Journal of proteome research

pubs.acs.org/jpr

Article

Theoretical Insights into the Anti-SARS-CoV-2 Activity of Chloroquine and Its Analogs and In Silico Screening of Main Protease Inhibitors

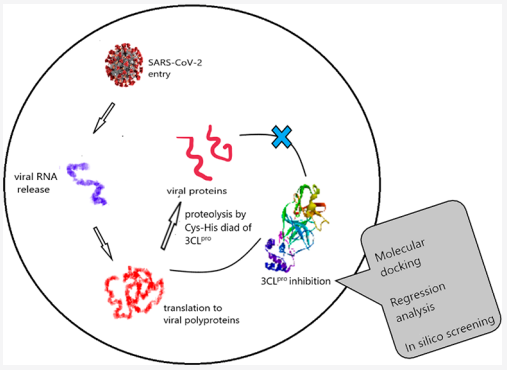
A. S. Achutha, V. L. Pushpa,* and Surendran Suchitra

Cite This: *J. Proteome Res.* 2020, 19, 4706–4717

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Corona virus disease (COVID-19) is a dangerous disease rapidly spreading all over the world today. Currently there are no treatment options for it. Drug repurposing studies explored the potency of antimalarial drugs, chloroquine and hydroxychloroquine, against SARS-CoV-2 virus. These drugs can inhibit the viral protease, called chymotrypsin-like cysteine protease, also known as Main protease ($3CL^{pro}$); hence, we studied the binding efficiencies of 4-aminoquinoline and 8-aminoquinoline analogs of chloroquine. Six compounds furnished better binding energies than chloroquine and hydroxychloroquine. The interactions with the active site residues especially with Cys145 and His41, which are involved in catalytic diad for proteolysis, make these compounds potent main protease inhibitors. A regression model correlating binding energy and the molecular descriptors for chloroquine analogs was generated with $R^2 = 0.9039$ and $Q^2 = 0.8848$. This model was used to screen new analogs of primaquine and molecules from the Asinex compound library. The docking and regression analysis showed these analogs to be more potent inhibitors of $3CL^{pro}$ than hydroxychloroquine and primaquine. The molecular dynamic simulations of the hits were carried out to determine the binding stabilities. Finally, we propose four compounds that show drug likeness toward SARS-CoV-2 that can be further validated through in vitro and in vivo studies.



KEYWORDS: *SARS-CoV-2, COVID-19, $3CL^{pro}$, chloroquine, hydroxychloroquine, molecular docking, regression, antiviral screening, molecular dynamics*

1. INTRODUCTION

In December 2019, patients were admitted in hospitals in Wuhan, China due to pneumonia with fever and other symptoms.[1,2] In January 2020, it was discovered that a novel coronavirus, which is similar to SARS-CoV and MERS-CoV genomes, is responsible for this disease.[2] This virus is named as Severe Acute Respiratory Syndrome Corona Virus-2 (SARS-CoV-2) and the disease is called Corona Virus Disease-2019 (COVID-19).[3,4] It was declared as pandemic on March 11 by the World Health Organization (WHO).[5] As of September 1, 2020, there are 25,327,098 confirmed cases with 848,255 deaths worldwide according to WHO reports.[6]

Coronavirus is a single-stranded positive sense RNA with a membrane envelope composed of four types of viruses: α, β, γ, and δ coronaviruses. SARS-CoV-2 belongs to β-coronavirus family.[5,7] Its genome encodes for structural proteins mainly spike glycoprotein (S), envelope glycoprotein (E), membrane glycoprotein (M), nucleocapsid protein (N), and nonstructural proteins consisting of chymotrypsin-like cysteine protease also known as Main protease ($3CL^{pro}$) and papain-like protease (PL^{pro}).[1,8,9] The viral RNA is released to the host cell, which is then translated to polyproteins inside the cell. These viral polyproteins are cleaved to functional proteins by proteolysis, which is catalyzed by $3CL^{pro}$ and PL^{pro} [1,8,9]. Since the main protease or $3CL^{pro}$ cleaves at 11 sites of the polyprotein to produce smaller proteins for viral replication and because of its nonsimilarity with human proteins, $3CL^{pro}$ is a potential target in anti-COVID-19 drug design.[8,10]

The main protease of SARS-CoV-2 has a Cys-His catalytic diad involving Cys145 and His41 and has four conserved binding subsites S1, S2, S4, and S1′. Main chains of Cys145, Phe140, and Leu141 and side chains of Asn142, Glu166, His163, and His172 residues are present in the S1 subsite. The side chains of His41, Met49, Tyr54, and Met165, along with the alkyl side chain of Asp187 forms the S2 subsite. The S4 subsite consists of Met165, Leu167, Phe185, and Gln192 side chain residues and the main chain of Gln189. S1′ includes Thr24 and Thr25. Therefore, a molecule to show $3CL^{pro}$

Special Issue: Proteomics in Pandemic Disease

Received: September 2, 2020
Published: September 22, 2020

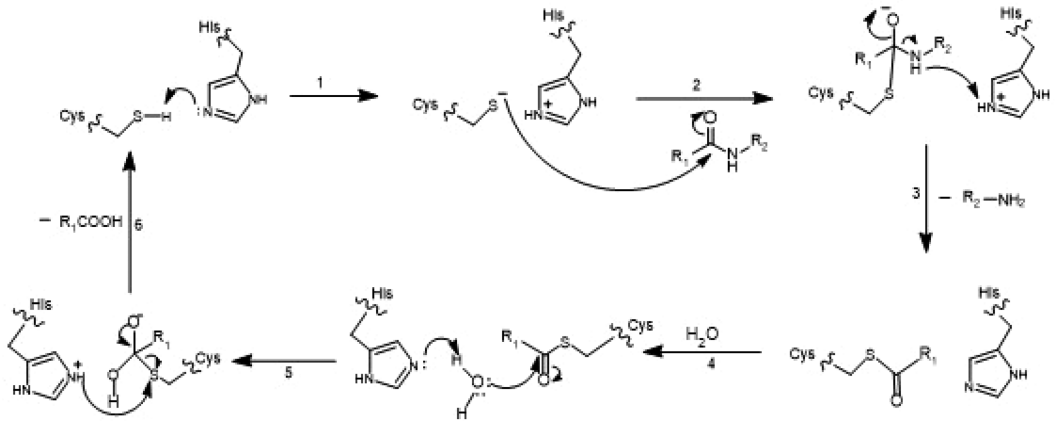


Figure 1. Mechanism of proteolysis by Cys-His diad by hydrolysis of amide substrate. (1) Deprotonation from cysteine by histidine. (2) Nucleophilic attack on the histidine. (3) Release of an amine resulting in the formation of a thioester deprotonation of histidine. (4) Addition of water. (5) Thioester is hydrolyzed and cysteine−substrate bond is broken. (6) Regeneration of enzyme with the elimination of a carboxylic acid molecule.[9,11]

inhibitory property should interact with one or more residues among these amino acids. Also, the inhibitory action can be assumed to enhance if they interact with Cys145 or His41, which is involved in the catalytic diad of Main protease[10] (Figure 1). Cysteine acts as a nucleophile and Histidine as an acid/base catalyst in this catalytic reaction.[9,11]

Currently there are no specific approved treatment options for COVID-19. Researches are in search of vaccine development, druggable small molecule, monoclonal antibodies and cell-based therapies.[12] Several drug repurposing studies have been conducted so far, and many of them gave positive results. The 4-aminoquinolines, chloroquine (CQ), and its hydroxyl analog hydroxychloroquine (HCQ) have been found to disrupt the viral replication and infectivity.[13−15] CQ [4-*N*-(7-chloroquinolin-4-yl)-1-*N*,1-*N*-diethylpentane-1,4-diamine] and HCQ (2-[4-[(7-chloroquinolin-4-yl)amino]pentyl-ethylamino]ethanol) are used particularly as antimalarial drugs. Their broad spectrum of activity as antibacterial, antiviral, and antifungal infections has also been reported.[15,16] HCQ has been found out to be less toxic than CQ, and it is currently being used to treat COVID-19 patients.[17]

The present study deals with theoretical perspectives on the anti-SARS-CoV-2 activity of CQ and HCQ analogs by observing their main protease inhibitory property. Molecular docking studies with the $3CL^{pro}$ protein were performed to analyze the drug likeness as well as to correlate the binding energy of the docked complex with various physicochemical properties of the active molecules, which will aid in the design of new anti-COVID-19 medicatives.

2. METHODOLOGY

2.1. Computational Details

Analogs of CQ and HCQ as well as primaquine derivatives were obtained through literature search, from the database of small molecules Pubchem as well as created manually.[18−22] All the molecules were drawn in Marvin Sketch and minimized their energy using Avogadro software.[23,24] The structure of $3CL^{pro}$ protein having PDB ID 6LU7 was retrieved from RCSB Protein Data Bank.[10,25] Characterization studies of the protein were done by using Expasy-protparam, an online server, which provides all the details regarding the protein.

2.2. Molecular Docking

The energy minimization of the retrieved protein was done by employing Swiss PBD Viewer.[26,27] Molecular docking was done using Autodock 4.2.6.[28] Both protein and ligands were prepared in pdbqt format. Polar hydrogens and Gasteiger charges were added to the receptor. Grid was prepared with 60 × 60 × 60 $Å^3$ with spacing 0.375 Å and centered at −18.617, 13.208, and 60.061 Å along *x*, *y*, and *z* axes, respectively. Grids were prepared for every atom present in the ligand data set. The genetic algorithm was employed as search parameter with 50 runs, 300 and 27,000 population size and number of generations, respectively.

2.3. Regression Analysis

A multiple linear regression (MLR) model was generated to find the correlation with binding energy obtained from docking and the physicochemical properties using QSARINS.[29,30] The molecular descriptors for the ligands were calculated utilizing PaDEL-Descriptor.[31] To build the model, 80% of the data set were randomly divided as training set and test set in 80:20 ratio. Thirty-five compounds (28 training and 7 test) including CQ and HCQ were employed for modeling purpose (Supplementary Tables S2 and S3). The models were generated using training set with MLR analysis of 1−5 variables. The generated models were examined by internal validation and external validation. Internal validation was done by cross-validated leave-one-out (LOO) method, which involves iteratively leaving one compound from the training set and generating regression model with the remaining molecules and predicting the value of response for the excluded one. For a good model, the regression coefficient (R^2) as well as the cross validated R^2, that is, Q^2 has to be close and >0.6.[29,32] The model was validated using test set, which was not included in model generation.[29,33] The best model was used to predict the binding energy of novel set of compounds.

2.4. Molecular Dynamic Simulations

All the MD simulations were done using the NAMD package (Version 2.14) developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana−Champaign[34] with the CHARMM 36 force field.[35,36] The docked structures were used as the initial coordinates for MD simulations. TIP3P water box was used to

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

solvate the complex with 5 Å padding. The solvated system was minimized for 100 ps with a time step of 2 fs. The simulations were done with Particle mesh Ewald electrostatics and periodic boundary conditions. Atom-based cutoff of 12 Å was applied for nonbonding interactions. The temperature and pressure were kept constant at 310 K and 1 atm, respectively, using Langevin dynamics. The system was equilibrated for 100 ps followed by a production simulation for 20 ns using the NPT ensemble. All the preparation and analysis steps were done using the VMD package 1.9.3 (available at http://www.ks.uiuc.edu/Research/vmd/) (see Figure 2).[37]

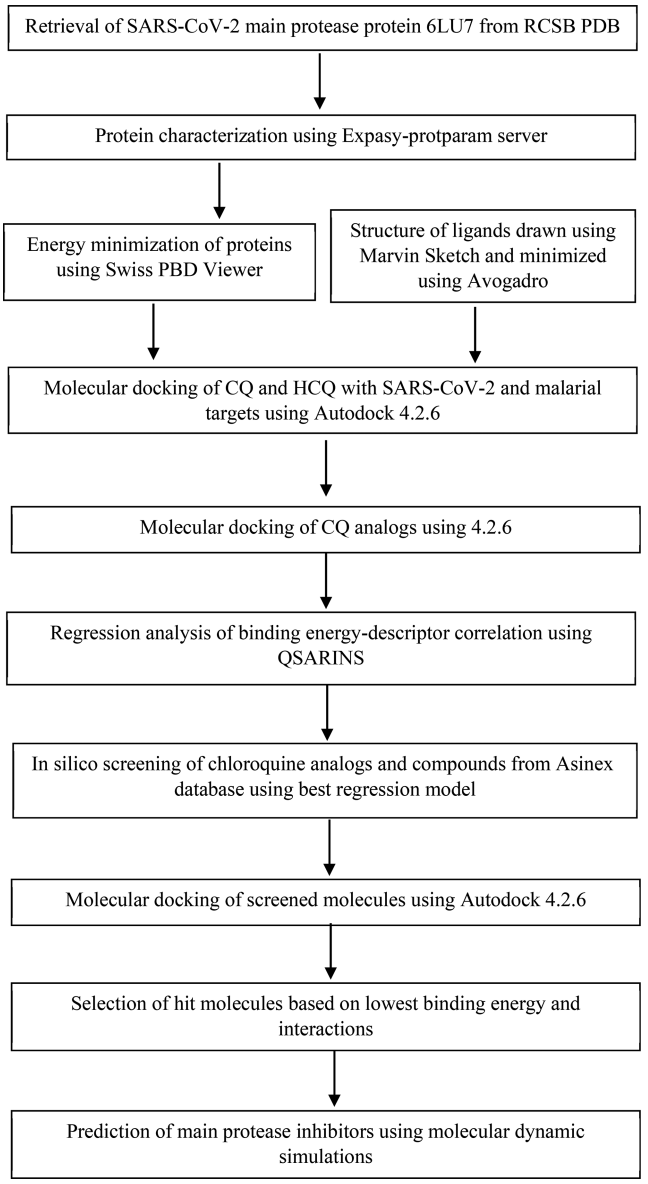


Figure 2. Flowchart for the methodology used in prediction of main protease inhibitor.[38]

3. RESULTS AND DISCUSSION

3.1. Protein Characterization

Being an inevitable enzyme for replication and transcription of the virus, the 3-chymotrypsin-like-cysteine protease (3CL^{pro}) having PDB ID 6LU7 was retrieved from RCSB PDB and characterized by primary and secondary structure analysis as shown in Table 1.

Table 1. Values of Features for Protein Characterization

properties	values
molecular weight	333,797.64 kDa
energy	−16,473.465 kJ/mol
resolution	2.16 Å
theoretical pI	5.95
aliphatic index	82.12
GRAVY	−0.019
instability index	27.65

The pI value (isoelectric point) of 5.95 shows the slightly acidic nature of the protein. The aliphatic index (AI) is a measure of the relative volume occupied by the side chains Ala, Leu, Ile, and Val, which are the aliphatic residues in a protein. High AI ranging from 66.5 to 84.33 implies high thermal stability and hydrophobicity, which help it for membrane penetration in biological system.[39,40] The value for grand average of hydropathy (GRAVY) indicates the hydrophilicity of the protein.[40] The GRAVY value of −0.019, which is close to zero, shows that the protein is hydrophobic. Instability index <40 shows that in vivo half-life is greater than 60 h, which indicates the stability of the protein.[41,39] All the parameters show that the 6LU7 is a stable, hydrophobic protein.

The 6LU7 protein consists of three domains in which the catalytically active site is present between Domain I and II. Figure S1 (Supporting Information) shows the domains and active site of 6LU7 protein bound with its cocrystallized inhibitor N3.[10,42]

3.2. Comparative Docking of CQ and HCQ with 3CL^{pro} and PfDHFR-TS Targets

The antimalarial drugs chloroquine (CQ) and its hydroxy analog hydroxychloroquine (HCQ) are currently used as drugs for the treatment of COVID-19.[13,17] To compare its efficiency as an anti-COVID-19 drug to that of an antimalarial drug, docking studies were carried out for both targets, 3CL^{pro} and PfDHFR-TS. *Plasmodium falciparum* dihydrofolate reductase thymidylate synthase (PfDHFR-TS) is one of the important targets for antimalarial drugs[43,44] (Supplementary Table S1).

It was identified that for both the targets HCQ has lower binding energy owing to its higher binding affinity to the protein. Both the drugs are more active on PfDHFR-TS than 3CL^{pro}. The 2D interaction diagrams and docked poses of CQ and HCQ with the target 3CL^{pro} are given in Figures S2 and S3, respectively. HCQ fits more perfectly to the active site of the target than CQ, and hence, HCQ has high negative binding energy, which imparts more stability to the docked complex (Supplementary Figure S3).

CQ forms three hydrogen bonds with Gly143, Cys145, and His164; van der Waals interaction with Asn142 and Gln189; π-alkyl interaction with His163; and two π-donor hydrogen bond interactions with the −SH group of Cys145 at distances 3.78 Å, 5.16 Å, 5.31 Å, 4.30 Å, 5.10 Å, 6.63 Å, 4.93 Å, and 4.95 Å, respectively (Supplementary Figure S2a, Table 2). The Cys145-His41 diad is blocked by these interactions. The interaction of alkyl group of the CQ with the π-cloud of imidazole ring of histidine may disturb the delocalization of π electrons. The basicity of the imidazole ring is due to the

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

Table 2. Binding Energy of Chloroquine Analogs and Their Interactions

lig name	binding energy (kcal/mol)	interaction with amino acid residues	
		hydrogen bond	hydrophobic and other interactions
CQ	−6.13	Gly143, Cys145, His164	His41, Asn142, His163
HCQ	−6.58	Phe140, Asn142, Ser144, Glu166	Phe140, Asn142, Cys145, Glu166, His172
1	−6.76	Phe140, Glu166	Cys145, Ser144
2	−5.67	Asn142	Leu141, Cys145
3	−5.9		Val104, Ile106, Phe112, Ile136, Ser158, Cys160, Tyr182,
4	−6.55	His41, Cys145, Glu166	Leu27, Phe140, His163, Met165, His172
5	−5.31	Ser144, Glu166	Cys145, His163
6	−5.91	Asn142	Cys145
7	−5.32	Thr111	Val104, Ile106, Asn151, Thr292, Asp295
8	−7.19	Cys145, Glu166	Met49, Cys145
9	−5.29	Asn142	Phe140, Cys145
10	−6.62	Glu166	His41, Met165, Glu166
11	−5.67		Met49, Phe140, Leu141, Met165, Glu166
12	−5.23	Asn142	Asn142, Cys145, Glu166, Leu167, Pro168
13	−6	Asn142, Ser144, Glu166	Asn142, Cys145
14	−5.89	Asn142, Ser144	Asn142, Cys145
15	−6	Asn142, Ser144	Cys145, Asn142
16	−5.98	Phe140, Asn142, Glu166	His41, Ser139, Leu141, Cys145, His172
17	−5.9	Glu166	Phe140, Cys145, Glu166
18	−5.98	Asn142, Glu166	Asn142, Cys145, Met165, Glu166, Pro168,
19	−5.81	Ser139, Asn142, Ser144	Phe140, Leu141 Asn142, Cys145, His172
20	−5.81	Leu141, Glu166	His41, Leu141, Cys145, Met165
21	−5.65	Asn142, Glu166	Asn142, Cys145, Glu166
22	−7.21	Ser139, Ser144, Asn142	Phe140, Leu141, Asn142, Cys145, His172
23	−5.8	Phe140, Asn142, Glu166	His41, Phe140, Leu141, Asn142, Gly143, Cys145, His163, Glu166
24	−7.05	Gly143, His164	Phe140, Cys145, His163, Met165, Glu166
25	−7.28	Asn142, Ser144, Glu166	Phe140, Leu141, Cys145
26	−5.25	Asn142, Glu166	His41, Phe140, Leu141, Cys145, Glu166, Pro168
27	−6.15	Gly143, His164	Cys145, Met165, Glu166, Leu167, Pro168
28	−6.47	Glu166	His41, Phe140, Leu141, Asn142, Cys145, His163, Met165, Glu166
29	−7.33	Phe140, Gly143, Ser144, Cys145, Glu166	Cys145
30	−5.49	Asn142, Ser144, Glu166	Phe140, leu141, Cys145
31	−6.03	Leu141, Glu166	Leu27, His41, Phe140, Gly143, Cys145, Glu166
32	−5.97	Gly138, Phe140, Asn142	Phe140, Asn142, Gly143, Ser144, Cys145, Glu166
33	−6.05	Phe140, Asn142, Glu166	Ser139, Cys145, His163, Met165, Glu166, His172

resonance stabilization of the positive imidazolium ion (Figure 3).[45]

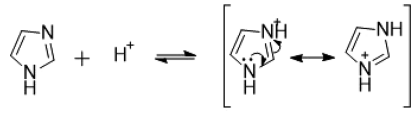


Figure 3. Resonance stabilization of imidazolium ion.

The alkyl−π interaction of the alkyl donor and aromatic acceptor increases the stability of the docked complex as well as reduces the ability of the imidazole ring to accept H^+ from the Cys residue, which is the initial step in the catalytic cycle in proteolysis, thereby reducing the activity.[46,11]

HCQ forms two hydrogen bonds with Phe140, and one each with Asn142, Ser144, and Glu166 at distances 5.15 Å, 5.0 Å, 3.94 Å, 4.31 Å, and 5.04 Å, respectively. It forms two van der Waals interactions with Phe140, and one each with Glu166 and His172 at distances 6.59 Å, 7.12 Å, 5.30 Å, and 5.63 Å, respectively. It also forms a $\pi-\sigma$ interaction with Asn142 and two π−alkyl interactions with Cys145 (Supplementary Figure S2b, Table 2). The π−alkyl interaction with Cys145 may disturb the conformation required for the interaction with polypeptide in the catalytic cycle.

There have been reports about the entry of SARS coronavirus as acidic pH dependent in which the activation occurs by the fusion of the membranes of virus and cellular endosomes and viral genome enters the cytoplasm.[47,48,15] The presence of amino side chain in CQ and HCQ makes them basic and increases the pH of cell organelles like endosomes and lysosomes, which interferes with the replication and inhibits the viral activation.[17,15]

3.3. Molecular Docking Analysis of CQ and HCQ Analogs

To find more potent inhibitors of $3CL^{pro}$, analogs of chloroquine were subjected to molecular docking, which were obtained from literature, Pubchem database, by replacement with isosteres and functional group variations[18−20] of chloroquine (Supplementary Tables S2 and S3). The binding energy thus obtained from docking with 6LU7 protein is given in Table 2.

All the chloroquine analogs show some of the key interactions with the active site residues. The negative value shows a release of energy while forming a protein−ligand docked complex, which imparts stability. The more negative the binding energy, the higher will be the stability and binding affinity. The compounds **29**, **25**, **22**, **8**, **24**, and **10** have more negative binding energy than CQ and HCQ. Their 2D ligand interaction diagrams are shown in Supplementary Figure S4. From the structures of these six compounds, we can infer that if the side chain has an extra −NH group, it can introduce additional hydrogen bonds, which results in a decrease in binding energy. In compound **24**, the tail of side chain has become a part of a ring, which can interact with Met165 residue via π−alkyl interaction. Because of the presence of an additional −NH group, the basicity of these compounds is higher than CQ and HCQ, and thus, they can act as better inhibitors against SARS-CoV-2.

Compounds **29**, **25** **22**, **8**, **24**, and **1** form π-donor hydrogen bonds with Cys145, which will make the −SH hydrogen less available for catalytic cycle. Compounds **29**, **25**, **22**, **24**, and **1** have a π−alkyl interaction with Cys145. Overall, these interactions may affect the catalysis of proteolysis reaction

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

Table 3. Parameters of Top Five Regression Models Fitting the Criteria

model no.	descriptors used	R^2	R_{adj}	Q^2_{loo}	F	s	K_{xx}	δ_k	$RMSE_{tr}$
1	SHBint4, minsNH2, n6Ring, GraphFP567, KRFP567	0.9101	0.8897	0.8765	44.500	0.2144	0.2642	0.0591	0.1900
1	SCH-7, nHsNH2, minHBint6, FP402, KRFPC476	0.9039	0.8830	0.8488	43.2819	0.2331	0.2378	0.1058	0.2076
2	SCH-7, SHBint4, minsNH2, KRFP434	0.8792	0.8599	0.8384	45.5069	0.2392	0.2415	0.0793	0.2184
3	SHBint4, nTRing, ExtFP698, GraphFP893	0.8612	0.8422	0.8162	45.0798	0.2140	0.1672	0.0695	0.2129
4	Vp-5, maxHBint3, GraphFP409, KRFP607, KRFPC397	0.8596	0.8252	0.8081	26.5003	0.2402	0.3641	0.0669	0.2129
5	AlogP, VP-5, minHBint3, KRFP493	0.7819	0.7439	0.7182	20.6106	0.2908	0.3085	0.0406	0.2635

and hence block viral replication. Compound **29** is primaquine (an 8-aminoquinoline), which is an antimalarial drug.[49,50]

The effect of the side chain in the chloroquinoline scaffold was studied by changing the chain length and methyl substitution of the side chain. Compounds **12** and **23** are obtained by the removal of the branching methyl group from CQ and HCQ structures, which results in an increase in the binding energy to −5.23 and −5.80 kcal/mol, respectively. Therefore, it is understood that the branching next to the −NH− group in 4-aminopentyl side chain is important. Isosteric substitution of $-CH_3$ group in compounds **21**, **22**, and **26** with −OH, $-NH_2$, and $-CF_3$ results in binding energies −5.75, −7.21, and −5.25 kcal/mol, respectively. Substitution with NH_2 resulted in a decrease in binding energy compared to HCQ by a factor of −0.63 kcal/mol, which means that an extra 0.63 kcal/mol of energy is released while forming a docked complex, which is more stable than that of HCQ.

An increase in binding energy due to the variation of side chain length in CQ and HCQ structures of compounds **6**, **9**, **11**, **14**, **15**, **16**, **17**, **18**, and **19** makes them less stable complexes with the protein than with CQ and HCQ. Therefore, the chain length of 4-aminopentyl group is optimal for showing biological activity. When an additional −OH group is added to the ethyl chain of HCQ resulting in 2-[(4-aminopentyl)(2-hydroxyethyl)amino]ethan-1-ol side chain as in compound **13**, as a result, the binding energy is decreased to −6.0 kcal/mol. Therefore, the −OH group at one terminal ethyl chain only is required for their inhibitory activity. Rearrangement of the amino side chain from the fourth position to fifth position of quinoline ring in HCQ (molecule 33) increased the binding energy by 0.53 kcal/mol. Replacement of quinoline scaffold by isoquinoline scaffold also increased the binding energy by an amount of 0.61 kcal/mol.

3.4. Regression Analysis

The CQ, HCQ, and its 33 analogs taken for this study were subjected to MLR analysis. Regression models were generated with binding energy as the response and molecular descriptors as the variables. Many models were generated by varying 1−5 variables at a time from a set of 1602 descriptors. The fitness criteria of a model include $R^2 \geq 0.6$, $R^2_{adj} \geq 0.6$, $Q^2_{loo} \geq 0.6$, high value for Fisher ratio (F), low regression standard deviation (s), lower correlation between descriptors (K_{xx}), positive value for δ_k, which is the difference in correlation between descriptors and the response (y) value and correlation among descriptors, and smaller root−mean−squared error for training calculation ($RMSE_{tr}$).[51,52] The parameters of the best five models calculated are given in Table 3.

Considering all the models generated, model 1 has the highest fitness score. However, upon inspecting its external validation parameters $RMSE_{ext} = 0.9003$ and $R^2_{ext} = 0.3245$, the values did not meet the cutoff and hence show poor predictability. For a good prediction, $RMSE_{ext}$ should not vary much from $RMSE_{tr}$ (difference should be less) and R^2_{ext} should be greater than 0.6.

Model 2 was selected for further studies since it was the best fit model with next highest R^2 and Q^2_{loo}, which represents the fitness, robustness of the model, and the proximity of actual and predicted binding energy values. The model was built by incorporating the compounds **2**, **5**, **11**, **21**, **26**, **28**, and **30** in the test set and all others in the training set. Thus, the binding energy, the response value (y), can be expressed in terms of the descriptors considered for the generation of model 2 in the form of a linear relationship $y = \Sigma m_i x_i + c$ where m_i is the coefficient obtained for each descriptor in MLR model and x_i is the value for the selected descriptor. The internal and external validation parameters determine the feasibility and predictivity of the model[29,32] (Table S4). The linear equation correlating binding energy and descriptors generated by model 2 can be written as shown in eq 1:

$$\begin{aligned}\text{Binding energy} = {} & -9.8164(\text{SCH-7}) + 1.0801(\text{nHsNH2}) \\ & + 4.4844(\text{minHBint6}) - 0.4483(\text{FP402}) \\ & - 0.5525(\text{KRFPC476}) - 2.9987 \end{aligned} \quad (1)$$

The root−mean−square error for cross-validation $RMSE_{cv} = 0.2296$ is greater than $RMSE_{tr} = 0.2008$, making the model stable and predictive. The criteria for Y scrambling R^2_y scr < 0.2, Q^2_y scr < 0.2, and R^2_y scr > Q^2_y scr implies that there is no correlation between descriptors, and hence, the model is not simply correlated by chance. Model 2 was assessed with external validation parameters $R^2_{ext} > 0.6$ and small difference between $RMSE_{ext}$ and $RMSE_{tr}$ and was hence acceptable due to its high predictivity. The predicted values of binding energy for both training and test set of molecules are given in Table S5. The plot of observed versus predicted binding energy of training and test set molecules is given in Figure 4. Substitution of −Cl with $-NH_2$ in quinoline scaffold in compound **30** led to inaccurate prediction and hence was detected as an outlier.

Model 2 shows that the binding energy is a linear combination of the descriptors SCH-7, nHsNH2, minHBint6, FP402, and KRFP402. SCH-7 is simple seventh order chi chain that is a topological descriptor based on interatomic distances calculated by the bonds between them representing molecular connectivity as a chemical graph like a bond-line formula of chemical structures. It considers the specificity of the structures at a fragment level rather than the whole molecule.[53] The order 7 represents the number of edges in the graph, which indicates the branching.[54,55] The negative value of its coefficient indicates its negative impact on biological activity. Thus, a low value for SCH-7 is required for a compound to show low binding energy, which makes a restriction on branching. nHsNH2 represents the number of NH_2 groups and amine hydrogens.[56] The positive coefficient of 1.0801 indicates that as the number of NH_2 groups increases, the binding energy decreases, which is in perfect agreement

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

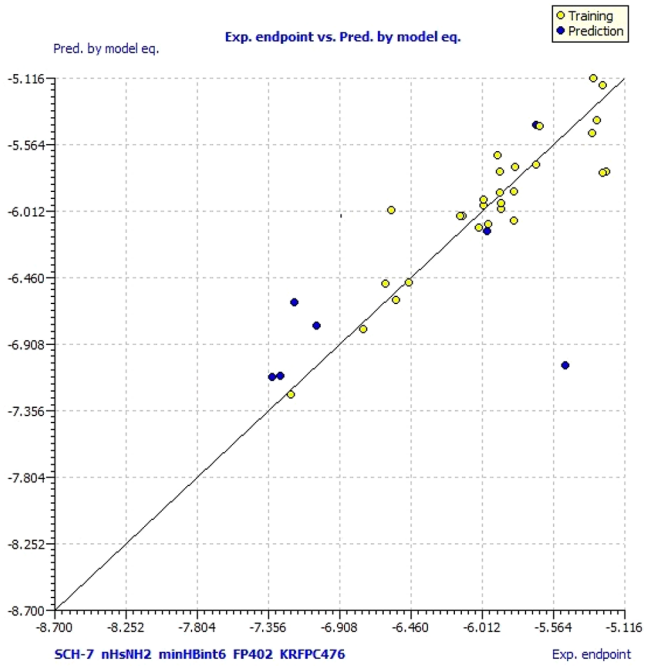


Figure 4. Observed versus predicted binding energies calculated by Model 2 for training and test set.

with the docking results calculated. minHBint6 is the minimum E-state descriptor of strength for potential hydrogen bonds of path length 6.[57] It is an atom-type electro-topological state descriptor in which both electronic and topological functions are considered.[53] There is a positive dependence of minHBint6 with the binding energy because as the chances of H-bond increases, the binding energy will be more negative. FP402 and KRFPC476 are fingerprint descriptors, which account for structure, specific properties, connectivity, or pharmacophores.[58] KRFPC476 accounts for the count of substructures also.[56] Both of these fingerprints are inversely related to binding energy. Designing molecules in such a way that having high values for nHsNH2 and minHBint6 and low values for SCH-7, FP402, and KRFPC476 will help to reduce the binding energy considerably.

3.5. Binding Energy Prediction and Molecular Docking Analysis of Primaquine Analogs

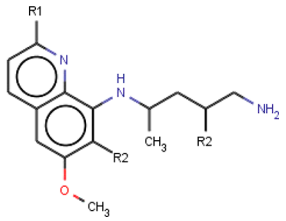
Compound **29**, primaquine, shows the least binding energy among the CQ analogs. It is the only antimalarial drug that can interfere with a majority of the life cycle stages of the malarial protozoan species.[59,60] The combined usage of chloroquine and primaquine may mask chloroquine resistance.[61] However, its toxic effect is mainly due to the hemolytic lesions caused by methemoglobin production.[59,60] It is reported that introduction of a *tert*-butyl group at C2 position (R1 in scaffold given in Table 4) will increase the efficiency and is completely drained of from methemoglobin production as in compound pq1.[59,60]

3.6. In Vitro Screening of Antiviral Compounds

By using the formulated regression Model 2, we predicted the binding energy of some primaquine analogs obtained from the literature and PubChem database and then carried out their molecular docking studies on $3CL^{pro}$ target to check the inhibitory potency of the ligands, given in Table 4. Also screened, 10,584 Asinex compounds having antiviral properties using the regression model 2. It was observed that the predicted binding energy of some compounds obtained from

Table 4. Actual and Predicted Binding Energies of Primaquine Analogs and Their Interactions with Active Site of 6LU7 Protein

Parent Backbone



compound	R1	R2	R3	predicted binding energy	actual binding energy	interactions	
						hydrogen bond	hydrophobic and other interactions
pq1	$-C(CH_3)_3$	H	H	−7.2615	−7.39	Phe140, Glu166	His41, Met165, Glu166, Leu167
pq2	$-CH(CH_3)_2$	H	H	−7.3131	−7.70	Phe140, Gly143, Glu166	Cys145, Met165
pq3	$-OCH_3$	H	H	−7.8998	−7.56	Phe140, Gly143, Ser144, Cys145, Glu166	His163
pq4	$-C(CH_3)_3$	$-CH(CH_3)_2$	H	−7.1488	−7.74	Phe140, Glu166	His41, Met165, Glu166
pq5	$-CH(CH_3)_2$	$-CH(CH_3)_2$	H	−7.422	−7.04	Glu166, Leu167	His41, Leu141, Met165, Glu166, Leu167
pq6	H	H	CH_3	−7.2005	−7.35	Asn142, Gly143, Glu166	Cys145
pq7	H	H	$-CH(CH_3)_2$	−6.0402	−7.33	Phe140, Gly143, His163, Glu166	Cys145, Met165
pq8	H	H	$-CH_2CH_2NH_2$	−7.1204	−8.58	Phe140, Glu166, His172	Cys145
pq9	$-CH(CH_3)_2$	H	$-CH_2CH_2NH_2$	−8.2004	−8.48	Phe140, Glu166, His172	His41, Met49, Cys145, Met165
pq10	H	H	$-CH_2NH_2$	−8.3932	−8.32	Phe140, Gly143, His163, Glu166	Cys145
pq11	$-CH(CH_3)_2$	H	$-CH_2NH_2$	−7.3889	−7.98	Glu166, Leu167	Cys145, His163, His164, Met165, His164

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

Table 5. Binding Energies of Molecules Available in Asinex Compound Library Obtained through Docking Analysis and Predicted by the Model and Their Interactions with Active Site of 6LU7 Protein

			interactions	
compound	predicted binding energy	actual binding energy	hydrogen bond	hydrophobic and other interactions
A1	−23.9789	−7.14		His41, Cys145, Met165, Glu166
A2	−18.4293	−5.91	Glu166	His41, His163, Glu166
A3	−14.1327	−7.15		Leu27, Cys145, Met165
A4	−13.5978	−7.48	Glu166	Leu27, Gly143, Cys145, Glu166
A5	−12.7448	−7.81	Ser144, Cys145, Glu166	His41, Ser144, His164, Cys145, Met165, Glu166, Pro168
A6	−12.1664	−7.60	Leu141, Cys145	Thr26, Gly143, Cys145, Met165, Glu166
A7	−11.5557	−7.72	Gly143, Ser144, Cys145	Phe140, Asn142, Cys145, His163, Met165, Glu166, His172
A8	−11.5058	−6.33	Gly143	Leu27, Lys137, Leu141, Asn142, Cys145, Glu166, His172
A9	−11.4991	−7.17		His41, Phe140, gly143, Cys145,
A10	−11.4704	−7.05	Phe140, Gly143, Glu166	His41, Cys145, Lys137, Glu166
A11	−11.2018	−7.66	Cys145, Glu166	His41, Leu141, Asn142, Cys145, Met165, Glu166, Pro168
A12	−11.1328	−8.18	Gly143, Cys145, Glu166	Leu27, His41, Cys145, Met165, Glu166, Pro168
A13	−10.9315	−7.72	Ser144, Cys145	Leu27, Cys145, Glu166, Pro168
A14	−10.8543	−7.22	Cys145, Glu166	Leu27, Cys145, His163
A15	−10.7756	−6.98	Cys145	Leu27, His41, Asn142, Gly143, Cys145, Met165, Gln189
A16	−10.602	−6.51	Phe140, Gly143, Glu166	His41, Lys137, Asn142, Gly143, Cys145, Glu166
A17	−10.3794	−5.97		Thr26, Leu27, His41, His163, Cys145
A18	−10.3451	−7.51	Thr26	Thr25, Leu27, Cys145, Met165, Glu166, Pro168
A19	−10.1495	−7.57	Cys145, Glu166	Thr26, Cys145, Met165, Leu167, Pro168
A20	−10.1483	−7.65	Gly143, Glu166	Leu141, Asn142, Ser144 Cys145, Glu166, Leu167, Pro168
A21	−10.1232	−6.89	Ser144, His163, Glu166	Cys145
A22	−10.0867	−6.77	Phe140	Leu27, His41, Cys145, Met165, Glu166
A23	−9.34084	−6.93	Cys145	Leu27, His41, Gly143, Cys145, Glu166
A24	−9.21909	−7.12	Glu166	His41, Leu141, His163, Met165, Glu166, His172
A25	−9.09923	−7.85	Ser144, Cys145	Asn142, Pro168
A26	−8.94774	−7.17	Thr26, Ser144, Cys145	Cys145, Glu166
A27	−8.88306	−7.08	Gly143, Glu166	Leu27, His41, Asn142, Gly143, Cys145, His163, Glu166
A28	−8.51347	−6.23	Glu166	Cys145, His163, Met165, Glu166, Leu167, Pro168
A29	−8.06925	−7.22		Val104, Ile106, Thr111, Phe112, Met130, Ile136, Cys160, Tyr182
A30	−8.05939	−7.45	Ser144, Cys145, Glu166	Cys145, His163, Met165

the Asinex database differed largely from its actual value, which was obtained from docking analysis. This may be due to their difference in scaffold from that of the compounds used to build the regression model. Even though the predicted binding energy was less accurate, they have furnished better binding energies on docking with the target protein implying the reliability of the model. Thirty molecules that showed lower binding energies were subjected to molecular docking analysis to identify the potent $3CL^{pro}$ inhibitors (Supplementary Figure S5). Binding energies obtained through both the methods and their interactions with the receptor protein are given in Table 5.

All the primaquine analogs showed lower binding energies than CQ and HCQ. All the compounds except pq5 furnished binding energies lower than that of primaquine. The binding energies calculated using the regression Model 2 are in good agreement with the value obtained from docking analysis. All the compounds exhibited interactions with some of the active site residues including the catalytic diad residues either Cys145 or His41 or both, which make them potent $3CL^{pro}$ inhibitors. The compounds pq8, pq9, and pq10 exhibited lower binding energies of −8.58, −8.48, and −8.32 kcal/mol. The 2D interaction diagrams of these ligands with the protein are given in Figure 5. In accordance with the findings based on docking studies, the regression analysis of CQ analogs reveals the presence of one extra $-NH_2$ bond in their structure.

Inspecting the 2D interaction diagrams of compounds A6, A9, A10, A15, A16, A23, A29, A30 it can be seen that there are unfavorable donor−donor or acceptor-acceptor interactions present between the compounds and target protein which may negatively affect the stability of the docked complexes (Supplementary Figure S6). Hence these can be considered as poor inhibitors of $3CL^{pro}$ protein even though they have binding energies less than HCQ. From these results it is clear that binding energy as well as interactions play decisive roles in determining the druggability of compounds.

Asinex database compounds A12, A5, A13, A11, A20, A19, A18, A4, A14, and A3 obtained lower binding energies among all (Figure 5 and Figure S7). These compounds form interactions with Cys145, His 41, or both along with other active site residues, which are responsible for their inhibitory property. The compounds A12, A5, A13, A11, A19, and A14 form hydrogen bonding interactions with Cys145, which will obstruct the initial deprotonation step in the catalytic cycle of $3CL^{pro}$ as shown in Figure 1. The $\pi-\pi$ interaction of His 41 with π systems of the compounds A12 and A5 will affect the proton accepting ability of His41. The π−sulfur interactions are present in aromatic rings with −SH group of the compounds A13, A20, A18, A4, A14, and A3 with Cys145 residue. It involves the interaction with the aromatic π orbitals and σ^* acceptor orbital of SH bond of Cys145. S−H/π interactions are the important contributors to binding energy.[62] The thiol group of Cys145 is engaged in S−H/π

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

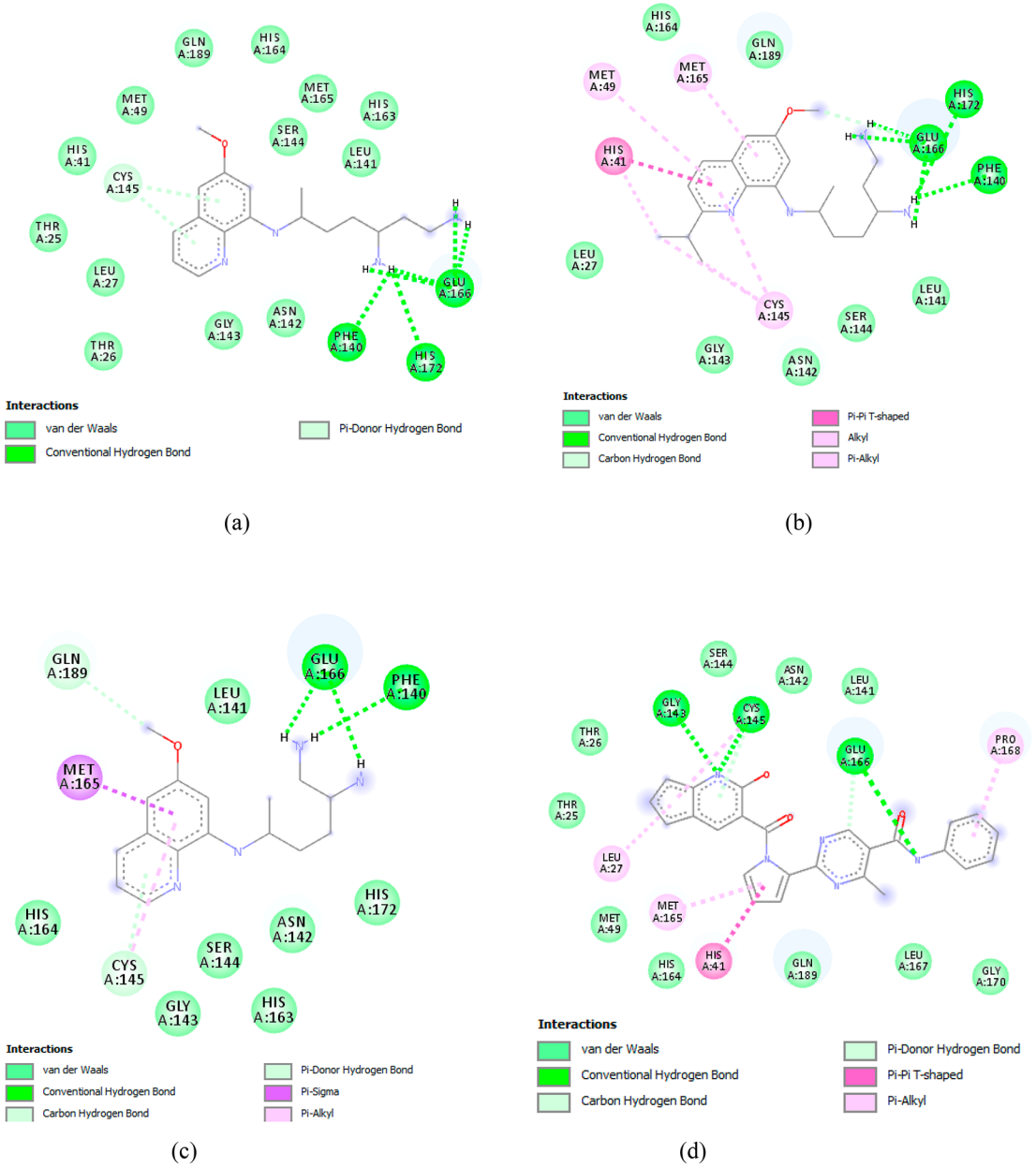


Figure 5. 2D interaction diagrams of pq8, pq9, pq10, and A12 with target 6LU7.

with the aromatic rings of the inhibitors and is barely available for catalyzing the proteolysis. The combined effect of all the interactions with the target protein makes these molecules potent inhibitors of the main protease enzyme of SARS-CoV-2 virus.

Comparing the structures of the hit compounds pq8, pq9, pq10, and A12 with the rest, it was clear that the high flexibility of the structures owing to more single bonds and less compact rings makes them able to undergo conformational changes at the active site of the target leading to more interactions with Cys145 and His41 along with other active site residues. Hence, these four compounds were selected for further analysis.

3.7. Molecular Dynamics Simulation

Even though docking studies give the best binding poses, they do not account for the conformational changes taking place upon ligand binding as the docking analysis preferably treats protein as rigid. The conformational changes affecting the binding affinity could be analyzed using molecular dynamic simulations.[63] Four top ranked compounds via docking (pq8, pq9, pq10, and A12) were selected for 20 ns molecular dynamics simulation studies to understand the conformational changes in the protein on binding of the ligand. RMSD, RMSF, and Radius of gyration plots were used to evaluate the stability of the complexes. The binding conformations of the compounds at the last frame of simulation are given in Figure 6.

The root−mean−square deviation (RMSD) plot of apoprotein $3CL^{pro}$, pq8, pq9, pq10, and A12 with average RMSDs 1.75, 1.65, 1.93, 2.02, and 1.86 Å, respectively, is given in Figure 7a. It can be inferred that for the apo structure as well as the docked complexes except pq10, RMSD values rose to 2.54 Å and then formed a plateau after 12 ns of simulation, which indicates that the system has reached stable equilibrium state and has been confined in the active site. The RMSD of pq10 requires an extension of stabilization time to reach

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

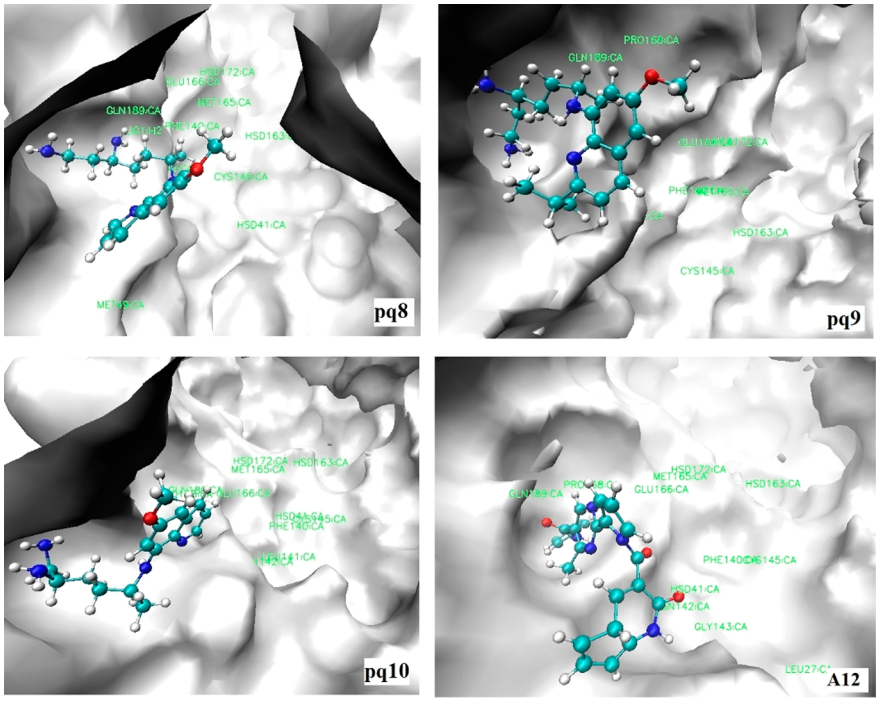


Figure 6. Final poses (20 ns) of ligand protein-complexes in the MD simulations.

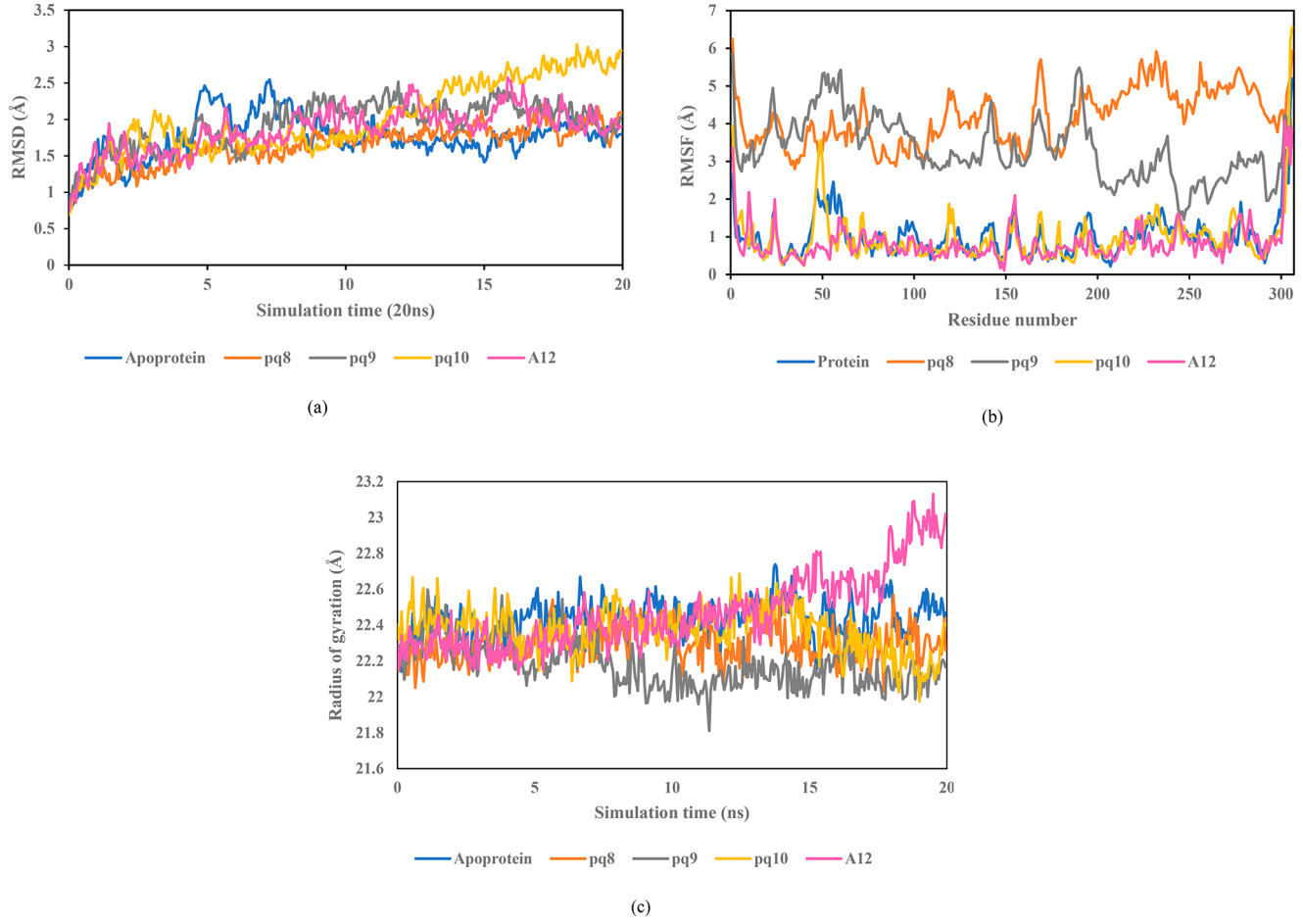


Figure 7. (a) RMSD. (b) RMSF. (c) Radius of gyration plots for pq8, pq9, pq10, and A12.

equilibrium. The higher RMSD fluctuations for the holo structures compared to the apo structure indicate larger conformational changes followed by repositioning of the ligand in the active site.

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

The root−mean−square fluctuation (RMSF) plot gives the fluctuations of individual residues in the protein backbone.[64] Higher RMSF denotes higher flexibility and vice versa.[65] The RMSF plot (Figure 7b) showed greater residue flexibility for pq9 and pq10 compared to the unbound form. pq8 and A12 have comparable RMSF values with that of the apo form. The Cys145 and His41 residues involved in the catalytic cycle have fluctuations with greater than 0.7 Å in apoprotein. In pq8 and pq9, it fluctuates more than 3 Å, while in pq10 and A12, there is fluctuation less than 0.6 Å. In all the systems, the active site residues from 140 to 190 have the highest fluctuations.

The structural flexibility or compactness of the protein molecule is analyzed by the radius of gyration.[66] The protein is supposed to maintain a relatively steady value for R_g. From inspection of Figure 7c, it can be seen that R_g has almost no significant variance with the average R_g values of 22 Å. The ligand bound form has lower R_g values compared to apoprotein, which indicates that the bound structures are more compact.

4. CONCLUSION

Molecular docking analysis of chloroquine and hydroxychloroquine on main protease ($3CL^{pro}$) protein has shown that HCQ is more efficient than CQ. More potent inhibitors of $3CL^{pro}$ have been identified through docking studies. The interactions with the active site residues, especially with Cys145 and His41, which are the key residues involved in catalytic diad for proteolysis in the enzyme, are responsible for the activity of these compounds. A regression model with R^2 = 0.9039 and Q^2 = 0.8848 was developed correlating binding energy with the molecular descriptors. This model was used to predict the binding energies of novel molecules. The primaquine derivatives and antiviral compounds obtained from Asinex compound library were screened using the regression model, and the hit compounds thus obtained were further docked with $3CL^{pro}$ protein. They showed less binding energies and proper interactions with the target, main protease ($3CL^{pro}$). Molecular dynamics studies on these top four compounds pq8, pq9, pq10, and A12 revealed their binding stability and conformational changes associated with the protein−ligand complexes. Hence, these four compounds are proposed for validation and further studies as more efficient medication for COVID-19.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00683.

Binding energy obtained for docking CQ and HCQ with $3CL^{pro}$ and PfDHFR-TS; structure of CQ analogs with 4-aminoquinoline scaffold; structure of chloroquine analogs with quinoline scaffold, values for internal and external validation criteria, binding energy values of training and test molecules predicted with Model 2; domains and active site of 6LU7 bound with inhibitor N3; 2D interaction diagrams of docked complexes of CQ, HCQ with $3CL^{pro}$; docked poses of CQ and HCQ in active site of 6LU7; 2D interaction diagrams of **29**, **25**, **22**, **8**, **24**, **1**, and **10** with 6LU7; structures of screened compounds from Asinex compound library; 2D interaction diagram of compounds having unfavorable interaction with target; 2D interaction diagrams of Asinex compounds with 6LU7 protein (PDF)

■ AUTHOR INFORMATION

Corresponding Author

V. L. Pushpa − *PG and Research Department of Chemistry, Sree Narayana College, Kollam, Kerala 691001, India;* orcid.org/0000-0001-5332-7806; Email: drvlpushpa2017@gmail.com

Authors

A. S. Achutha − *PG and Research Department of Chemistry, Sree Narayana College, Kollam, Kerala 691001, India*

Surendran Suchitra − *PG and Research Department of Chemistry, Sree Narayana College, Kollam, Kerala 691001, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00683

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge generous funding from Kerala State Council for Science, Technology and Environment (KSCSTE) for the Junior Research Fellowship (No. 414/2020/KSCSTE).

■ REFERENCES

(1) Liu, C.; Zhou, Q.; Li, Y.; Garner, L. V.; Watkins, S. P.; Carter, L. J.; Gregg, A. C.; Daniels, A. D.; Jervey, S.; Albaiu, D.; et al. Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Cent. Sci.* **2020**, *6*, 315−331.

(2) Udugama, B.; Kadhiresan, P.; Kozlowski, H. N.; Malekjahani, A.; Osborne, M.; Li, V. Y. C.; Chen, H.; Mubareka, S.; Gubbay, J. B.; Chan, W. C. W. Diagnosing COVID-19: The Disease and Tools for Detection. *ACS Nano* **2020**, *14* (4), 3822−3835.

(3) Wang, L.; Wang, Y.; Ye, D.; Liu, Q. A Review of the 2019 Novel Coronavirus (COVID-19) Based on Current Evidence. *Int. J. Antimicrob. Agents* **2020**, *55* (6), 105948.

(4) Chang, L.; Yan, Y.; Wang, L. Coronavirus Disease 2019: Coronaviruses and Blood Safety. *Transfus. Med. Rev.* **2020**, *34* (2), 75−80.

(5) Astuti, I.; Ysrafil. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An Overview of Viral Structure and Host Response. *Diabetes Metab. Syndr.* **2020**, *14* (4), 407−412.

(6) *Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update*; WHO, 2019. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessed September 2, 2020).

(7) Zhang, L.; Liu, Y. Potential Interventions for Novel Coronavirus in China: A Systematic Review. *J. Med. Virol.* **2020**, *92* (5), 479−490.

(8) Tahir ul Qamar, M.; Alqahtani, S. M.; Alamri, M. A.; Chen, L.-L. Structural Basis of SARS-CoV-2 3CLpro and Anti-COVID-19 Drug Discovery from Medicinal Plants. *J. Pharm. Anal.* **2020**, *10* (4), 313−319.

(9) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S. An Overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy Department of Medicinal Chemistry, Tokyo University of Pharmacy and Life Sciences, Tokyo. *J. Med. Chem.* **2016**, *59* (14), 6595−6628.

(10) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu,

F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from COVID-19 Virus and Discovery of Its Inhibitors. *Nature* **2020**, *582*, 289−293.
(11) Elsasser, B.; Zauner, F. B.; Messner, J.; Soh, W. T.; Dall, E.; Brandstetter, H. Distinct Roles of Catalytic Cysteine and Histidine in the Protease and Ligase Mechanisms of Human Legumain As Revealed by DFT-Based QM/MM Simulations. *ACS Catal.* **2020**, *7* (9), 486−500.
(12) Lythgoe, M. P.; Middleton, P. Ongoing Clinical Trials for the Management of the COVID-19 Pandemic. *Trends Pharmacol. Sci.* **2020**, *41* (6), 363−382.
(13) Singh, A. K.; Singh, A.; Shaikh, A.; Singh, R.; et al. Chloroquine and Hydroxychloroquine in the Treatment of COVID-19 with or without Diabetes: A Systematic Search and a Narrative Review with a Special Reference to India and Other Developing Countries. *J. Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14* (3), 241−246.
(14) Pastick, K. A.; Okafor, E. C.; Wang, F.; Lofgren, S. M.; Skipper, C. P.; Nicol, M. R.; Pullen, M. F.; Mcdonald, E. G.; Lee, T. C.; Schwartz, I. S.; Kelly, L. E.; Lother, S. A.; Mitjà, O.; Letang, E.; Abassi, M.; Boulware, D. R.; et al. Review: Hydroxychloroquine and Chloroquine for Treatment of SARS-CoV-2 (COVID-19). *Open forum Infect. Dis.* **2020**, *7* (4), No. ofaa130.
(15) Devaux, C. A.; Rolain, J.; Colson, P.; Raoult, D. New Insights on the Antiviral Effects of Chloroquine against Coronavirus: What to Expect for COVID-19? *Int. J. Antimicrob. Agents* **2020**, *5* (5), 105938.
(16) Rolain, J.; Colson, P.; Raoult, D. Recycling of Chloroquine and Its Hydroxyl Analogue to Face Bacterial, Fungal and Viral Infections in the 21st Century. *Int. J. Antimicrob. Agents* **2007**, *30* (4), 297−308.
(17) Liu, J.; Cao, R.; Xu, M.; Wang, X.; Zhang, H.; Hu, H.; Li, Y.; Hu, Z.; Zhong, W.; Wang, M. Hydroxychloroquine, a Less Toxic Derivative of Chloroquine, Is Effective in Inhibiting SARS-CoV-2 Infection in Vitro. *Cell Discovery* **2020**, *6* (16), 6−9.
(18) Parhizgar, A. R.; Tahghighi, A. Introducing New Antimalarial Analogues of Chloroquine and Amodiaquine: A Narrative Review. *Iran. J. Med. Sci.* **2017**, *42* (2), 115−128.
(19) Aguiar, A. C. C.; Murce, E.; Cortopassi, W. A.; Pimentel, A. S.; Almeida, M. F. S.; Barros, D. C. S.; Guedes, J. S.; Meneghetti, M. R.; Krettli, A. U. Chloroquine Analogs as Antimalarial Candidates with Potent in Vitro and in Vivo Activity. *Int. J. Parasitol.: Drugs Drug Resist.* **2018**, *8* (3), 459−464.
(20) Al-Bari, A. A. Chloroquine Analogues in Drug Discovery: New Directions of Uses, Mechanisms of Actions and Toxic Manifestations from Malaria to Multifarious Diseases. *J. Antimicrob. Chemother.* **2015**, *70* (6), 1608−1621.
(21) *PubChem. Database*; National Center for Biotechnology Information, 2020. https://pubchem.ncbi.nlm.nih.gov/ (accessed April 17, 2021).
(22) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2019 Update: Improved Access to Chemical Data. *Nucleic Acids Res.* **2019**, *47* (D1), 1102−1109.
(23) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012**, *4* (1), 1−17.
(24) *Avogadro: An Open-Source Molecular Builder and Visualization Tool. Version 1.XX*; Avogadro Chemistry, 2018. http://avogadro.cc/ (accessed February 7, 2020).
(25) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235−242.
(26) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An Environment for Comparative Protein Modeling. *Electrophoresis* **1997**, *18* (15), 2714−2723.
(27) *SwissPdb Viewer*; Swiss Institute of Bioinformatics, 2019. http://www.expasy.org/spdbv/ (accessed June 1, 2019).
(28) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Software News and Updates AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785−2791.
(29) Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. *J. Comput. Chem.* **2013**, *34* (24), 2121−2132.
(30) Gramatica, P.; Cassani, S.; Chirico, N. QSARINS-Chem: Insubria Datasets and New QSAR/QSPR Models for Environmental Pollutants in QSARINS. *J. Comput. Chem.* **2014**, *35* (13), 1036−1044.
(31) Yap, C. W. E. I. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *J. Comput. Chem.* **2011**, *32* (7), 1466−1474.
(32) Sarithamol, S.; Pushpa, V. L.; Divya, V.; Manoj, K. B. Comparative QSAR Model Generation Using Pyrazole Derivatives for Screening Janus Kinase-1 Inhibitors. *Chem. Biol. Drug Des.* **2020**, *95*, 503−519.
(33) Abdulfatai, U.; Uzairu, A.; Uba, S. Quantitative Structure Activity Relationship Study of Anticonvulsant Activity of α _ Substituted Acetamido- N-Benzylacetamide Derivatives. *Cogent Chem.* **2016**, *2* (1), 1166538.
(34) Phillips, J. C.; Bernardi, C.; Hardy, D. J.; Radak, B. K.; Maia, J. D. C.; Skeel, R. D.; Stone, J. E.; Ribeiro, J. V.; Wang, Y.; Buch, R.; Fiorin, G.; Jiang, W.; Chipot, C.; Phillips, J. C.; Maia, J. D. C.; et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153* (4), 044130.
(35) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859−1865.
(36) Kim, S.; Lee, J.; Jo, S.; Brooks, C. L.; Lee, H. S.; Im, W. CHARMM-GUI Ligand Reader & Modeler for CHARMM Force Field Generation of Small Molecules. *J. Comput. Chem.* **2017**, *38* (21), 1879−1886.
(37) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33−38.
(38) Borgio, J. F.; Alsuwat, H. S.; Al Otaibi, W. M.; Ibrahim, A. M.; Almandil, N.; Al Asoom, L. I.; Salahuddin, M.; Kamaraj, B.; AbdulAzeez, S. Research Paper State-of-the-Art Tools Unveil Potent Drug Targets amongst Clinically Approved Drugs to Inhibit Helicase in SARS-CoV-2. *Arch. Med. Sci.* **2020**, *16* (3), 508−518.
(39) Panda, S.; Chandra, G. Physicochemical Characterization and Functional Analysis of Some Snake Venom Toxin Proteins and Related Non-Toxin Proteins of Other Chordates Abstract: Background: *Bioinformation* **2012**, *8* (18), 891−896.
(40) Sarithamol, S.; V, D.; R, S. V.; Surendran, S.; Pushpa, V. L.; Manoj, K. B. Genetic Involvement of Interleukin 4 for Asthma and Identification of Potential Phytochemical Scaffold Through Molecular Docking Studies. *Int. J. Curr. Pharm. Res.* **2018**, *10* (1), 43−54.
(41) Idicula-Thomas, S.; Balaji, P. V. Understanding the Relationship between the Primary Structure of Proteins and Its Propensity to Be Soluble on Overexpression in Escherichia Coli. *Protein Sci.* **2005**, *14* (5), 582−592.
(42) Jin, Z.; et al. Structure of Mpro from COVID-19 Virus and Discovery of Its Inhibitors. *Nature* **2020**, *582*, 289−293.
(43) Sivaprakasam, P.; Tosso, P. N.; Doerksen, R. J. Structure - Activity Relationship and Comparative Docking Studies for Cycloguanil Analogs as PfDHFR-TS Inhibitors. *J. Chem. Inf. Model.* **2009**, *49* (7), 1787−1796.
(44) Syahri, J.; Yuanita, E.; Nurohmah, B. A.; Armunanto, R.; Purwono, B. Chalcone Analogue as Potent Anti-Malarial Compounds against Plasmodium Falciparum: Synthesis, Biological Evaluation, and Docking Simulation Study. *Asian Pac. J. Trop. Biomed.* **2017**, *7* (8), 675−679.
(45) Morrison, R. T.; Boyd, R. N.; Battacharjee, S. K. *Organic Chemistry*; Pearson, 2011; pp 765−766.
(46) Brandl, M.; Weiss, M. S.; Jabs, A.; Suhnel, J.; Hilgenfeld, R. C-h...Pi-Interactions in Proteins. *J. Mol. Biol.* **2001**, *307* (1), 357−377.
(47) Wang, H.; Yang, P.; Liu, K.; Guo, F.; Zhang, Y.; Zhang, G.; Jiang, C. SARS Coronavirus Entry into Host Cells through a Novel

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717

Clathrin- and Caveolae-Independent Endocytic Pathway. *Cell Res.* **2008**, *18* (2), 290−301.
(48) Yang, Z.; Huang, Y.; Ganesh, L.; Leung, K.; Kong, W.; Schwartz, O.; Subbarao, K.; Nabel, G. J. PH-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus Is Mediated by the Spike Glycoprotein and Enhanced by Dendritic Cell Transfer through DC-SIGN. *J. Virol.* **2004**, *78* (11), 5642−5650.
(49) Graves, P. M; Choi, L.; Gelband, H.; Garner, P. Primaquine or Other 8-Aminoquinolines for Reducing Plasmodium Falciparum Transmission. *Cochrane database Syst. Rev.* **2018**, *2* (2), CD008152.
(50) Baird, J. K.; Hoffman, S. L. Primaquine Therapy for Malaria. *Clin. Infect. Dis.* **2004**, *39* (9), 1336−1345.
(51) Rastija, V.; Molnar, M.; Siladi, T.; Masand, V. H. QSAR Analysis for Antioxidant Activity of Dipicolinic Acid Derivatives. *Comb. Chem. High Throughput Screening* **2018**, *21* (3), 1−11.
(52) Chtita, S.; Ghamali, M.; Ousaa, A.; Aouidate, A.; Belhassan, A.; Taourati, A. I.; Masand, V. H.; Bouachrine, M.; Lakhlifi, T. QSAR Study of Anti-Human African Trypanosomiasis Activity for 2-Phenylimidazopyridines Derivatives Using DFT and Lipinski's Descriptors. *Heliyon* **2019**, *5* (3), e01304.
(53) Roy, K.; Dc, A. U.; Sengupta, C. QSAR with Electro Topological State Atom Index: Antialirenergic Activi Ty of N, N-Dimethyl-2-Bromo-2-Phenylethylamines. *Indian Jounal Chem.* **1999**, *38*, 942−949.
(54) Hall, L. H.; Kier, L. B. The Molecular Connectivity Chi Indexes and Kappa Shape Indexes in Structure-Property Modeling. *Reviews in Computational Chemistry* **2007**, *2*, 367−422.
(55) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Ter, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been ? Where Are You Going To ? *J. Med. Chem.* **2014**, *57* (12), 4977−5010.
(56) Yap, C. W. PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2013**, *32*, 1466−1474.
(57) Edache, E. I.; Uzairu, A.; Abechi, S. E. Investigation of 5,6-Dihydro-2-Pyrones Derivatives as Potent Anti-HIV Agents Inhibitors. *J. Comput. Methods Mol. Des.* **2015**, *5* (3), 135−149.
(58) Yin, Y.; Xu, C.; Gu, S.; Li, W.; Liu, G.; Tang, Y. Quantitative Regression Models for the Prediction of Chemical Properties by an Efficient Workflow. *Mol. Inf.* **2015**, *34*, 679−688.
(59) Jain, M.; Vangapandu, S.; Sachdeva, S.; Jain, R. Synthesis and Blood-Schizontocidal Antimalarial Activities of 2-Substituted/2, 5-Disubstituted-8-Quinolinamines and Some of Their Amino Acid Conjugates. *Bioorg. Med. Chem.* **2004**, *12* (5), 1003−1010.
(60) Jain, M.; Vangapandu, S.; Sachdeva, S.; Singh, S.; Singh, P. P.; Jena, G. B.; Tikoo, K.; Ramarao, P.; Kaul, C. L.; Jain, R. Discovery of a Bulky 2- Tert -Butyl Group Containing Primaquine Analogue That Exhibits Potent Blood-Schizontocidal Antimalarial Activities and Complete Elimination of Methemoglobin Toxicity. *J. Med. Chem.* **2004**, *47*, 285−287.
(61) Pukrittayakamee, S.; Vanijanonta, S.; Chantra, A.; Clemens, R.; White, N. J. Blood Stage Antimalarial Efficacy of Primaquine in Plasmodium Vivax Malaria. *J. Infect. Dis.* **1994**, *169* (4), 932−935.
(62) Forbes, C. R.; Sinha, S. K.; Ganguly, H. K.; Bai, S.; Yap, G. P. A.; Patel, S.; et al. Insights into Thiol-Aromatic Interactions: A Stereoelectronic Basis for S - H/π Interactions. *J. Am. Chem. Soc.* **2017**, *139* (5), 1842−1855.
(63) Radwan, A.; Mahrous, G. M.; Arabia, S. Docking Studies and Molecular Dynamics Simulations of the Binding Characteristics of Waldiomycin and Its Methyl Ester Analog to Staphylococcus Aureus Histidine Kinase. *PLoS One* **2020**, *15* (6), No. e0234215.
(64) Babu, S.; Nagarajan, S. K.; Madhavan, T. Understanding the Structural Features of JAK2 Inhibitors: A Combined 3D-QSAR, DFT and Molecular Dynamics Study. *Mol. Diversity* **2019**, *24* (4), 845−874.
(65) Priya, R.; Sumitha, R.; Doss, C. G. P.; Rajasekaran, C.; Babu, S.; Seenivasan, R.; Siva, R. Molecular Docking and Molecular Dynamics to Identify a Novel Human Immunodeficiency Virus Inhibitor from Alkaloids of Toddalia Asiatica. *Pharmacogn. Mag.* **2015**, *11* (3), 414−422.
(66) Chinnasamy, S.; Selvaraj, G.; Kaushik, A. C.; Kaliamurthi, S.; Chandrabose, S.; Singh, S. K.; Thirugnanasambandam, R.; Gu, K.; Wei, D.-Q. Molecular Docking and Molecular Dynamics Simulation Studies to Identify Potent AURKA Inhibitors: Assessing the Performance of Density Functional Theory, MM-GBSA and Mass Action Kinetics Calculations. *J. Biomol. Struct. Dyn.* **2020**, *38* (14), 4325−4335.

https://dx.doi.org/10.1021/acs.jproteome.0c00683
J. Proteome Res. 2020, 19, 4706−4717