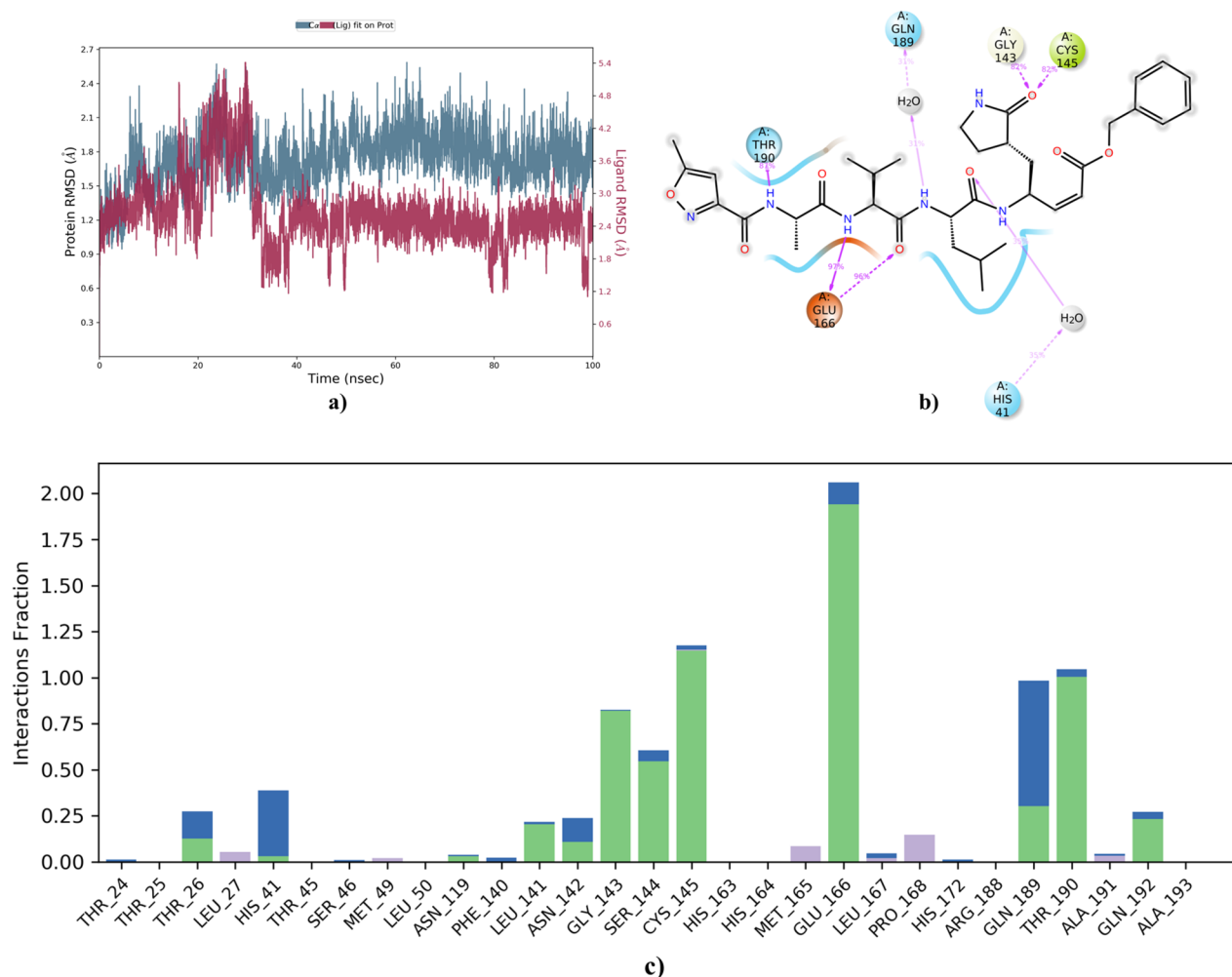


Fig. 8 **a** The root means square deviation of the atomic positions of the 6LU7 protease and the N3 inhibitor. **b** 2D visualization of contacts made between 6LU7 and N3 that persist beyond 30% of the MD

simulation. **c** Histogram of contacts between 6LU7 and N3 during 100 ns of the MD simulation

pocket. Likewise, we can note that the eleven candidate molecules were able to interact with most of the active amino acid residues with which the reference inhibitor **N3** (−13.5 kcal/mol) interacted within the active pocket of 3CL^{pro}. Thus, the high biological binding of the proposed small molecules to 6LU7 is likely to achieve excellent inhibition against SARS-CoV-2 3CL^{pro} compared to the template molecule **T40**.

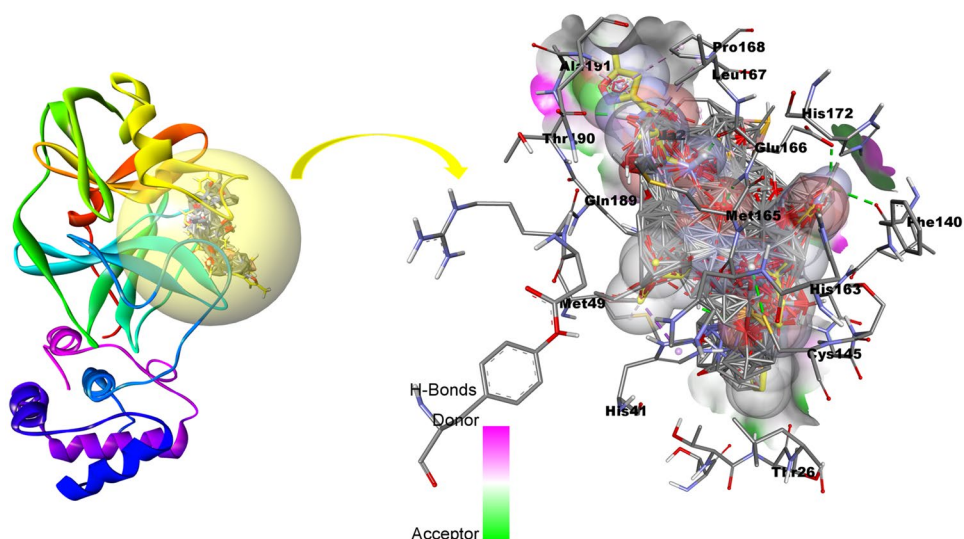
Non-covalent interaction analysis of examined ligands Figure S2 displays a 2D visualization plot of key non-covalent interactions between the active reference residues in the 3CL^{pro} pocket and the eleven filtered molecules adjacent to templates (**T40** and **DSF**). All non-covalent ligand–protein interactions identified in Fig. S2 are classified and summarized in Table 4.

From Fig. S2 (complex DSF-6LU7), we can observe that the structure of the **DSF** inhibitor used in the positive control

interacts via a carbon-hydrogen bond with the active reference sites occupied by Gln189 (5.82 Å), three hydrophobic bonds (alkyl and pi-alkyl) with reference active sites occupied by residues Leu27 (4.27 Å), Cys145 (5.06 Å), and His41 (5.59 Å), as well as other interactions of the van der Waals class.

However, the non-covalent interactions generated between **T40** (template) and the reference active sites in 3CL^{pro} (Fig. S2, complex T40-6LU7) indicate that the moiety (OH, R1) interacts with the active reference residues His164 (5.00 Å), Cys145 (6.77 Å) through two conventional hydrogen bonds. Moreover, the moiety (bromobenzyl, R2) interacted with one electrostatic bond with the reference site Glu166 (4.79 Å), and three hydrophobic bonds (stacked amide-Pi, alkyl, Pi-alkyl) between the moiety (4-phenylpyridine, R3) and the active reference sites Leu167 (6.54 Å), Pro18 (4.33 Å), Met165 (5.06 Å), as well as other interactions of van der Waals class. The overall non-covalent interactions and number of hydrogen

Fig. 9 Molecular docking of the best conformations of the selected compounds, **T40**, **N3**, and **DSF** in the active pocket of the SARS-CoV-2 3CL^{pro} (the re-docked N3 ligand is highlighted in yellow)



bonds generated in the complex (40-6LU7) led to higher binding affinity energy for template 40 (BAE = -10.2 kcal/mol) compared to DSF inhibitor (BAE = -4.7 kcal/mol). Also, from Fig. S2, (complex 40-6LU7), we can notice the main contribution of pharmacophore sites (R1, R2, R3) in the generation of most non-covalent inhibitory interactions. This implies that the 3D-QSAR pharmacophore modeling predictions were accurate in rationalizing the pharmacophore sites in the structure of template molecule **T40** that are favorable to achieving high inhibitory activity against SARS-CoV-2 3CL^{pro}. As a result, the structural modifications applied to the structure of **T40** explain the significant improvement in the inhibitory activity and binding affinity energies of the eleven novel molecules (**D06**, **D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, **D30**, and **D76**) screened as promising inhibitors of 3CL^{pro} enzymatic activity.

To sum up, the molecular docking simulation results indicated that the eleven novel small molecules achieved perfect docking poses in the active pocket of 3CL^{pro} where the peptide inhibitor **N3** interacts. The optimal binding modes, high binding affinity energies, and high predicted pIC₅₀ biological activity of these molecules provide a strong theoretical basis for the nomination of these molecules as novel drug agents against COVID-19.

ADME-Tox prediction

The discovery of coronavirus protease inhibitors targeting 3CL^{pro} has been a major innovation due to the poor pharmacokinetic properties of large peptide compounds such as the **N3** (PubChem Compound CID: 6,323,191) covalent inhibitor model [80, 81]. Thus, the discovery of novel non-covalent 3CL^{pro} inhibitors based on small molecule structures could be an effective therapeutic key against COVID-19 compared to covalent 3CL^{pro} inhibitors. The evaluation of

absorption, distribution, metabolism, excretion, and toxicity (ADME-Tox) properties is important in drug discovery and design to ensure effective and safe drugs. Based on this, we evaluate the pharmacokinetic properties (ADME-Tox) of eleven small molecules (**D06**, **D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, **D30**, and **D76**) candidates to reach the 3CL^{pro} enzyme with sufficient inhibitory concentration, biological binding, and high drug response. Table S7 presents ADME-Tox predictions assessed in this work.

From the ADME-Tox predictions presented in Table S7, it can be observed that all examined molecules can be a substrate of P-glycoprotein, which explains the high absorption rates of these molecules in the human intestine (> 90%). In contrast, the **N3** inhibitor showed a low absorption of 57.88%; so, the small intestine may not be able to absorb enough **N3** peptide, which may lead to rapid clearance before reaching the therapeutic target. Regarding distribution property, the negative VD_{ss} values for all molecules indicate that the distribution of these molecules is in the bloodstream of the body. The significantly low fraction unbound (FU) values of the examined molecules suggest that these molecules are more likely to bind the protease than to the plasma. Furthermore, all molecules were unable to penetrate the central nervous system (CNS) and the blood–brain barrier (BBB), except molecule **D12** (logBB = 0.29), which means that the proposed molecules have no potential impact on the brain. In terms of metabolism, all examined molecules could be substrates for the enzyme cytochrome 3A4 is responsible for drug metabolism in the human body; so, the drug compounds will not be rapidly excreted and metabolized by the human body. Regarding the excretion property, the low total clearance values of all molecules mean that the 11 proposed drug molecules as well as the template molecule **T40**, have significant half-life and stability in the body. This may allow them to reach their therapeutic target before excretion, unlike

Table 4 Summary of molecular docking predictions for best-modeled molecules, template molecules

Ligands	Binding energy (kcal/mol)	Hydrogen-binding interactions	Electrostatic interactions	van der Waals interactions	Hydrophobic interactions
D5F	−4.7	Gln189 (3.50 Å), Asn142 (3.51 Å)	-	Thr25, Thr26, Met49, Gly143, Glu166, Met165	Leu27 (4.06 Å), Cys145 (4.63 Å), His41 (4.99 Å)
T40	−10.2	Cys145 (3.02 Å), His164 (2.72 Å)	Cys145(5.24 Å), Glu166(3.95 Å)	His41, His163, Gly143, Ser144, Asn142, Phe140, Leu141,	Met165 (4.89,5.38, 5.41 Å), Pro168 (4.04 Å), Leu167 (4.50 Å)
D06	−10.5	Gln189 (3.35 Å)	Met49(4.92 Å), His41(4.54 Å), Cys145(5.20 Å)	Tyr54, Pro52, His164, Glu166, Gln192, Thr190, Asp187, Arg188, Leu167	Met165 (5.50 Å), Pro168 (4.72 Å)
D07	−10.7	Ser144 (3.53 Å), Asn142 (2.81 Å), Gly143 (2.83 Å)	-	Glu166, Met165, His163, Phe140, Leu141, Met49, Ser46, Pro168	Cys145 (5.06,4.56 Å)
D08	−10.8	Ser144 (2.54 Å), Gly143 (2.82 Å), Gly166 (3.06 Å), Thr25 (2.81 Å), Thr24 (2.69 Å)	-	Met165, His163, Leu141, Asn142, His41, Leu27, Thr45, Met49, Gln189	Cyc145 (4.87,5.06 Å)
D12	−10.6	Ser144 (2.42 Å), Asn142 (2.64 Å), Gly143 (2.64 Å)	-	Glu166, Met165, His163, Phe140, Leu141, Leu27, Ser46, Met49	Cys145 (4.94,5.13 Å)
D18	−10.4	Ser144 (2.33 Å), Gly143 (2.80 Å), Asn142 (3.48 Å)	-	Glu166, Met165, His163, Phe140, Leu141, Leu27, Ser46, Met49, Thr24, Thr25	Cys145 (5.13,4.82 Å)
D23	−10.9	Gln192 (2.61 Å), Arg188 (2.55 Å), Met165 (2.84 Å)	-	Leu167, Thr190, Asp187, His164, His41, Met49, Leu27, Cys145, Thr26, Glu166, Gly143	Met165 (5.32, 4.22 Å), Gln189 (3.48 Å), Pro168 (5.13,5.23 Å)
D25	−10.6	Asn142 (2.71 Å), Ser144 (2.36 Å), Gly143 (2.65 Å)	-	Pro168, Glu166, Met165, His163, Leu141, Phe140, Leu27, Met49, Ser46, Met49	Cys145 (5.07,5.10 Å)
D26	−10.5	Arg188 (2.54 Å), Cys145 (3.95 Å), Gln189 (3.42 Å), Gln192 (unfavorable donor-donor, 2.14 Å)	His41(3.49 Å)	Pro168, Thr190, Asp187, His164, Met49, Gly143, Glu166	Met165 (5.16, 4.36 Å)
D27	−10.5	Thr190 (3.09 Å), Gln192 (2.97 Å), Arg188 (2.53 Å)	-	Leu167, Gln189, His164, Gly143, Asn142, Glu166	Pro168 (5.42 Å), Met165 (4.40, 4.92 Å), Cys145 (5.07 Å)
D30	−10.5	Asn142 (2.68), Ser144 (2.43), Gly143 (2.67)	-	Glu166, Met165, His163, Phe140, Leu141, Leu27, Ser46, Met49	Cys145 (4.96,5.17 Å)
D76	−11.1	Arg188 (2.46 Å), Met165 (3.02 Å), Glu166 (2.17 Å), Gln189 (3.22 Å)	-	Gln192, His164, Asn142, Ser144, Gly143, Leu141, Leu167	Met165 (4.78, 5.44 Å), Pro168 (5.30 Å), Cys145 (5.12 Å)

Table 5 Prime MM-GBSA energies for binding of ligands to the active site of 3CL^{PRO} compared to the references N3 and T40

MM-GBSA/(kcal/mol)	Selected ligands/6LU7									Reference ligands/6LU7	
	D07	D08	D12	D18	D23	D25	D26	D27	D76	T40	N3
ΔG_{bind}	-53.51	-55.15	-51.93	-50.43	-59.54	-52.24	-48.92	-47.83	-57.30	-47.77	-83.84
ΔG_{bind} Vdw	-45.89	-56.18	-44.44	-44.09	-52.87	-48.12	-42.18	-42.97	-51.58	-48.34	-84.31
ΔG_{bind} Hbond	-2.00	-2.21	-1.85	-1.99	-0.16	-1.86	-0.34	-0.44	-0.31	-2.72	-0.56
ΔG_{bind} Coulomb	-15.84	-20.62	-12.05	-11.07	-17.54	-13.56	-4.69	-6.17	-6.05	-7.24	-30.72
ΔG_{bind} Lipo	-19.33	-16.97	-15.47	-15.51	-21.99	-18.39	-20.00	-20.40	-19.92	-22.83	-20.61
ΔG_{bind} Packing	-2.54	-2.71	-2.28	-2.47	-1.01	-2.28	-2.21	-1.73	-2.26	-2.71	-0.03
ΔG_{bind} Solv_GB	26.11	28.50	21.15	21.87	21.02	23.87	21.34	24.24	22.76	28.82	46.07
ΔG_{bind} Covalent	5.00	1.04	3.02	2.85	2.02	3.11	2.16	3.95	2.07	4.10	8.50

the **N3** inhibitor, which has a high total clearance value of 0.65. Regarding the risks of toxicity, Osiris predictions show that **D06** is likely to have slight toxicity as irritation, as well as **D30**, which is mutagenic and carcinogenic, whereas the remaining compounds did not indicate any potential toxicity risks. Due to the potential toxicity risks of **D06** and **D30**, we eliminate them and keep only those molecules that do not have any predicted toxic effect. Based on ADME-Tox predictions, the nine selected molecules (**D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, and **D76**) are identified as highly favorable for drug design.

MM-GBSA computations

Table 5 shows the binding free energy values obtained by MM-GBSA for the selected ligands (**D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, and **D76**) and also for references (**T40** and **N3**). Although the ligands (**T40** and **N3**) did not show a good fit for drug-like features, their presence as references were necessary during the in silico validation routines. The MM-GBSA computations presented in Table 6 indicate that the binding energies of the designed ligands range from the lowest value of -47.83 kcal/mol (**D27**) to the highest level of -59.54 kcal/mol (**D23**). The proposed drug ligands showed relatively higher binding energies than the reference molecule **T40** (-47.77 kcal/mol) and less than the **N3** peptide inhibitor (-83.84 kcal/mol). Among the ligand/protein binding patterns evaluated via MM-GBSA, we can notice the significant contribution of (ΔG_{bind} Vdw), (ΔG_{bind} H-bond), (ΔG_{bind} Coulomb), (ΔG_{bind} Lipo), and (ΔG_{bind} Packing) energies to the average binding energy (ΔG_{bind}) of the nine proposed drug molecules and also for the references (**T40** and **N3**). The positive energy contribution of (ΔG_{bind} Solv_GB) and (ΔG_{bind} Covalent) was not favorable to (ΔG_{bind}), which can lead to resistance against the binding. This means that the effect of the energies of non-binding interactions is much more favorable to achieving high and equilibrium stability of the examined ligands in the active

pocket of 3CL^{PRO} compared to potential covalent interactions. In brief, the obtained MM-GBSA computations confirm the molecular docking results related to the high binding affinity energies of the proposed small molecule (**D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, and **D76**) toward the active site inside the 3CL^{PRO} pocket.

Molecular dynamics simulations

Molecular dynamics simulations were performed to examine the stability level of potential non-covalent interactions between the active residues of the 3CL^{PRO} (6LU7) pocket and investigated 9,10-dihydrophenanthrene derivatives. For this purpose, we consider the samples D08-6LU7 ($\Delta G_{\text{bind}} = -55.15$ kcal/mol), D23-6LU7 ($\Delta G_{\text{bind}} = -59.54$ kcal/mol), and D76-6LU7 ($\Delta G_{\text{bind}} = -57.30$ kcal/mol) as test items and sample T40-6LU7 ($\Delta G_{\text{bind}} = -47.77$ kcal/mol) as a reference in MD simulations protocol.

RMSD and RMSF analysis

The root mean square deviation (RMSD) and fluctuation (RMSF) parameters were used to estimate the range of potential fluctuations in the backbone of 6LU7 proteases (uncomplexed and complexed). The RMSD and RMSF indices express the average α -carbon backbone for all atoms of amino acid residues that formed the 6LU7 systems. In this work, RMSD and RMSF variations of examined 6LU7 systems were evaluated based on the first frame of the uncomplexed 6LU7 backbone (6LU7 free). Figure 10 shows the RMSD and RMSF time scales obtained after 100 ns of MD simulation trajectory.

From Fig. 10a, it appears that the C α atoms of the 6LU7 protease backbone were not affected in their stability after **D08**, **D23**, and **D76** ligands docked into the active pocket of 6LU7 during 100 ns of the MD trajectory. Some slight fluctuations in the amino acid side chains were observed in the

Table 6 Detected 6LU7-ligand contacts along the 100 ns trajectory of MD simulations

Complex	Protein–ligand contacts		
	H-bonds	Hydrophobic	Water bridges
6LU7-D08	Thr24 (12%), Thr25 (18%), His41 (15%), Gln192 (13%), Ser144 (60%), Gln189 (20%) Cyl45 (10%), Thr45 (8%), Asn142 (3%), Glu166 (4%), Thr190 (5%) and Thr26 (16%)	Leu27 (6%), His41 (56%), Met49 (40%), Cys145 (10%), Met165 (37%), Leu167 (5%), Pro168 (6%) and Ala191 (3%)	Thr24 (26%), Cys44 (4%), Ser46 (16%), Leu141 (3%), Asn142(5%), Cys143(18%), Gly166(13%), Gln189 (18%), Thr190, Ala191, and Gln192 (<5%)
6LU7-D23	His41 (6%), Asn142 (5%), Val186 (6%), Glu166 (25%), Asp187 (54%) and Gln189 (18%)	His41 (16%), Met49 (18%), Cyl45 (10%), Met165 (38%), Pro168 (6%) and Leu166 (5%)	Thr25 (4%), Thr16 (3%), His41 (4%), Asn142 (7%), His164 (8%), Leu166 (24%) and Gln189 (42%)
6LU7-D76	Asp (83%), His164 (42%), Glu166 (23%), Gln189 (13%) and Val186 (5%)	Met49 (5%), Met165 (22%), Leu167 (6%) and Pro168 (26%)	His41 (5%), His164 (42%), Glu166 (25%) and Gln189 (13%),
6LU7-T40	Thr16 (<.3%), His41(22%), Glu166(41%) and Gln189(18%)	His41 (20%), Met49 (12%), Met165 (40%), Leu167 (50%), Pro168 (52%) and Ala191 (<10%)	Thr26 (5%), Ser46(6%), Glu166 (15%), Asp187 (41%), Arg188(3%) and Gln189 (18%)

D08-6LU7 complex from 70 to 80 ns; these fluctuations did not exceed 1.5 Å after which D08-6LU7 returned to equilibrium until the end of the MD simulation time set at 100 ns. The average RMSD values obtained for 6LU7 free and complexed with the ligands **D08**, **D23**, **D76**, and **T40**, respectively, were 1.239 Å, 1.179 Å, 1.876 Å, 1.305 Å, 1.284 Å, and 1.400 Å. The low values of fluctuations observed in the RMSD indicate that the proposed drug molecules reached a good equilibrium in the 6LU6 protease pocket.

The mean RMSF values of the α -carbon residues of the examined systems 6LU7, 6LU7-D08, 6LU7-D23, 6LU7-D76, and 6LU7-T40, respectively, were 0.821 Å, 0.961 Å, 0.886 Å, 0.873 Å, and 0.788 Å (Fig. 10b). Despite the presence of some fluctuations related to extreme residue sequences (SER_47, LEU_50, ASN_51, ASP_155, THR_304, PHE_305, and GLN_306), the average values of these fluctuations did not exceed 4.5 Å, indicating the expected high stability of the examined ligands in the active pocket of 6LU7.

The RMSF data presented in Fig. 10c indicate the presence of some fluctuations in the structures of the **D08**, **D23**, **D76**, and **T40** ligands. The observed fluctuations could be due to some structural properties of the ligands, such as rotation angles, torsion, and flexible interactions between the ligands and their binding sites in the active 6LU7 pocket.

Dynamics of protein–ligand interactions (PL-contacts)

As previously mentioned in Fig. 7b, the active site in the 6LU7 protease pocket contains the polar amino acids threonine (Thr190, Thr24), glutamine (Gln189, Gln192), asparagine (Asn142) and aspartic (Asp187), serine (Ser144), and the negatively charged glutamic acid (Glu166), positively charged amino acids such as histidine (His163, His172, His41, His164) and nonpolar amino acids such as phenylalanine (Phe140), methionine (Met165, Met49), leucine (Leu167), proline (Pro168), alanine (Ala191), and the hydrophobic amino acid cysteine (Cys145). Interactions of the tested ligands with these residues and not moving away from the active pocket containing them can be considered a mechanism of inhibition of the enzymatic activity of 3CL^{PRO}. Figures 11 and 12 show synthetic diagrams of the key contacts that occur between the protein and the examined ligands during the MD simulations.

From Fig. 11 and 12, we can see that the examined ligands were able to maintain most interactions with the residues predicted by molecular docking, meaning that the small designed molecules were able to dock into the active 3CL^{PRO} pocket throughout the MD simulation and did not move away from the target active site. Table 6 presents a summary of the key protein–ligand contacts for the examined complexes 6LU7-D08, 6LU7-D23, 6LU7-D76, and 6LU7-T40, as well as the contact ratios identified at all 100 ns of MD simulations.

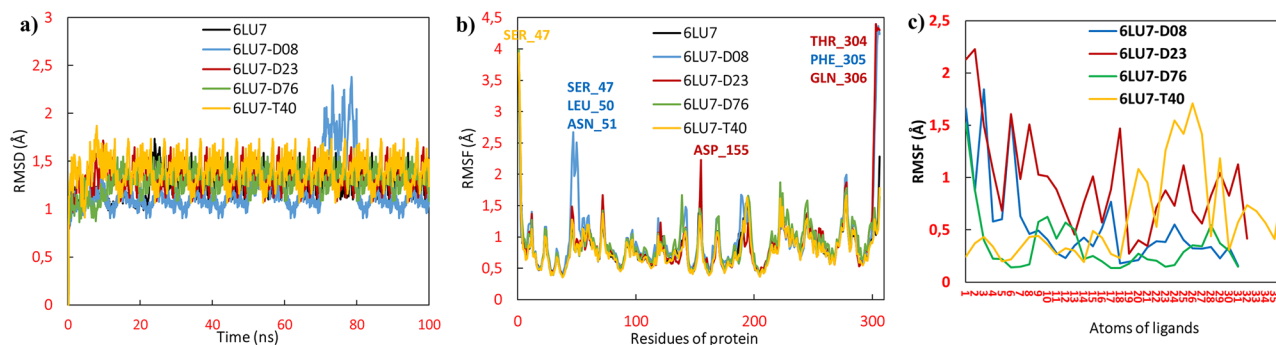


Fig. 10 **a** RMSD of free protease 6LU7, complexed with ligands **D08**, **D23**, **D76**, and **T40**. **b** RMSF of backbone atoms in free 6LU7, complexed with the ligands **D08**, **D23**, **D76**, and **T40**. **c** RMSF of ligands **D08**, **D23**, **D76**, and **T40** complexed with 6LU7.

From Table 6, we can notice that ligand **D8** was able to contact by conventional hydrogen bonds with the residues Thr24 (12%), Thr25 (18%), His41 (15%), Gln192 (13%), Ser144 (60%), Gln189 (20%) Cyl45 (10%), Thr45 (8%), Asn142 (3%), Glu166 (4%), and Thr190 (5%). Additionally, a hydrogen bond formed through a water bridge with the amino acid Thr26 (16%). **D08** is also contacted by hydrophobic interactions with Leu27 (6%), His41 (56%), Met49 (40%), Cys145 (10%), Met165 (37%), Leu167 (5%), Pro168 (6%), and Ala191 (3%).

For ligand **D23**, it was able to bind via hydrogen bonds to amino acid residues His41, Asn142 (ion bridge and water), and Val186 for periods less than 10%, as well as bind to

Glu166 (25%), Asp187 (54%), and Gln189 (18%). Hydrophobic interactions were also formed with His41 (16%), Met49 (18%), Cyl45 (10%), Met165 (38%), and with Pro168 and Leu166 (< 10%).

For ligand **D76**, there is a hydrogen bond contact with residues Asp (83%), His164 (42% water-bridge-aided), Glu166 (23%), Gln189 (13% water-bridge aided), and Val186 (5%). Additionally, hydrophobic interactions were formed with Met49 (5%), Met165 (22%), Leu167 (6%), and Pro168 (26%).

Ligand **T40** was contacted by hydrogen bonds with the amino acid residues Thr16 (< 3%), His41(22%), Glu166(41%), and Gln189(18%). In addition, hydrophobic interactions were formed between **T40** and the amino acid

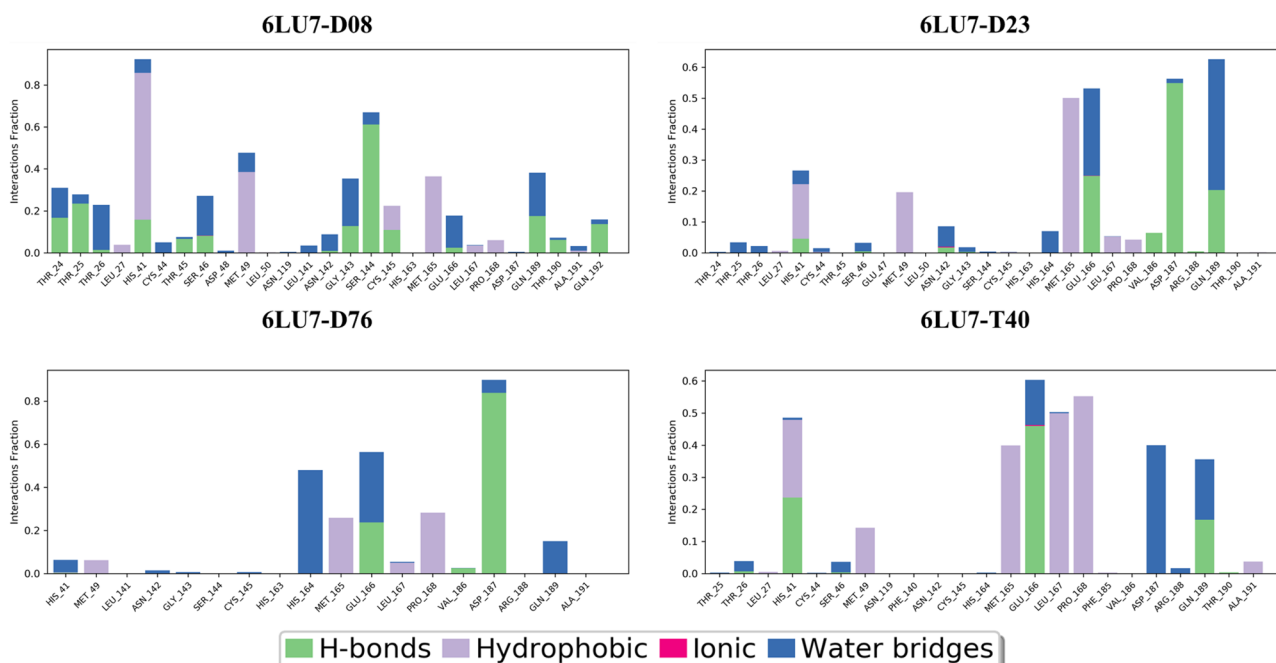


Fig. 11 Contact histogram of 6LU7-D08, 6LU7-D23, 6LU7-D76, and 6LU7-T40 along the MD time course

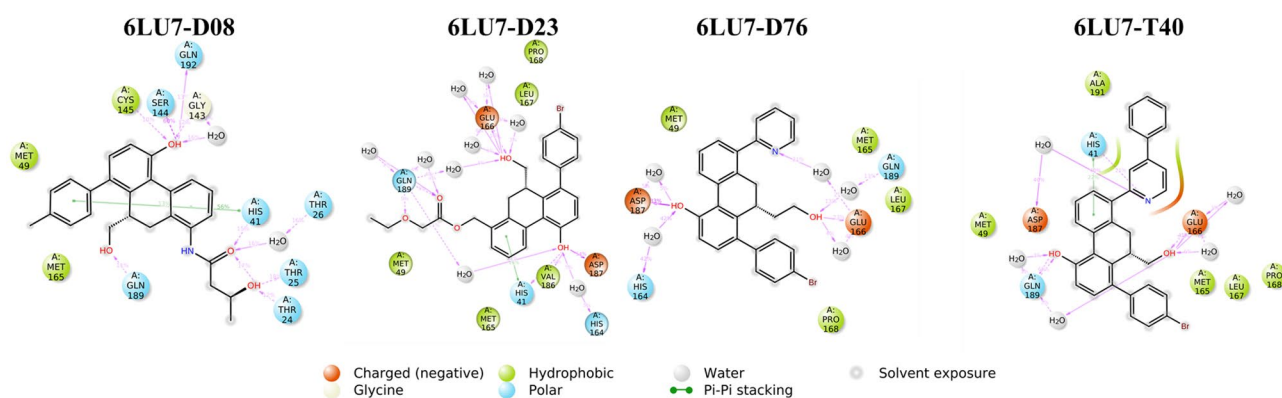


Fig. 12 2D visualization of summary contacts between 6LU7 and the ligands **D08**, **D23**, **D76**, and **T40** throughout the MD simulations time course

residues His41 (20%), Met49 (12%), Met165 (40%), Leu167 (50%), Pro168 (52%), and Ala191 (< 10%).

Moreover, from Table 6, we can also notice that the water bridge interactions contributed to the stability of ligands **D8**, **D23**, **D76**, and **T40** inside the 3CL^{pro} active pocket. The medium to low contact ratios of non-covalent interactions formed between ligands **D08**, **D23**, **D76**, and **T40** with active amino acid residues in the 6LU7 pocket can be explained by their high ΔG_{bind} (from -47.77 to -57.30 kcal/mol). This means that the structures of the proposed drug molecules have a flexible structure and good structural properties that allow them to make many contacts in the active pocket through different interactional modes.

From the summary of protein–ligand interactions presented in Figs. 11, 12, and Table 6, it can be concluded that weak H-bonds, hydrophobic bonds, and water bridges strongly contribute to the stability of the drug ligands (**D8**, **D23**, **D76**, and **T40**) with 3CL^{pro}. These weak non-covalent interactions are very appropriate for drug ligands to achieve protein binding compatibility and reach the desired therapeutic target, as well as to facilitate the removal of drug compounds after reaching the therapy. This is because strong covalent interactions between the ligand and the receptor are difficult to remove and can result in the opposite effect of covalent drug compounds. Therefore, non-covalent small molecule drugs may have very comparable therapeutic and pharmacokinetic activity against SARS-CoV-2 3CL^{pro} compared to covalent drug molecules such as covalent Michael inhibitor (**N3**).

Properties of ligands

Figure 13 presents the properties of **D08**, **D23**, **D76**, and **T40** ligands estimated at over 100 ns of MD simulation. A total of six properties of ligands were evaluated: ligand RMSD,

radius of gyration (rGyr), intramolecular hydrogen bonds (intra-HB), molecular surface area (MolSA), solvent accessible surface area (SASA), and polar surface area (PSA).

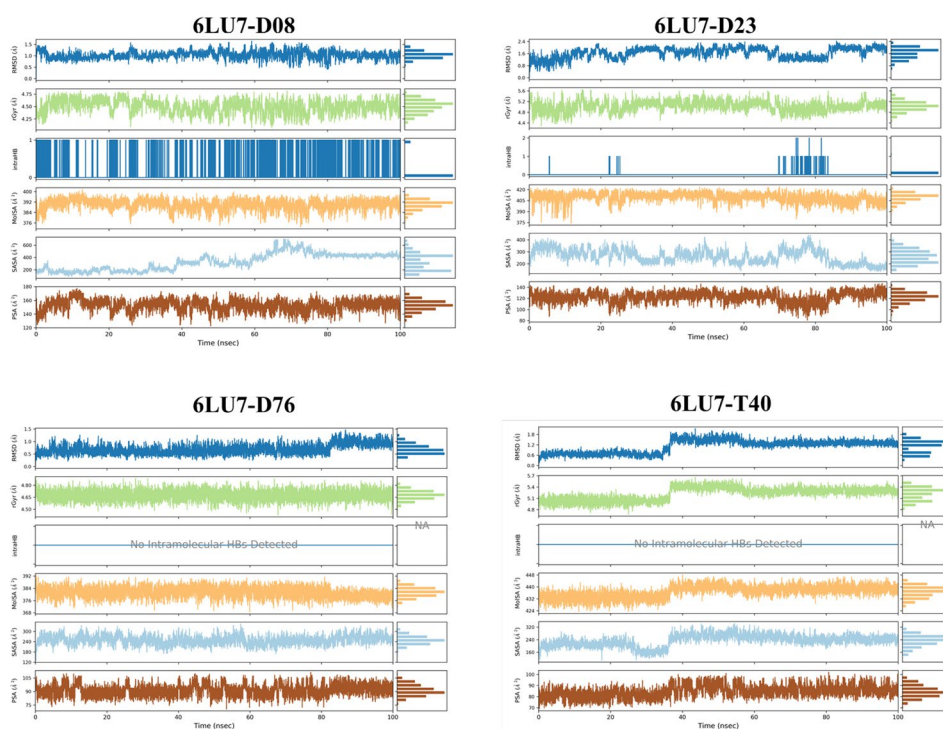
For ligand **D08** in 6LU7-D08, the RMSD ranged from 0.5 to 1.5 Å, and its equilibrium was approximately at 1 Å. rGyr was limited to the range (4.25–4.75 Å), and its equilibrium was around 4.5 Å. The intra-HB was high throughout the MD simulation. MolSA was in the range (376–400 Å²), and its equilibrium was about 390 Å². SASA showed strong fluctuations ranging from about (200–600 Å²) over the time interval (40–80 ns), then stabilized at around 400 Å². PSA was in the range (120–180 Å²), and its equilibrium was around 160 Å².

In the 6LU7-D23 complex, the ligand **D23** showed many fluctuations in RMSD ranging from 0.8–2.4 Å and stabilized at about 1.6 Å. rGyr ranged from 5.6 to 4.4 Å and stabilized at about 5 Å. Intra-HB was low throughout the MD simulation. MolSA was ranged between 375 and 420 Å² and stabilized perfectly along the simulation time at about 405 Å². SASA showed strong fluctuations around the range 200–400 Å² and later stabilized at around 200 Å². PSA was found limited in the range 80–140 Å² and stabilized at around 120 Å².

In the 6LU7-D76 complex, the ligand **D76** showed perfect stability in terms of RMSD in the range 0.5–1.2 Å over the simulation period with some slight fluctuations at about 80 ns. rGyr was in the range 4.8–4.5 Å and stabilized at about 4.65 Å. Intra-HB was not detected in the **D76** ligand. MolSA was in the range 376–392 Å² and stabilized at approximately 384 Å². The SASA was in the range 180–300 Å² and almost stabilized at 240 Å². PSA was in the range of 75–105 Å² and stabilized at about 90 Å².

In the 6LU7-T40 complex, the RMSD of the ligand **T40** was in the range 0.5–1.8 Å and stabilized at about 1.2 Å.

Fig. 13 Timeline of the properties of the ligands **D08**, **D23**, **D76**, and **T40** complexed with 6LU7 during 100 ns of MD trajectory



rGyr was in the range 4.8–5.7 Å and stabilized after 40 ns at about 5.4 Å. Intra-HB was not detected in the **T40** ligand. MolSA was in the range 424–448 Å² and stabilized at about

440 Å². SASA was in the range of about 160–320 Å²; it stabilizes at about 240 Å² after 40 ns of the MD trajectory. PSA was in the range 70–100 Å² and stabilized at about 80 Å².

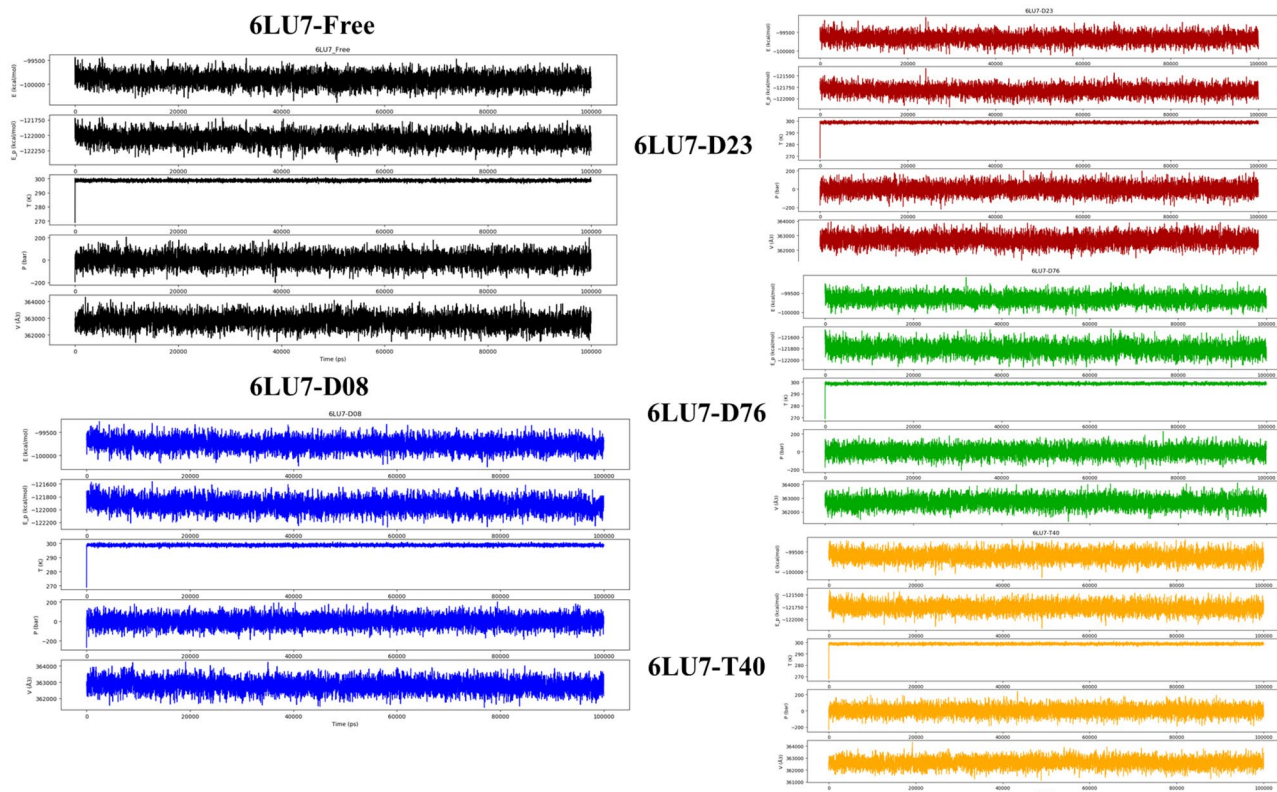


Fig. 14 The thermodynamic properties of the 6LU7 systems (6LU7 Free, 6LU7-D08, 6LU7-D23, 6LU7-D76, and 6LU7-T40)

Table 7 Scores average of the thermodynamics properties of analyzed systems

Systems	Average property values				
	Total energy (kcal/mol)	Potential energy (kcal/mol)	Temperature (K)	Pressure (bar)	Volume (\AA^3)
6LU7-Free	-99,905.370	-122,063.495	298.705	1.409	362,814.329
6LU7-D08	-99,756.824	-121,927.605	298.714	1.565	362,806.668
6LU7-D23	-99,661.259	-121,823.752	298.702	1.193	362,702.992
6LU7-D76	-99,649.019	-121,807.582	298.709	1.247	362,706.347
6LU7-T40	-99,603.363	-121,760.057	298.704	1.885	362,615.184

Thermodynamic properties analysis

The summary of the quality of the MD simulation is analyzed through the computation of the stability profile for the investigated systems in terms of the variation in total energy (E), potential energy (E_P), temperature (T), pressure (P), and volume (V) over the 100 ns of MD simulation trajectory. Figure 14 shows the thermodynamic properties diagrams generated for the 6LU7, 6LU7-D08, 6LU7-D23, 6LU7-D76, and 6LU7-T40 systems.

The generated thermodynamic properties show that the scores (E, EP, T, P, and V) of the complexes (6LU7-D08, 6LU7-D23, 6LU7-D76, and 6LU7-T40) remained stable and close to those of free 6LU7. This finding can be confirmed by the average values (E, EP, T, P, and V) of the 6LU7 complexes that are very close to those of free 6LU7 (Table 7). These results further prove that the structures of small 9,10-dihydrophenanthrene compounds can reach perfect stability in the active pocket of 3CL^{PRO}.

All in all, molecular dynamics analyses show the high stability of the samples (**D08**, **D23**, **D76**, and **T40**) inside the 3CL^{PRO} active pocket. Therefore, the choice of the nine molecules (**D07**, **D08**, **D12**, **D18**, **D23**, **D25**, **D26**, **D27**, and **D76**) screened through 3D-QSAR, molecular docking, drug-like, ADMET, and MM-GBSA studies can be validated as promising non-covalent inhibitors of SARS-CoV 3CL^{PRO}.

Conclusion

The outbreak of coronavirus 2019 (COVID-19) has negatively impacted daily life in all regions of the world. Due to the severity of COVID-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), it is necessary to advance the search for an appropriate and effective drug against COVID-19. In this context, 3-chymotrypsin-like cysteine protease (3CL^{PRO}) is an indispensable input to viral replication. Therefore, inhibition of 3CL^{PRO} enzymatic activity becomes an attractive target against COVID-19. To search for novel pharmacological compounds against COVID-19 through computer-aided drug design, we used a series of 46

small molecules derived from 9,10-dihydrophenanthrene as potential inhibitors of SARS-CoV-2 3CL^{PRO} in the computational analysis and pharmaceutical parameters screening. In this study, the 3D quantitative structure–activity relationship (3D-QSAR) for 9,10-dihydrophenanthrene derivatives was carefully analyzed and described using CoMFA and CoMSIA techniques. As a result, based on the structure of the 9,10-dihydrophenanthrenes derivatives, two models CoMFA/SE and CoMSIA/SEHDA were developed. Both models showed a high ability to predict the biological activity of pIC₅₀ against SARS-CoV-2 3CL^{PRO}, the pharmacological sites were rationalized, and the most important features favorable for modeling and improving the biological activity of the studied molecules were identified. Accordingly, 96 new drug molecules were generated based on the structure of the synthesized template molecule T40 exhibiting the highest biological activity pIC₅₀ observed in vitro, followed by bioavailability parameter screening to select candidate drug compounds.

Then, the bioactivity (pIC₅₀) of the modeled molecules was predicted by 3D-QSAR models, their non-covalent interaction to 3CL^{PRO} (PDB code:6LU7) was investigated via molecular docking, and in silico pharmacokinetics, ADME properties, and toxicity were evaluated, as well as free binding energies (ΔG_{bind}) were scored by MM-GBSA computations. The results of this study demonstrated that the generated nine compounds generated D07, D08, D12, D18, D23, D25, D26, D27, and D76 have high biological inhibitory activity (pIC₅₀), excellent non-covalent binding to 3CL^{PRO}, good pharmacokinetic suitability and less potential toxicity compared to the template synthesized compound T40 and N3 peptidic inhibitor. The results obtained were confirmed by molecular dynamics simulations of the tested systems (6LU7 uncomplexed and complex). For this purpose, the structural stability and dynamics of free and complexed with the tested ligands (D08, D23, D76, and T40) in an aqueous environment were discussed.

Finally, we have shown that nine small molecules modeled on 9,10-dihydrophenanthrene structures have the potential to act as a promising non-covalent drug candidate against COVID-19 by inhibiting the enzymatic activity of 3CL^{PRO}.

Therefore, the adoption of the small molecule structures proposed in this study will be useful as a key starting point for the development of therapy against COVID-19. Thus, the retrosynthesis of these molecules and the evaluation of their bioactivity in vitro and in vivo may be of interest in the context of SARS-CoV-2 3CL^{pro} drug design and discovery. Also, the potential activity of the proposed small molecules against other protein pathways of coronaviruses can be investigated.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11224-022-02004-z>.

Author contribution Ossama Daoui: conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, writing-original draft, reviewing, and editing. Souad Elkhattabi: data curation, original draft preparation, conceptualization, supervision, review, and editing. Samir Chtita: visualization, original draft preparation, software, conceptualization, review, and editing.

Data availability All data used in this work are private.

Code availability The codes used in this work are not available.

Declarations

Conflicts of interest The authors declare no competing interests.

References

- Fouedjou RT, Chtita S, Bakhouch M et al (2021) Cameroonian medicinal plants as potential candidates of SARS-CoV-2 inhibitors. *J Biomol Struct Dyn* 0:1–15. <https://doi.org/10.1080/07391102.2021.1914170>
- Zhang J-W, Xiong Y, Wang F et al (2022) Discovery of 9,10-dihydrophenanthrene derivatives as SARS-CoV-2 3CL^{pro} inhibitors for treating COVID-19. *Eur J Med Chem* 228:114030. <https://doi.org/10.1016/j.ejmech.2021.114030>
- Wu T, Kang S, Peng W et al (2021) Original hosts, clinical features, transmission routes, and vaccine development for coronavirus disease (COVID-19). *Front Med* 8
- Cihan P (2021) Forecasting fully vaccinated people against COVID-19 and examining future vaccination rate for herd immunity in the US, Asia, Europe, Africa, South America, and the World. *Appl Soft Comput* 111:107708. <https://doi.org/10.1016/j.asoc.2021.107708>
- Nasreen S, He S, Chung H et al (2021) Effectiveness of COVID-19 vaccines against variants of concern, Canada. *Medrxiv*
- He X, He C, Hong W et al (2021) The challenges of COVID-19 Delta variant: prevention and vaccine development. *MedComm* 2:846–854
- SARS-CoV-2 Evolution. <https://www.who.int/news-room/questions-and-answers/item/sars-cov-2-evolution>. Accessed 8 Apr 2022
- Liu C, Ginn HM, Dejnirattisai W et al (2021) Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell* 184:4220–4236.e13. <https://doi.org/10.1016/j.cell.2021.06.020>
- Singh J, Rahman SA, Ehtesham NZ et al (2021) SARS-CoV-2 variants of concern are emerging in India. *Nat Med* 27:1131–1133. <https://doi.org/10.1038/s41591-021-01397-4>
- Pillaiyar T, Meenakshisundaram S, Manickam M (2020) Recent discovery and development of inhibitors targeting coronaviruses. *Drug Discovery Today* 25:668–688. <https://doi.org/10.1016/j.drudis.2020.01.015>
- Kumar V, Kar S, De P et al (2022) Identification of potential anti-virals against 3CL^{pro} enzyme for the treatment of SARS-CoV-2: a multi-step virtual screening study. *SAR QSAR Environ Res* 0:1–30. <https://doi.org/10.1080/1062936X.2022.2055140>
- De P, Bhayye S, Kumar V, Roy K (2022) In silico modeling for quick prediction of inhibitory activity against 3CL^{pro} enzyme in SARS CoV diseases. *J Biomol Struct Dyn* 40:1010–1036. <https://doi.org/10.1080/07391102.2020.1821779>
- De P, Kumar V, Kar S et al (2022) Repurposing FDA approved drugs as possible anti-SARS-CoV-2 medications using ligand-based computational approaches: sum of ranking difference-based model selection. *Struct Chem*. <https://doi.org/10.1007/s11224-022-01975-3>
- Amin SA, Banerjee S, Ghosh K et al (2021) Protease targeted COVID-19 drug discovery and its challenges: insight into viral main protease (M^{pro}) and papain-like protease (PL^{pro}) inhibitors. *Bioorg Med Chem* 29:115860
- Chan JF-W, Yip CC-Y, To KK-W et al (2020) Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-PCR assay validated in vitro and with clinical specimens. *J Clin Microbiol* 58:e00310-e320
- Wang L, Xiang Y (2020) Spike glycoprotein-mediated entry of SARS coronaviruses. *Viruses* 12:1289
- Jankun J (2020) COVID-19 pandemic; transmembrane protease serine 2 (TM^{PRSS2}) inhibitors as potential drugs. *Translation: The University of Toledo Journal of Medical Sciences* 7:1–5
- Ni W, Yang X, Yang D et al (2020) Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. *Crit Care* 24:1–10
- Rossi GP, Sanga V, Barton M (2020) Potential harmful effects of discontinuing ACE-inhibitors and ARBs in COVID-19 patients. *Elife* 9:e57278
- Anand K, Ziebuhr J, Wadhvani P et al (2003) Coronavirus main proteinase (3CL^{pro}) structure: basis for design of anti-SARS drugs. *Science* 300:1763–1767. <https://doi.org/10.1126/science.1085658>
- Cui W, Yang K, Yang H (2020) Recent progress in the drug development targeting SARS-CoV-2 main protease as treatment for COVID-19. *Front Mol Biosci* 7
- Khan SA, Zia K, Ashraf S et al (2021) Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *J Biomol Struct Dyn* 39:2607–2616. <https://doi.org/10.1080/07391102.2020.1751298>
- Deshmukh MG, Ippolito JA, Zhang C-H et al (2021) Structure-guided design of a perampanel-derived pharmacophore targeting the SARS-CoV-2 main protease. *Structure* 29:823–833.e5. <https://doi.org/10.1016/j.str.2021.06.002>
- Liu S, Zheng Q, Wang Z (2020) Potential covalent drugs targeting the main protease of the SARS-CoV-2 coronavirus. *Bioinformatics* 36:3295–3298
- Delre P, Caporuscio F, Saviano M, Mangiatordi GF (2020) Repurposing known drugs as covalent and non-covalent inhibitors of the SARS-CoV-2 papain-like protease. *Front Chem* 8:1032
- Osipiuk J, Azizi S-A, Dvorkin S et al (2021) Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nat Commun* 12:1–9
- Halford B (2021) Pfizer unveils its oral SARS-CoV-2 inhibitor. *Chem Eng News* 7–7
- Boras B, Jones RM, Anson BJ et al (2021) Discovery of a novel inhibitor of coronavirus 3CL protease for the potential treatment of COVID-19. 2020.09.12.293498

29. Jin Z, Du X, Xu Y et al (2020) Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582:289–293. <https://doi.org/10.1038/s41586-020-2223-y>
30. Zhang L, Lin D, Sun X et al (2020) Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* 368:409–412. <https://doi.org/10.1126/science.abb3405>
31. Qiao J, Li Y-S, Zeng R et al (2021) SARS-CoV-2 Mpro inhibitors with antiviral activity in a transgenic mouse model. *Science* 371:1374–1378. <https://doi.org/10.1126/science.abf1611>
32. Ma C, Sacco MD, Hurst B et al (2020) Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res* 30:678–692. <https://doi.org/10.1038/s41422-020-0356-z>
33. Xiong Y, Zhu G-H, Wang H-N et al (2021) Discovery of naturally occurring inhibitors against SARS-CoV-2 3CLpro from Ginkgo biloba leaves via large-scale screening. *Fitoterapia* 152:104909. <https://doi.org/10.1016/j.fitote.2021.104909>
34. Kitamura N, Sacco MD, Ma C et al (2022) Expedited approach toward the rational design of noncovalent SARS-CoV-2 main protease inhibitors. *J Med Chem* 65:2848–2865. <https://doi.org/10.1021/acs.jmedchem.1c00509>
35. Han SH, Goins CM, Arya T et al (2022) Structure-based optimization of ML300-derived, noncovalent inhibitors targeting the severe acute respiratory syndrome coronavirus 3CL protease (SARS-CoV-2 3CLpro). *J Med Chem* 65:2880–2904. <https://doi.org/10.1021/acs.jmedchem.1c00598>
36. Zhang C-H, Stone EA, Deshmukh M et al (2021) Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. *ACS Cent Sci* 7:467–475. <https://doi.org/10.1021/acscentsci.1c00039>
37. Balupuri A, Balasubramanian PK, Cho SJ (2020) 3D-QSAR, docking, molecular dynamics simulation and free energy calculation studies of some pyrimidine derivatives as novel JAK3 inhibitors. *Arab J Chem* 13:1052–1078
38. De Vita E (2021) 10 years into the resurgence of covalent drugs. *Future Med Chem* 13:193–210. <https://doi.org/10.4155/fmc-2020-0236>
39. Erak M, Bellmann-Sickert K, Els-Heindl S, Beck-Sickinger AG (2018) Peptide chemistry toolbox – transforming natural peptides into peptide therapeutics. *Bioorg Med Chem* 26:2759–2765. <https://doi.org/10.1016/j.bmc.2018.01.012>
40. Daoui O, Mazoir N, Bakhouch M et al (2022) 3D-QSAR, ADME-Tox, and molecular docking of semisynthetic triterpene derivatives as antibacterial and insecticide agents. *Struct Chem*. <https://doi.org/10.1007/s11224-022-01912-4>
41. Daoui O, Elkhatabi S, Chtita S et al (2021) QSAR, molecular docking and ADMET properties in silico studies of novel 4,5,6,7-tetrahydrobenzo[D]-thiazol-2-Y1 derivatives derived from dimedone as potent anti-tumor agents through inhibition of C-Met receptor tyrosine kinase. *Heliyon* 7:e07463. <https://doi.org/10.1016/j.heliyon.2021.e07463>
42. Aanouz I, Belhassan A, El-Khatabi K et al (2021) Moroccan medicinal plants as inhibitors against SARS-CoV-2 main protease: computational investigations. *J Biomol Struct Dyn* 39:2971–2979. <https://doi.org/10.1080/07391102.2020.1758790>
43. PhD SAL MD (2021) DISulfiram for COvid-19 (DISCO) Trial: a phase 2 double-blind, randomized placebo-controlled trial of disulfiram compared to standard care in patients with symptomatic COVID-19. clinicaltrials.gov
44. Lee SA, Elliott JH, McMahon J et al (2019) Population pharmacokinetics and pharmacodynamics of disulfiram on inducing latent HIV-1 transcription in a phase IIb trial. *Clin Pharmacol Ther* 105:692–702. <https://doi.org/10.1002/cpt.1220>
45. Sybyl-X 2.0, Tripos International, St. Louis, Missouri, 63144, USA. In: Software Informer. <https://sybyl-x.software.informer.com/2.0/>. Accessed 17 Jul 2021
46. Zhao X, Chen M, Huang B et al (2011) Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) studies on α 1A-adrenergic receptor antagonists based on pharmacophore molecular alignment. *Int J Mol Sci* 12:7022–7037
47. Daoui O, Elkhatabi S, Chtita S (2022) Design and prediction ADME/Tox properties of novel magnolol derivatives as anticancer agents for NSCLC using 3D-QSAR, molecular docking, MOLCAD and MM-GBSA studies. *Lett Drug Des Discov* 19:1–1. <https://dx.doi.org/10.2174/1570180819666220510141710>
48. Golbraikh A, Tropsha A (2000) Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *Mol Divers* 5:231–243. <https://doi.org/10.1023/A:1021372108686>
49. Tropsha A (2010) Best practices for QSAR model development, validation, and exploitation. *Mol Inf* 29:476–488
50. Golbraikh A, Tropsha A (2002) Beware of q²! *J Mol Graph Model* 20:269–276. [https://doi.org/10.1016/S1093-3263\(01\)00123-1](https://doi.org/10.1016/S1093-3263(01)00123-1)
51. Rücker C, Rücker G, Meringer M (2007) y-Randomization and its variants in QSPR/QSAR. *J Chem Inf Model* 47:2345–2357. <https://doi.org/10.1021/ci700157b>
52. Chalkha M, Akhazzane M, Moussaid FZ et al (2021) Design, synthesis, characterization, in vitro screening, molecular docking, 3D-QSAR, and ADME-Tox investigations of novel pyrazole derivatives as antimicrobial agents. *New J Chem*. <https://doi.org/10.1039/D1NJ05621B>
53. Salama I, Abdel-Fattah MAO, Hany MS et al (2012) CoMFA and CoMSIA studies of 1,2-dihydropyridine derivatives as anticancer agents. *Med Chem* 8:372–383
54. Skalic M, Jiménez J, Sabbadin D, De Fabritiis G (2019) Shape-based generative modeling for de novo drug design. *J Chem Inf Model* 59:1205–1214. <https://doi.org/10.1021/acs.jcim.8b00706>
55. Pires DEV, Blundell TL, Ascher DB (2015) pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem* 58:4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104>
56. Molecular Properties Prediction - Osiris Property Explorer. <https://www.organic-chemistry.org/prog/peo/>. Accessed 27 Mar 2022
57. Bank RPD RCSB PDB - 6LU7: The crystal structure of COVID-19 main protease in complex with an inhibitor N3. <https://www.rcsb.org/structure/6LU7>. Accessed 11 Apr 2022
58. M.J.E.A. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, H. Nakatsuji, et al., Gaussian 09. <https://gaussian.com/g09citation/>. Accessed 23 Jan 2022
59. Mohapatra RK, El-ajaily MM, Alassbaly FS et al (2021) DFT, anticancer, antioxidant and molecular docking investigations of some ternary Ni(II) complexes with 2-[(E)-[4-(dimethylamino)phenyl]methyleneamino]phenol. *Chem Pap* 75:1005–1019. <https://doi.org/10.1007/s11696-020-01342-8>
60. Morris GM, Huey R, Lindstrom W et al (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30:2785–2791
61. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461
62. Dassault Systèmes BIOVIA Discovery Studio Modeling Environment. Release (2017) Dassault Systèmes 2016. <https://discover.3ds.com/discovery-studio-visualizer-download>. Accessed 7 May 2021
63. Choudhary MI, Shaikh M, tul-Wahab A, ur-Rahman A (2020) In silico identification of potential inhibitors of key SARS-CoV-2 3CL hydrolase (Mpro) via molecular docking, MMGBSA

- predictive binding energy calculations, and molecular dynamics simulation. PLoS ONE 15:e0235030
64. Alnajjar R, Mostafa A, Kandeil A, Al-Karmalawy AA (2020) Molecular docking, molecular dynamics, and in vitro studies reveal the potential of angiotensin II receptor blockers to inhibit the COVID-19 main protease. *Heliyon* 6:e05641. <https://doi.org/10.1016/j.heliyon.2020.e05641>
 65. Bakal RL, Jawarkar RD, Manwar JV et al (2022) Identification of potent aldose reductase inhibitors as antidiabetic (anti-hyperglycemic) agents using QSAR based virtual screening, molecular docking, MD simulation and MMGBSA approaches. *Saudi Pharmaceut J*
 66. Abchir O, Daoui O, Belaidi S et al (2022) Design of novel benzimidazole derivatives as potential α -amylase inhibitors using QSAR, pharmacokinetics, molecular docking, and molecular dynamics simulation studies. *J Mol Model* 28:106. <https://doi.org/10.1007/s00894-022-05097-9>
 67. Al-Jumaili MHA, Siddique F, Abul Qais F et al (2021) Analysis and prediction pathways of natural products and their cytotoxicity against HeLa cell line protein using docking, molecular dynamics and ADMET. *J Biomol Struct Dyn* 1–13
 68. Schrödinger Release 2020–3 (2020) Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2020. Maestro-Desmond Interoperability Tools. Schrödinger. https://scholar.google.com/scholar_lookup?hl=en&publication_year=2020&author=Schr%C3%B6dinger+Release+2020-3&title=Desmond+Molecular+Dynamics+System%2C+D.+E.+Shaw+Research%2C+New+York%2C+NY%2C+2020.+Maestro-Desmond+Interoperability+Tools. Accessed 9 Jan 2022
 69. Zielkiewicz J (2005) Structural properties of water: comparison of the SPC, SPCE, TIP4P, and TIP5P models of water. *J Chem Phys* 123:104501. <https://doi.org/10.1063/1.2018637>
 70. Kumar BK, Faheem SKVGC et al (2022) Pharmacophore based virtual screening, molecular docking, molecular dynamics and MM-GBSA approach for identification of prospective SARS-CoV-2 inhibitor from natural product databases. *J Biomol Struct Dyn* 40:1363–1386. <https://doi.org/10.1080/07391102.2020.1824814>
 71. Cheatham TEI, Miller JL, Fox T et al (2002) Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. In: ACS Publications. <https://doi.org/10.1021/ja00119a045>. Accessed 15 Feb 2022
 72. Cheng A, Merz KM (1996) Application of the Nosé–Hoover chain algorithm to the study of protein dynamics. *J Phys Chem* 100:1927–1937. <https://doi.org/10.1021/jp951968y>
 73. Guo Z, Mohanty U, Noehre J et al (2010) Probing the α -helical structural stability of stapled p53 peptides: molecular dynamics simulations and analysis. *Chem Biol Drug Des* 75:348–359
 74. kumar BH, Manandhar S, Mehta CH et al (2022) Structure-based docking, pharmacokinetic evaluation, and molecular dynamics-guided evaluation of traditional formulation against SARS-CoV-2 spike protein receptor bind domain and ACE2 receptor complex. *Chem Pap* 76:1063–1083. <https://doi.org/10.1007/s11696-021-01917-z>
 75. Umar HI, Ajayi A, Bello RO et al (2022) Novel molecules derived from 3-O-(6-galloylglucoside) inhibit main protease of SARS-CoV 2 in silico. *Chem Pap* 76:785–796. <https://doi.org/10.1007/s11696-021-01899-y>
 76. Lipinski CA (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 1:337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>
 77. Martin YC (2005) A bioavailability score. *J Med Chem* 48:3164–3170. <https://doi.org/10.1021/jm0492002>
 78. Fukunishi Y, Kurosawa T, Mikami Y, Nakamura H (2014) Prediction of synthetic accessibility based on commercially available compound databases. *J Chem Inf Model* 54:3259–3267. <https://doi.org/10.1021/ci500568d>
 79. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem* 1:55–68
 80. Belhassan A, Chtita S, Zaki H et al (2022) In silico detection of potential inhibitors from vitamins and their derivatives compounds against SARS-CoV-2 main protease by using molecular docking, molecular dynamic simulation and ADMET profiling. *J Mol Struct* 1258:132652. <https://doi.org/10.1016/j.molstruc.2022.132652>
 81. Chtita S, Fouedjou RT, Belaidi S et al (2022) In silico investigation of phytoconstituents from Cameroonian medicinal plants towards COVID-19 treatment. *Struct Chem*. <https://doi.org/10.1007/s11224-022-01939-7>

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