Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Molecular Structure 1237 (2021) 130380

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

Rational design of flavonoid based potential inhibitors targeting SARS-CoV 3CL protease for the treatment of COVID-19

Shipra Bhati[a,*], Vikas Kaushik[b], Joginder Singh[b,*]

[a] Department of Chemistry, The Oxford College of Engineering, Bommanhalli, Bangalore-560068, Karnataka, India
[b] Department of Biotechnology, Lovely Professional University, Phagwara-144411, Punjab, India

ARTICLE INFO

Article history:
Received 16 September 2020
Revised 23 March 2021
Accepted 24 March 2021
Available online 7 April 2021

Keywords:
COVID-19
SARS-CoV-2
3CL protease
Flavonoid
Molecular docking

ABSTRACT

The current outbreak of Coronavirus Disease 2019 (COVID-19) pandemic has reported thousands of deaths worldwide due to the rapid transmission rate and the lack of antiviral drugs and vaccinations. There is an urgent need to develop potential antiviral drug candidates for the prevention of COVID-19 infection. In the present study, a series of potential inhibitors targeting SARS-CoV 3CL protease were rationally designed by incorporating gamma lactam ring, and various fluoro substituted heterocyclic ring systems to the flavonoid scaffold. The prediction of drug-likeness, oral bioavailability, toxicity, synthetic accessibility, and ADMET properties was made by computational means. Quercetin was used as standard. The binding affinity of the ligands towards the 3CL protease target was examined using docking simulations. The designed ligands possess favourable pharmacokinetic and pharmacodynamic properties. Ligand L4, L8, and L14 appeared to be the lead compounds in the series and can be considered for further *in-vivo* and *in-vitro* validation.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has emerged as a global challenge and threatens millions of lives. The World Health Organization (WHO), on 11 March 2020, declared coronavirus disease-19 (COVID-19) a pandemic [1]. Globally, as of 22 March, there are 122,822,505 active cases, and the number of deaths attributed to COVID-19 worldwide has already surpassed 2709,041 [2]. According to the World Health Organization reports, human coronaviruses are spread by respiratory droplets and communication pathways and cause mild upper respiratory problems. Fever, cough, shortness of breath to pneumonia, kidney failure, and even death are some of the symptoms of this disease [3]. A much higher human-to-human transmissibility together with greater lethality relative to influenza is some of the characteristics of the COVID-19 [4,5]. Due to the situation of a "public health emergency", Chloroquine, Hydroxychloroquine, Remdesivir, Ribavirin, Lopinavir/Ritonavir, Favipiravir, Umifenovir, Oseltamivir along with several other antiviral drugs have been tried recently in the treatment of COVID-19 [6]. The strategy of repurposing the drug does not offer a permanent solution, and further clinical trials are required to test the effectiveness of these drugs to treat COVID-19. The approach of rational drug designing based on the pathogenic mechanism of COVID-19 and target protein structure of SARS-CoV-2 has been suggested to fight against COVID-19 [7,8].

The discovery of anti-SARS-CoV-2 drugs includes the de novo development of new, specific drugs based upon the genomic and biophysical understanding of this virus. SARS-CoV-2, belonging to the coronaviridae family, has two viral proteases, papain-like protease (PL^{pro}) and 3C-like protease ($3CL^{pro}$) (also known as M^{pro}) [9]. A recent study has revealed that $3CL^{pro}$ of SARS-CoV-2 is a crucial enzyme that is correlated to viral transcription and replication [10]. It is a three-domain (domains I to III) cysteine protease where domains I, II, and III spans across 8–101, 102–184, and 201–306 amino acid residues, respectively. The active site of binding is located at the cleft of domains I and II, which consists of two catalytic residues, namely, HIS41 and CYS145. 3CLpro is involved in most maturation cleavage events within the precursor polyprotein [11]. Jin et al. in their study reported that $3CL^{pro}$ was involved in the release of polypeptides which are functional extensive proteolysis and cleavage of the enzyme itself from the sites of the genome to stop the multiplication of virus within the cells in the infected patients [12]. $3CL^{pro}$ is a homodimer in its active form and releases viral polypeptides to the body and contribute to the infection to occur. The maturation of viral particles and cleavage of the viral capsid is the main function of $3CL^{pro}$ [13]. The coronaviral proteases, 3C-like protease (3CLpro), are attractive antiviral drug targets because they are essential for coronaviral replication

* Corresponding author.
E-mail address: joginder.15005@lpu.co.in (J. Singh).

https://doi.org/10.1016/j.molstruc.2021.130380
0022-2860/

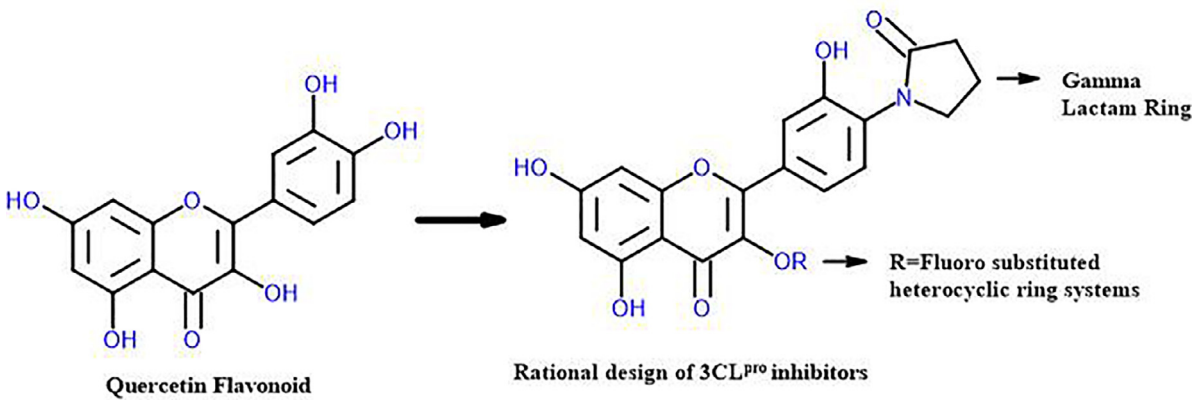

Fig. 1. Design strategy of SARS-CoV-2 3C-like protease ($3CL^{pro}$) inhibitors.

[14]. $3CL^{pro}$ is a promising drug target in the discovery of specific anti–SARS-CoV-2 drugs. Inhibition of the activity of the main protease is expected to block the viral replication as no human proteases are known to have similar cleavage specificity [15].

Flavonoids are naturally occurring polyphenolic compounds characterized by the flavan nucleus. Flavanoids are promising natural compounds against viral infection [16–19]. Quercetin (2-(3,4-dihydroxy phenyl)−3,5,7-trihydroxychromen-4-one) is a widely distributed polyphenolic flavonoid. Quercetin possesses enormous antiviral potential in inhibiting proteases, reverse transcriptase, polymerases, and anti-influenza A viruses [20–23]. Nguyen et al. reported that quercetin inhibits the SARS-CoV main protease ($3CL^{pro}$) with an IC_{50} (50% inhibitory concentration) *in vitro* of 73 μM [24]. Park and coworkers in their study found that both of the SARS-CoV proteases ($3CL^{pro}$ and PL^{pro}) were markedly inhibited by quercetin derivatives [25]. A recent study suggests the combined administration of quercetin and vitamin C for the early treatment of SARS-CoV-2 related disease (COVID-19) [26].

Since no clinically effective vaccines or specific antiviral drugs are currently available for the prevention and treatment of COVID-19 infections, there is an urgent need for potential SARS-CoV-2 drugs. Zhai and his group reported that SARS-CoV $3CL^{pro}$ inhibitors often have an (*S*)-γ-lactam ring that occupies the S1 site of $3CL^{pro}$ [27]. Dai et al. designed and synthesized potential SARS-CoV-2 main protease inhibitors are possessing γ-lactam and fluoro substituted ring systems [10]. In continuation of our work of designing potential therapeutics by virtual screening [28,29], the present investigation includes rational design and exploitation of the pharmacological potential of flavonoid based therapeutics as anti-SARS-CoV-2 for the treatment of COVID-19. A series of quercetin based derivatives were rationally designed by incorporating gamma lactam ring, and various fluoro substituted heterocyclic ring systems. The pharmacological potential, toxicity, and drug-likeness of the designed derivatives were tested employing computational techniques.

2. Materials and methods

The 3C-like protease ($3CL^{pro}$) inhibitors (L1-L15) were designed with the help of ACD Labs ChemSketch 12.0 software (http://www.acdlabs.com). The design strategy of these inhibitors has been shown in Fig. 1. The IUPAC names, the molecular formulae of newly designed derivatives are listed in Table 1. ChemSpider (http://www.chemspider.com) and Zinc compounds database (http://www.zinc.docking.org) were used to determine the originality of the designed derivatives.

2.1. Molecular properties, bioactivity score and toxicity potential

Molinspiration online tool (http://www.molinspiration.com) was used to calculate various molecular properties such as partition coefficient (log P), Topological polar surface area (TPSA), hydrogen bond donors and acceptors, number of rotatable bonds, number of atoms, and molecular weight. Lipinski's rule [30] was used to predict the oral bioavailability of the designed derivatives. Bioactivity scores for drug targets including GPCR ligands, ion channel modulators, kinase inhibitors, nuclear receptors ligands, protease inhibitors, and enzyme inhibitors, were also determined. Pharmacokinetic parameters such as solubility and toxic properties such as mutagenicity, tumorigenicity, irritant, and reproductive effect were computed using OSIRIS DataWarrior programs [31]. The theoretical prediction of synthetic accessibility (SA) of the ligands was made using Ambit-SA (http://ambit.sourceforge.net/reactor.html) software tool. The model for SA uses four weighted molecular descriptors, which represent different structural and topological features, combined within an additive scheme [32]. The synthetic accessibility score of the ligands L1-L15 was found between the range 64 to 65, suggesting that the ligands are easily synthesizable.

2.2. Prediction of ADMET properties

The pharmacokinetic properties such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) of the designed compounds L1-L15 were determined using the admetSAR prediction tool (http://lmmd.ecust.edu.cn:8000/). The ADMET profile of the compounds was studied by estimation of properties like Human Intestinal Absorption, blood-brain-barrier (BBB) penetration, Caco-2 permeability, CYP inhibitory promiscuity, AMES toxicity, carcinogenicity, and rate acute toxicity LD50.

2.3. Molecular docking

Molecular docking is an important part of the drug design and discovery to identify interactions between amino acid residues of the receptor protein and the ligands. In the present study, Autodock 4.2 software [33] was used to perform molecular docking simulations to estimate the binding affinity of the ligand-receptor complex. The SARS-CoV protease 3C-like protease ($3CL^{pro}$) (Fig. 2) (PDB ID: 6LU7) was used as the receptor. The X-ray crystallographic structure of 3CL protease with a resolution of 2.16 Å was retrieved from the protein data bank [34]. The 3CL Pro-is composed of 306 amino acids, with a molecular weight of 34.51 kDa. The N3 peptidomimetic inhibitor was selected and removed. Docking was performed on chain A containing 2 beta-sheets, 7 beta-

Table 1
The IUPAC names of the designed $3CL^{pro}$ inhibitors (L1-L15).

Sl. No.	Compd No.	R	IUPAC name	Molecular Formula
1	L1	4-fluoro-1H-imidazol-1-yl	1-(4-{3-((4-fluoro-1*H*-imidazol-1-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{22}H_{16}FN_3O_7$
2	L2	2-fluoro-1,3-oxazol-5-yl	1-(4-{3-((2-fluoro-1,3-oxazol-5-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{22}H_{15}FN_2O_8$
3	L3	2-fluoro-1,3-thiazol-5-yl	1-(4-{3-((2-fluoro-1,3-thiazol-5-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{22}H_{15}FN_2O_7S$
4	L4	5-fluoro-2*H*-1,2,3-triazol-4-yl	1-(4-{3-((5-fluoro-2*H*-1,2,3-triazol-4-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{21}H_{15}FN_4O_7$
5	L5	5-fluoro-1,3,4-oxadiazol-2-yl	1-(4-{3-((5-fluoro-1,3,4-oxadiazol-2-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{21}H_{14}FN_3O_8$
6	L6	5-fluoro-1,3,4-thiadiazol-2-yl	1-(4-{3-((5-fluoro-1,3,4-thiadiazol-2-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{21}H_{14}FN_3O_7S$
7	L7	2-fluoro-4,5-dihydrpyrimidin-4-yl	1-(4-{3-((2-fluoro-4,5-dihydrpyrimidin-4-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{23}H_{18}FN_3O_7$
8	L8	2-fluoro-4H-1,3-oxazin-4-yl	1-(4-{3-((2-fluoro-4H-1,3-oxazin-4-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{23}H_{17}FN_2O_8$
9	L9	2-fluoro-2H-1,4-thiazin-3-yl)oxy	1-(4-{3-((2-fluoro-2H-1,4-thiazin-3-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{23}H_{17}FN_2O_7S$
10	L10	3-fluoro-1,2,4-triazin-6-yl	1-(4-{3-((3-fluoro-1,2,4-triazin-6-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{22}H_{15}FN_4O_7$
11	L11	3-fluoropiperazin-2-yl	1-(4-{3-((3-fluoropiperazin-2-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{23}H_{22}FN_3O_7$
12	L12	2-fluoromorpholin-3-yl	1-(4-{3-((2-fluoromorpholin-3-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{23}H_{21}FN_2O_8$
13	L13	2-fluorothiomorpholin-3-yl	1-(4-{3-((2-fluorothiomorpholin-3-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{23}H_{21}FN_2O_7S$
14	L14	3-fluoropiperidin-4-yl	1-(4-{3-((3-fluoropiperidin-4-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{24}H_{23}FN_2O_7$
15	L15	6-fluoro-1,2,4,5-tetrazin-3-yl	1-(4-{3-((6-fluoro-1,2,4,5-tetrazin-3-yl)oxy)−5,7-dihydroxy-4-oxo-4*H*-1-benzopyran-2-yl}−2-hydroxyphenyl)pyrrolidin-2-one	$C_{21}H_{14}FN_5O_7$

Fig. 2. The structure of 3 CL protease (6LU7).

hairpin, 10 helices, and 16 beta turns in its secondary structure. To identify the most probable binding pose, the receptor was kept rigid, and the designed ligands (L1-L15) were set flexible. The results were evaluated based on the binding compatibility, *i.e.* binding energy (kcal/mol) and inhibition constant (μM).

The Lamarckian genetic algorithm (LGA) was employed to model the interactions between the designed ligands (L1-L15) and the receptor with the following settings: a maximum number of 2500,000 energy evaluations, an initial population of 150 randomly placed individuals, a maximum number of 27,000 generations, a maximum mutation rate of 0.02 and crossing over rates of 0.08. The grid box size was set to 120 × 120 × 120 points in x, y, and z directions, and −23.276, 11.233, 54.639 grid centre was set for 3 CL protease receptors. The box was centred based on the cognate ligand with a spacing of 0.375 Å. Cluster analysis was done on the docked results using an RMSD (Root Mean Square Deviation) tolerance of 2 Å. Ten best poses were generated for each ligand and scored using AutoDock 4.2 scoring functions. The schematic two-dimensional representations of the interaction between best docking pose for the selected ligand, and protease receptor 6LU7 were generated using Ligplot+ [35], a computer program.

3. Result and discussion

3.1. Lipinski's rule of five, bioactivity score, drug-likeness, solubility and toxicity

The molecular properties of the designed 3 CL protease inhibitors L1-L15 were calculated using molinspiration cheminformatics software and are presented in Table 2.

Lipinski's rule states that an orally active drug has no more than one violation of the following criteria: (i)An octanol-water partition coefficient log P not greater than 5 (ii) A molecular mass less than 500 daltons (iii) Not more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms) (iv) Not more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms). The test ligands conformed Lipinski's rule of five; however, one violation was observed for ligand L4, L5, L10, and L15. The molecular hydrophobicity or lipophilicity of a molecule was determined by calculating LogP (octanol/water partition coefficient). The LogP value of the designed ligands was found to be well under 5. The positive value of Log P suggested that the designed ligands are lipophilic and permeable through the cell membrane. The topological polar surface area (TPSA) of the designed derivatives was found in the range of 131.35–172.01 Å. A higher TPSA value above the limit of 160 Å was observed for the Ligand L4 and Ligand L15. The number of hydrogen bond acceptors (O and N atoms) was found between 9 and 12, and insignificant violations

Table 2
Prediction of molecular properties descriptors of the title compounds L1-L15.

Sl.No.	Compd No.	LogP	TPSA	natom	MW	nON	nOHN	nVio	nRot	Volume
1	L1	1.94	138.26	33	453.26	10	3	0	4	361.92
2	L2	2.40	146.47	33	454.37	10	3	0	4	358.12
3	L3	3.04	133.33	33	470.43	9	3	0	4	367.26
4	L4	1.96	162.01	33	454.37	11	4	1	4	357.38
5	L5	2.06	159.36	33	455.35	11	3	1	4	353.96
6	L6	2.70	146.22	33	471.42	10	3	0	4	363.10
7	L7	2.18	145.17	34	467.41	10	3	0	4	378.60
8	L8	2.21	142.04	34	468.39	10	3	0	4	374.94
9	L9	2.42	132.80	34	484.46	9	3	0	4	384.09
10	L10	2.06	159.11	34	466.38	11	3	1	4	368.24
11	L11	1.68	144.49	34	471.44	10	5	0	4	390.52
12	L12	1.88	141.70	34	472.43	10	4	0	4	387.10
13	L13	2.42	132.47	34	488.49	9	4	0	4	396.24
14	L14	1.47	132.47	34	470.45	9	4	0	4	394.92
15	L15	1.73	172.01	34	467.37	12	3	1	4	364.08
16	Standard	1.68	131.35	22	302.24	7	5	0	1	240.08

LogP, the logarithm of compound partition coefficient between n-octanol and water; TPSA, topological polar surface area; natom, number of atoms; MW, molecular weight; nON, number of hydrogen bond acceptors; nOHN, number of hydrogen bond donors; nVio, number of violations; Nrotb, number of rotatable bonds.

Table 3
Prediction of bioactivity score of the title compounds L1-L15.

Sl.No.	Compd No.	GPCR	ICM	KI	NRL	PI	EI
1	L1	0.02	−0.20	0.13	−0.04	−0.08	0.11
2	L2	−0.02	−0.24	0.10	0.05	−0.10	0.14
3	L3	−0.02	−0.36	0.23	−0.06	−0.09	0.11
4	L4	0.08	−0.13	0.18	−0.02	−0.06	0.13
5	L5	0.01	−0.21	0.13	−0.06	−0.01	0.14
6	L6	−0.08	−0.29	0.02	−0.05	−0.07	0.19
7	L7	−0.12	−0.30	−0.08	−0.18	−0.14	0.01
8	L8	−0.03	−0.25	−0.06	−0.06	−0.06	0.05
9	L9	−0.10	−0.40	−0.24	−0.10	−0.17	−0.02
10	L10	0.05	−0.22	0.13	0.14	−0.13	0.18
11	L11	−0.06	−0.33	−0.03	−0.15	−0.12	0.03
12	L12	−0.04	−0.34	−0.01	−0.013	−0.10	0.03
13	L13	−0.06	−0.37	−0.07	−0.18	−0.10	0.04
14	L14	0.21	−0.17	0.14	0.04	0.10	0.23
15	L15	−0.00	−0.38	0.04	0.05	−0.18	0.16
16	Standard	−0.06	−0.19	0.28	0.36	−0.25	0.28

GPCR, GPCR ligand; ICM, Ion channel modulator, KI, Kinase inhibitor; NRL, Nuclear receptor ligand; PI, Protease inhibitor, EI, Enzyme inhibitor.

were seen for ligand L4, L5, L10, and L15. The number of hydrogen bond donors (NH and OH) and rotatable bonds (nRB) in the ligands were per the Lipinski's rule of five, *i.e.* less than 5 and 10 respectively.

The results of the bioactivity scores of the designed ligands for the drug targets were calculated by molinspiration and are illustrated in Table 3. A bioactivity score value of more than 0.00 suggests considerable biological activity, while values between −0.50 to 0.00 suggest moderate activity of the compound. The compounds showing a bioactivity score value of less than −0.50 are expected to be inactive. The results indicated considerable to moderate interaction of the ligands with all the drug targets.

Solubility and toxicity assessment of the designed 3CL ^{pro} inhibitors was done using Osiris property explorer, and the results are presented in Table 4.

Drug solubility is an important parameter that significantly affects its absorption and distribution characteristics. The lower value of ClogS indicates higher solubility and better absorption. The designed ligands showed good solubility in comparison to the quercetin standard. The toxicity risk assessment is indicative of a toxicity risk within one of the four major toxicity classes. The results suggested that the tested ligands had a high risk of being tumorigenic but did not possess mutagenic, reproductive, and irritant toxicity.

Table 4
Prediction of solubility and toxicity of the title compounds L1-L15.

Sl.No.	Compd No.	ClogS	Mutagenic	Tumorigenic	Reproductive	Irritant
1	L1	−6.721	None	High	None	None
2	L2	−5.507	None	High	None	None
3	L3	−5.145	None	High	None	None
4	L4	−4.874	None	High	None	None
5	L5	−5.388	None	High	None	None
6	L6	−5.172	None	High	None	None
7	L7	−5.227	None	High	None	None
8	L8	−5.308	None	High	None	None
9	L9	−5.021	None	High	None	None
10	L10	−4.603	None	High	None	None
11	L11	−3.752	None	High	None	None
12	L12	−3.888	None	High	None	None
13	L13	−4.391	None	High	None	None
14	L14	−4.351	None	High	None	None
15	L15	−4.484	None	High	None	None
16	Standard	−2.491	High	High	None	None

Table 5
Prediction of ADMET profile of the title compounds L1-L15.

Sl.No.	Compd No.	Blood-Brain Barrier (BBB)	Human Intestinal Absorption (HIA)	Caco-2 Permeability	CYP Inhibitory Promiscuity	AMES toxicity	Carcinogenicity	Rat Acute ToxicityLD50 mol/kg
1	L1	BBB+	HIA+	Caco2-	High	Toxic	Non- Carcinogenic	2.5621
2	L2	BBB+	HIA+	Caco2-	Low	Non-Toxic	Non- Carcinogenic	2.5129
3	L3	BBB+	HIA+	Caco2-	High	Non- Toxic	Non- Carcinogenic	2.4147
4	L4	BBB-	HIA+	Caco2-	High	Toxic	Non- Carcinogenic	2.4668
5	L5	BBB+	HIA+	Caco2-	Low	Non- Toxic	Non- Carcinogenic	2.4885
6	L6	BBB+	HIA+	Caco2-	High	Non- Toxic	Non- Carcinogenic	2.4385
7	L7	BBB+	HIA+	Caco2-	Low	Non- Toxic	Non- Carcinogenic	2.4664
8	L8	BBB+	HIA+	Caco2-	Low	Non- Toxic	Non- Carcinogenic	2.4984
9	L9	BBB+	HIA+	Caco2-	High	Non- Toxic	Non- Carcinogenic	2.4742
10	L10	BBB+	HIA+	Caco2-	Low	Non- Toxic	Non- Carcinogenic	2.4338
11	L11	BBB+	HIA+	Caco2-	Low	Non- Toxic	Non- Carcinogenic	2.5478
12	L12	BBB+	HIA+	Caco2-	High	Non- Toxic	Non- Carcinogenic	2.4537
13	L13	BBB+	HIA+	Caco2-	High	Non- Toxic	Non- Carcinogenic	2.4557
14	L14	BBB+	HIA+	Caco2-	High	Non- Toxic	Non- Carcinogenic	2.5856
15	L15	BBB+	HIA+	Caco2-	Low	Non-Toxic	Non- Carcinogenic	2.4544
16	Standard	BBB-	HIA+	Caco2+	High	Non- Toxic	Non- Carcinogenic	3.0200

Table 6
Prediction of Binding energy and Inhibition constant of ligands L1-L15.

Sl.No.	Compd No.	Binding Energy (kcal/mol)	Inhibition Constant (μM)
1	L1	−6.60	14.58
2	L2	−7.54	2.97
3	L3	−8.43	0.659
4	L4	−8.84	0.330
5	L5	−8.35	0.753
6	L6	−8.41	0.686
7	L7	−8.41	0.681
8	L8	−8.65	0.456
9	L9	−8.14	1.08
10	L10	−8.29	0.837
11	L11	−7.13	5.97
12	L12	−7.54	2.95
13	L13	−7.93	1.53
14	L14	−8.60	0.50
15	L15	−8.07	1.22
16	Standard	−6.95	8.11

3.2. ADMET properties

The results of the ADMET profile prediction of the designed 3CL-protease inhibitors (L1-L15) are presented in Table 5. The results of various absorption models revealed that the tested ligands had absorption through the blood-brain barrier (except ligand L4), human intestinal absorption but did not show Caco-2 cell permeability. The results strongly support the ability of compounds to act as a drug. The CYP Inhibitory Promiscuity refers to the capacity for a drug or chemical to bind to and decrease or diminish the activity of multiple different CYP450 isoform enzymes [36]. The CYP Inhibitory Promiscuity of the tested ligands was found to be high for most of the ligands in the series except ligands L2, L5, L7, L8, L10, L11, and ligand L15. AMES toxicity test is employed to know whether a compound is mutagenic or not. Similar to the standard, the test ligands displayed non-toxicity (except ligand L1 and L4) which means that the ligands are nonmutagenic. The carcinogenic profile revealed that the designed ligands are non-carcinogenic. The computed LD50 dose of the tested ligands was found between 2.41–2.58 mol/kg. Ligand L3 had the lowest LD50 of 2.41 mol/Kg and was most toxic amongst the test ligands.

3.3. Docking and Ligplot analysis

The molecular docking simulations were performed to validate the anti-SARS-CoV-2 efficacy of the designed derivatives L1-L15 by investigating binding modes and orientation of ligands in the receptor pocket of 3CL ^{pro} target. The results of docking in terms of binding energy and inhibition constants are reported in Table 6.

Most of the ligands showed interactions with residues Ser144, Gly143, Leu141, Tyr54, Asp187, Cys145, Gln189, Thr190, Pro168, Gln192, Met165, His41, Arg188, Phe140, Glu166, Gln189, and Leu27. The lower value of binding energy indicates stronger interactions between ligand and the receptor. All the designed ligands had lower binding energy values than the quercetin standard, which confirmed their excellent affinity against the selected target. The binding energy value for the standard was found to be −6.95 kcal/mol whereas for the tested ligands it was ranged from −8.84 kcal/mol to −6.6 kcal/mol. In the series, the lowest binding energy was observed for ligand L4 (−8.84 kcal/mol). The constant inhibition values were found between 0.330 μM to 14.58 μM. Binding modes of Ligand L4, L8, and L14 with receptor 3CL protease were visualized using Pymol software (http://www.pymol.org/pymol) and are depicted in Fig. 3(a)-(c).

The relative contribution of hydrogen bonds and hydrophobic interactions of the docked ligands L4, L8, and L14 and functional

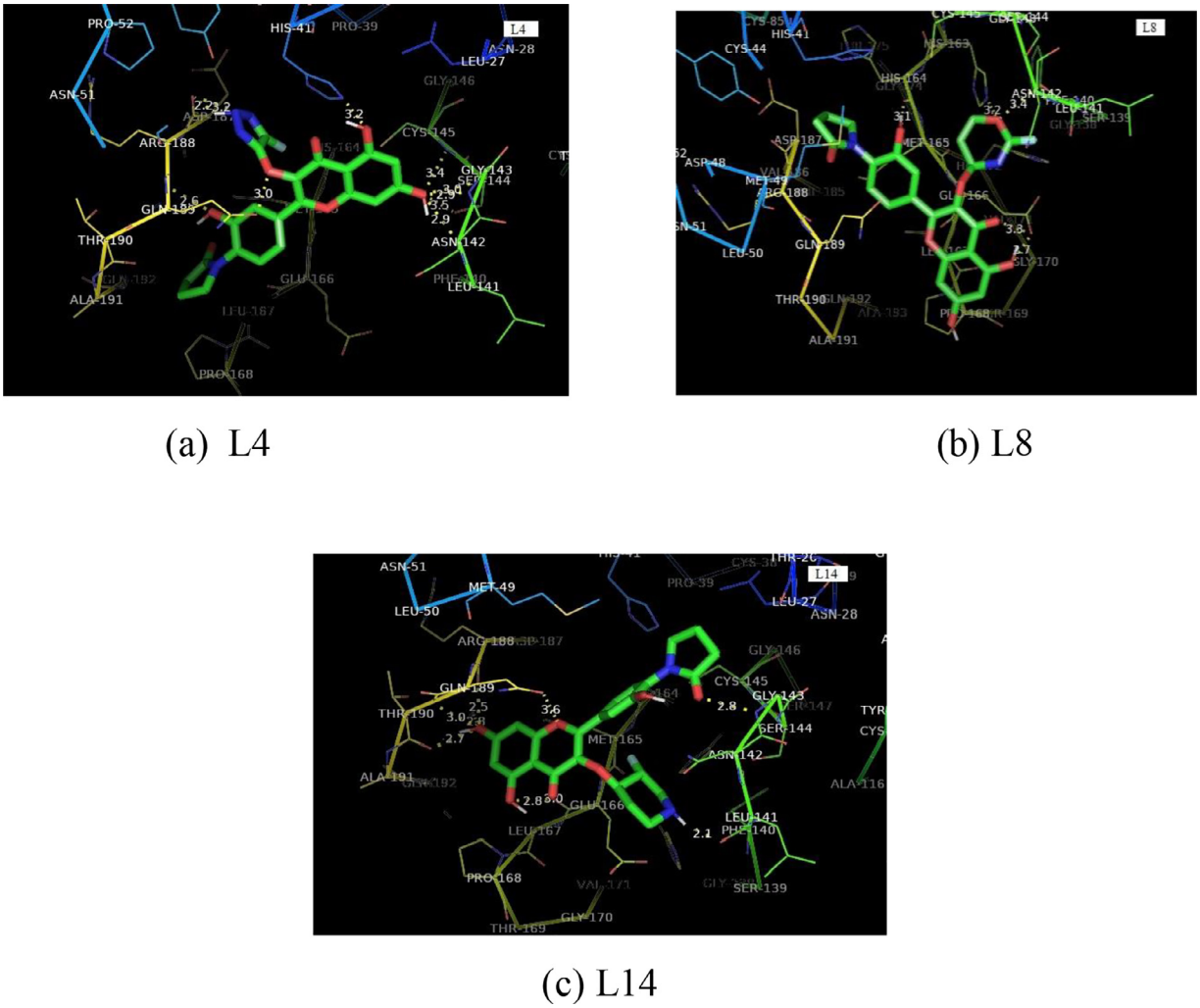

Fig. 3. (a)-(c) Binding modes of ligands L4, L8, and L14 respectively with 3CL protease 6LU7 visualized using Pymol software.

residues in the binding pocket of protease receptor 6LU7 are illustrated in Fig. 4(a)-(c) respectively. Ligand L4 demonstrated five hydrogen bond interactions with active site amino acid residues Ser144, Gly143, Leu141, Tyr54 and Asp187 at a distance of 3.00, 3.53, 3.93, 3.09, and 3.05 Å respectively, and the hydrophobic interactions with Cys145, Gln189, Thr190, Pro168, Gln192, Met165, and His41. Ligand L8 had only one hydrogen bond interaction with residue Glu166 at a distance of 2.66 Å but showed hydrophobic interactions with Pro168, Leu141, Met165, His163, Asp187, His41, and Gln189. Moreover, Ligand L14 found to possess several hydrogen bond interactions with Thr 190 (2.73 and 3.03 Å), Arg188 (2.46 Å), Gln192 (2.85 Å), Phe140 (3.06 Å), Glu166 (2.93 Å), and Gly143 (2.84 Å). The hydrophobic interactions of ligand L14 were observed with residues Met165, Gln189, Cys145, His41, and Leu27. On the other hand, the quercetin standard displayed five hydrogen bond interactions with Tyr54 (2.76 Å), Glu166 (2.81 and 2.97 Å), Gln192 (3.06 Å), and Thr190 (2.63 Å), but got hydrophobic interactions with Asp187, His41, Met165, Leu167, Arg188, Gln189, and Pro168. The results obtained here suggested that the hydrogen bonding and hydrophobic interactions played an important role in stabilizing the ligands in the binding pocket of the target receptor and increased their binding affinity.

4. Conclusion

The world is facing unforeseen challenges in the field of health services and the global economy in the COVID-19 pandemic. Till today there is no report of any clinically approved antiviral drugs or vaccines that are effectual against COVID-19. The design and discovery of novel drugs is the demand of the time to combat this pandemic. Rational drug design using computational tools provide unnumberable benefits such as cost-effectiveness, efficiency, resources, and time- saving. The present investigation proposes flavonoid based 3CL protease inhibitors designed rationally. The designed ligands demonstrated a better binding affinity towards 3 CL protease compared to the quercetin standard. The newly designed ligands confirmed to Lipinski's rule of five and had good bioavailability. Ligands L4, ligand L8 and ligand L14 possessing fluoro substituted triazole, oxazine and piperidine moieties respectively emerged as the lead compounds in the *in-silico* analysis. The compounds displayed significant pharmacokinetic and pharmacodynamic properties. The ADMET profile of the designed ligands was also found satisfactory compared to the standard. The key finding of the present research strengthens the relevance of newly designed flavonoid based 3 CL protease inhibitors as

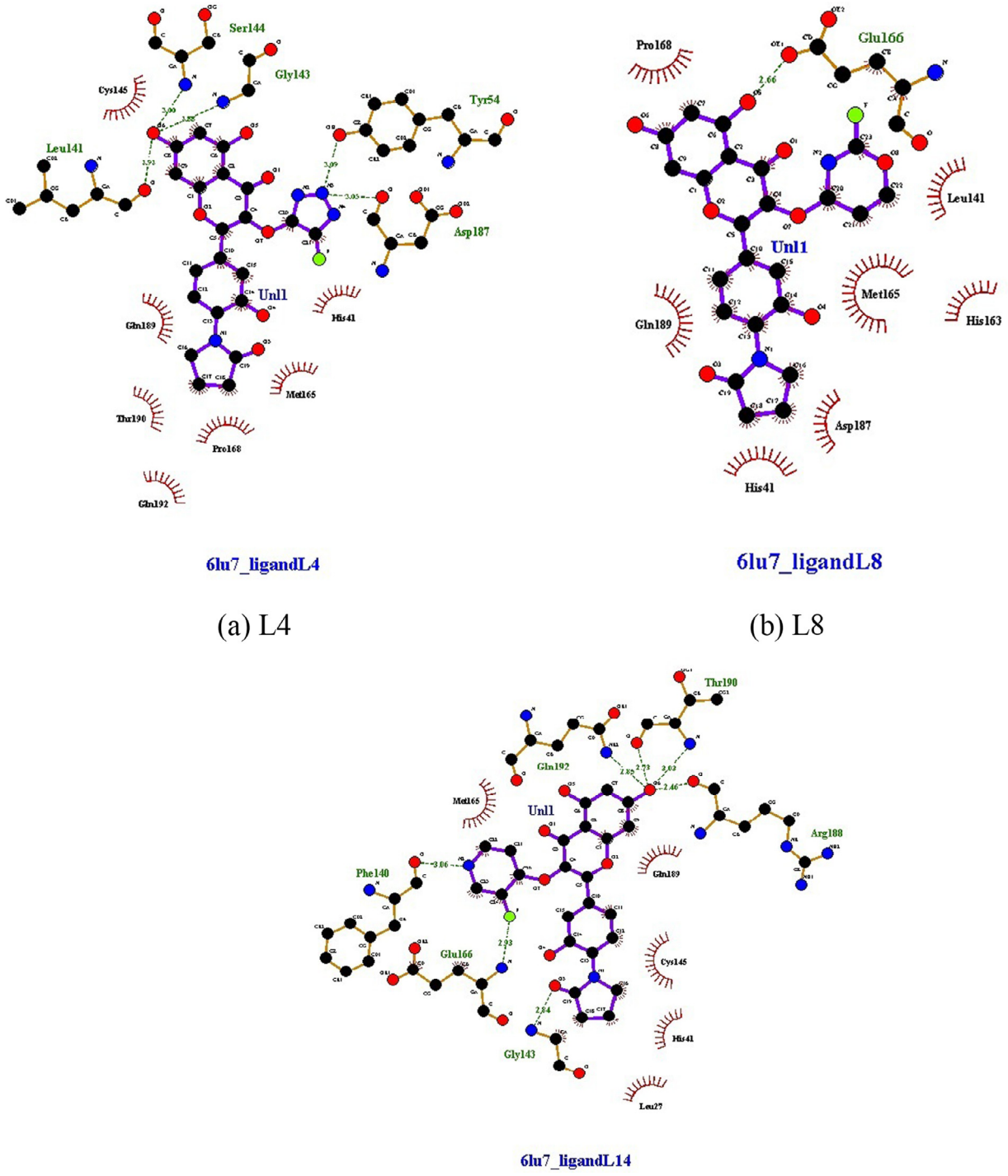

Fig. 4. (a)-(c) Ligplot⁺results showing the interaction of ligands L4, L8, and L14 respectively with 3CL protease 6LU7.

promising drug candidates for the treatment of COVID-19. Further *in vivo* and *in vitro* evaluation study for the designed 3CL protease inhibitors is suggested to validate the computational findings.

Declaration of Competing Interest

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2021.130380.

References

[1] World Health Organization- Coronavirus disease (COVID-19) pandemic, Situation Report-51. https://www.who.int/docs/default-source/coronaviruse/

situation-reports/20200311-sitrep-51-covid-19.pdf (accessed 20 September 2020)
[2] World Health Organization- Coronavirus disease (COVID-19) pandemic, WHO Coronavirus (COVID-19) Dashboard, https://covid19.who.int/ (accessed 22 March 2021)
[3] T. Singhal, A Review of Coronavirus Disease-2019 (COVID-19), Indian J. Pediatr. 87 (2020) 281–286, doi:10.1007/s12098-020-03263-6.
[4] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, et al., Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia, N. Engl. J. Med. 382 (13) (2020) 1199–1207, doi:10.1056/NEJMoa2001316.
[5] A.M. South, D.I. Diz, M.C. Chappell, COVID-19, ACE2, and the cardiovascular consequences, Am. J. Physiol. Heart Circ. Physiol. 318 (5) (2020) H1084–H1090, doi:10.1152/ajpheart.00217.2020.
[6] H. Lu, Drug treatment options for the 2019-new coronavirus (2019-nCoV), Biosci. Trends 14 (1) (2020) 69–71, doi:10.5582/bst.2020.01020.
[7] Z. Jin, J.Y. Liu, R. Feng, L. Ji, Z.L. Jin, H.B. Li, Drug treatment of coronavirus disease 2019 (COVID-19) in China, Eur. J. Pharmacol. 883 (2020) 173326, doi:10.1016/j.ejphar.2020.173326.
[8] S. Bhati, Structure-based drug designing of naphthalene based SARS-CoV PLpro inhibitors for the treatment of COVID-19, Heliyon 6 (11) (2020) e05558, doi:10.1016/j.heliyon.2020.e05558.
[9] A. Zumla, J.F. Chan, E.I. Azhar, D.S. Hui, K.Y. Yuen, Coronaviruses - drug discovery and therapeutic options, Nat. Rev. Drug. Discov. 15 (5) (2016) 327–347, doi:10.1038/nrd.2015.37.
[10] W. Dai, B. Zhang, X.M. Jiang, H. Su, J. Li, Y. Zhao, et al., Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease, Science 368 (6497) (2020) 1331–1335, doi:10.1126/science.abb4489.
[11] K. Anand, G.J. Palm, J.R. Mesters, S.G. Siddell, J. Ziebuhr, R. Hilgenfeld, Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain, EMBO. J. 21 (13) (2002) 3213–3224, doi:10.1093/emboj/cdf327.
[12] Z. Jin, X. Du, Y. Xu, et al., Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, Nature 582 (2020) 289–293, doi:10.1038/s41586-020-2223-y.
[13] K. Anand, H.Yang, M.Bartlam, Z. Rao, R. Hilgenfeld, Coronavirus main proteinase: target for antiviral drug therapy, Coronaviruses with Special Emphasis on First Insights Concerning SARS (2005) 173–199. 10.1007/3-7643-7339-3_9
[14] Y.M. Báez-Santos, S.E. St John, A.D. Mesecar, The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds, Antiviral Res 115 (2015) 21–38, doi:10.1016/j.antiviral.2014.12.015.
[15] S. Boopathi, A.B. Poma, P. Kolandaivel, Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment, J. Biomol. Struct. Dyn. (2020) 1–10, doi:10.1080/07391102.2020.1758788.
[16] R. Lani, P. Hassandarvish, M.H. Shu, W.H. Phoon, J.J. Chu, S. Higgs, D. Vanlandingham, S. Abu Bakar, K. Zandi, Antiviral activity of selected flavonoids against Chikungunya virus, Antiviral Res 133 (2016) 50–61, doi:10.1016/j.antiviral.2016.07.009.
[17] H. Zakaryan, E. Arabyan, A. Oo, K. Zandi, Flavonoids: promising natural compounds against viral infections, Arch. Virol. 162 (2017) 2539–2551, doi:10.1007/s00705-017-3417-y.
[18] P. Ninfali, A. Antonelli, M. Magnani, E.S. Scarpa, Antiviral Properties of Flavonoids and Delivery Strategies, Nutrients 12 (9) (2020) 2534, doi:10.3390/nu12092534.
[19] M. Russo, S. Moccia, C. Spagnuolo, I. Tedesco, G.L. Russo, Roles of flavonoids against coronavirus infection, Chem. Biol. Interact. 328 (2020) 109211, doi:10.1016/j.cbi.2020.109211.
[20] G. Spedding, A. Ratty, E. Middleton Jr., Inhibition of reverse transcriptases by flavonoids, Antiviral Res 12 (2) (1989) 99–110, doi:10.1016/0166-3542(89)90073-9.
[21] K. Shinozuka, Y. Kikuchi, C. Nishino, A. Mori, S. Tawata, Inhibitory effect of flavonoids on DNA-dependent DNA and RNA polymerases, Experientia 44 (1988) 882–885, doi:10.1007/BF01941188.
[22] L. Bachmetov, M. Gal-Tanamy, A. Shapira, M. Vorobeychik, T. Giterman-Galam, P. Sathiyamoorthy, A. Golan-Goldhirsh, I. Benhar, R. Tur-Kaspa, R. Zemel, Suppression of hepatitis C virus by the flavonoid quercetin is mediated by inhibition of NS3 protease activity, J. Viral Hepat. 19 (2) (2012) e81–e88, doi:10.1111/j.1365-2893.2011.01507.x.
[23] W. Wu, R. Li, X. Li, J. He, S. Jiang, S. Liu, J. Yang, Quercetin as an Antiviral Agent Inhibits Influenza a Virus (IAV) Entry, Viruses 8 (1) (2015) 6, doi:10.3390/v8010006.
[24] T.T. Nguyen, H.J. Woo, H.K. Kang, V.D. Nguyen, Y.M. Kim, D.W. Kim, S.A. Ahn, Y. Xia, D. Kim, Flavonoid-mediated inhibition of SARS coronavirus 3C-like protease expressed in Pichia pastoris, Biotechnol. Lett. 34 (5) (2012) 831–838, doi:10.1007/s10529-011-0845-8.
[25] J.Y. Park, H.J. Yuk, H.W. Ryu, S.H. Lim, K.S. Kim, K.H. Park, Y.B. Ryu, W.S. Lee, Evaluation of polyphenols from Broussonetia papyrifera as coronavirus protease inhibitors, J. Enzyme. Inhib. Med. Chem. 32 (1) (2017) 504–515, doi:10.1080/14756366.2016.1265519.
[26] R. Colunga Biancatelli, M. Berrill, J.D. Catravas, P.E. Marik, Quercetin and Vitamin C: an Experimental, Synergistic Therapy for the Prevention and Treatment of SARS-CoV-2 Related Disease (COVID-19), Front. Immunol. 11 (2020) 1451, doi:10.3389/fimmu.2020.01451.
[27] Y. Zhai, X. Zhao, Z. Cui, M. Wang, Y. Wang, L. Li, Q. Sun, X. Yang, D. Zeng, Y. Liu, Y. Sun, Z. Lou, L. Shang, Z. Yin, Cyanohydrin as an Anchoring Group for Potent and Selective Inhibitors of Enterovirus 71 3C Protease, J. Med. Chem. 58 (23) (2015) 9414–9420, doi:10.1021/acs.jmedchem.5b01013.
[28] S. Bhati, V. Kaushik, J. Singh, In Silico Identification of Piperazine Linked Thiohydantoin Derivatives as Novel Androgen Antagonist in Prostate Cancer Treatment, Int. J. Pept. Res. Ther. 25 (2018) 845–860, doi:10.1007/s10989-018-9734-5.
[29] N.M. Ray, R. Singh, J. Singh, S. Bhati, V. Kaushik, Computational screening of Thiohydantoin Derivatives for antitumor activity, Research J. Pharm. Tech 13 (2) (2020) 795–800, doi:10.5958/0974-360X.2020.00150.X.
[30] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings, Adv. Drug Deliv. Rev. 64 (1997) 4–17, doi:10.1016/j.addr.2012.09.019.
[31] T. Sander, J. Freyss, M. Von Korff, C. Rufener, DataWarrior: an open-source program for chemistry aware data visualization and analysis, J. Chem. Inf. Model. 55 (2) (2015) 460–473, doi:10.1021/ci500588j.
[32] N. Kochev, S. Avramova, P. Angelov, N. Jeliazkova, Computational Prediction of Synthetic Accessibility of Organic Molecules with Ambit-Synthetic Accessibility Tool, Org. Chem. Ind. J. 14 (2) (2018) 123.
[33] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function, J. Comput. Chem. 19 (1998) 1639–1662.
[34] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242, doi:10.1093/nar/28.1.235.
[35] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple Ligand-Protein Interaction Diagrams for Drug Discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786, doi:10.1021/ci200227u.
[36] F. Cheng, Y. Yu, Y. Zhou, Z. Shen, W. Xiao, G. Liu, W. Li, P.W. Lee, Y. Tang, Insights into molecular basis of cytochrome p450 inhibitory promiscuity of compounds, J. Chem. Inf. Model 51 (10) (2011) 2482–2495, doi:10.1021/ci200317s.