Available online at www.derpharmachemica.com

ISSN 0975-413X
CODEN (USA): PCHHAX

Der Pharma Chemica, 2020, 12(5): 32-40
(http://www.derpharmachemica.com/archive.html)

Predicted Treatment for COVID-19 Infection: Design, Docking Study and ADMET Prediction of Novel Small Molecules as Serine Protease Inhibitors

Phoebe F. Lamie* and John N. Philoppes

Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Beni-Suef University, Beni-Suef 62514, Egypt

***Corresponding author:** Phoebe F. Lamie, Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Beni-Suef University, Beni-Suef 62514, Egypt. Email: feby.farag@yahoo.com*

ABSTRACT

Nowadays, pandemic coronavirus disease COVID-19, first appeared in Wuhan, China, in December 2019, results in serious global threats to public health, economic and social habits. The aim of this work is to design new compounds with general pharmacophores that mimic seine protease inhibitors (such as camostat mesylate), TMPRSS2 inhibitors, antiviral agents and mucolytic cough suppressant agents. The designed eight compounds were subject to molecular docking study, drug-likeness and in-silico ADMET prediction. Most of the compounds, especially TDa, TDd, TDe, TDf and TDh exhibited good results in docking study inside TMPRSS2 active site and bioavailiability and toxicity evaluation. They could be promising candidates as drug leads for development anti-SARS-Cov-2, the fatal pandemic disease, and therapeutics in the near future.

Keywords: COVID-19, Docking study, ADMET, TMPRSS2, Serine protease inhibitors.

INTRODUCTION

From highly pathogenic viruses that have arisen in the last 40 years, Ebola viruses, Sever Acute Respiratory Syndrome (SARS-coV) and Middle East Respiratory Syndrom (MERS-coV) which appeared in 2012 [1-3]. Until now, there is no vaccines or therapeutic drugs approved from FDA, and their treatments depends on drug overlapping using broad spectrum antiviral drugs [4].

Nowadays, pandemic coronavirus disease COVID-19, first appeared in Wuhan, China, in December 2019, results in serious global threats to public health, economic and social habits [5-7].

COVID-19 (SARS-coV-2), causes acute infection to respiratory system, especially lungs leading to symptoms such as fever, tiredness, dry cough, nasal congestion, sore throat, difficulty breathing and diarrhea [8-10].

COVID-19 is an enveloped virus with a single positive strand RNA genome (+SSRNA). This envelop is covered by a glycosylated fusion spike (S) protein [11-13]. The s-protein consists of two functional domains, S1 and S2 subunits. The receptor binding domain (RBD), S1, is responsible for virus binding to the cell surface receptor, while, S2 subunit contains functional elements involved in membrane fusion [14-16], Figure 1.

SARS-coV-2 targets human angiotensin converting enzyme-2 (ACE-2) receptor enzyme in lungs via its RBD.9 High membrane fusion activity occurs through host cell seine protease enzyme called transmembrane serine protease 2 (TMPRSS2) at two proteolytic cleavage sites, S1/S2 boundary and S2' site which is located within upstream of putative fusion peptide [17].

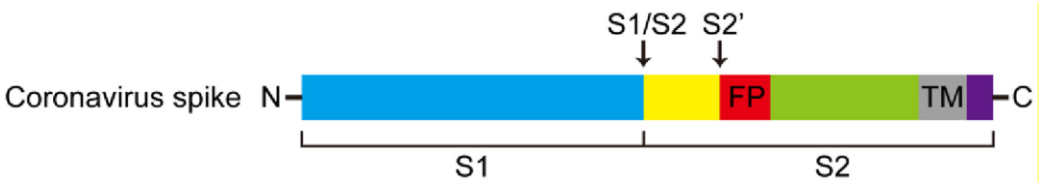

Figure 1: Structure of coronavirus spike protein and its cleavage sites. Arrows=cleavage site; FP=putative fusion peptide; TM=transmembrane domain [17].

This step activates the glycoprotein of COVID-19 and facilitate its entry into the target cell. Once the virus enters cytosol of human host cell, +SSRNA released, then RNA transcription and replication occur. New viral RNA is directed to endoplasmic reticulum, and assembly takes place by the aid of Golgi intermediate. Then viral RNA and differentiated protein fragments assemble and form vesicles which are transported to host cell surface and released by exocytosis process to be ready for attacking new alveoli cells [17]. Figure 2.

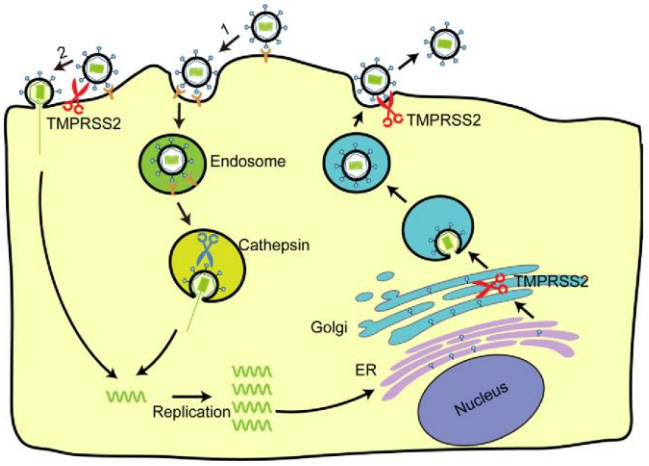

Figure 2: The replication cycle of coronavirus and the role of proteolytic enzyme TMPRSS2 [17].

Development of serine protease inhibitors may help in preventing the entrance of COVID-19 into alveoli cells and block viral infection without affecting on ACE-2 receptor enzyme and its vital activity in treatment of hypertension.

Camostat mesylate (Figure 3), a seine protease inhibitor, showed antiviral activity for SARS-coV infection, that is why we decided to design new serine protease inhibitors aiming at finding new compounds with promising activity in treatment of SARS-coV2 infection [18].

Moreover, compounds (2a&b, Figure 3), 3-amidinophenylalanine derivatives could inhibit TMPRSS2 possessing *Ki* values ≤ 5 nM [19]. In addition, bromohexine hydrochloride (3, Figure 3), a mucolytic cough suppressant, is also effective as TMPRSS2 inhibitor with IC_5 =0.75 µM, in case of influenza virus and coronavirus infections [17].

4-(2-Aminomethyl) benzenesulfonyl fluoride hydrochloride (4, Figure 3), is one of FDA-approved serine protease inhibitors. It showed varying degrees of anti-viral activities [20].

It was also reported that benzamide –derived peptide mimetic derivative (5, Figure 3) exhibited potent TMPRSS2 inhibitory activity against H1N1 influenza virus [21].

Inhibition of serine proteases could be achieved by preparation of peptidyl inverse substrates (6a&b, Figure 3) by esterification of β-amino alcohols with *p*-methoxybenzoic acid. These inhibitors exhibited high specificity to certain seine protease enzymes. Thus Boc-Val and Ac-Val afforeded substrates with specific inhibitory activity to chymotrypsin serine protease enzyme [22]. Guided by the aforementioned reported compounds and from deep inspection to the important pharmacophores in the above reported compounds, we decided to design novel serine protease inhibitors containing:

1- Amide linkages,
2- Urea derivative,
3- Piperidyl as cyclic aliphatic amine,
4- Sulfonyl group,
5- Amino acid residues (valine, leucine, isoleucine and phenyl alanine)

All the designed compounds TDa-h were subjected to molecular docking study as well as drug-likeness and *in*-silico ADMET prediction.

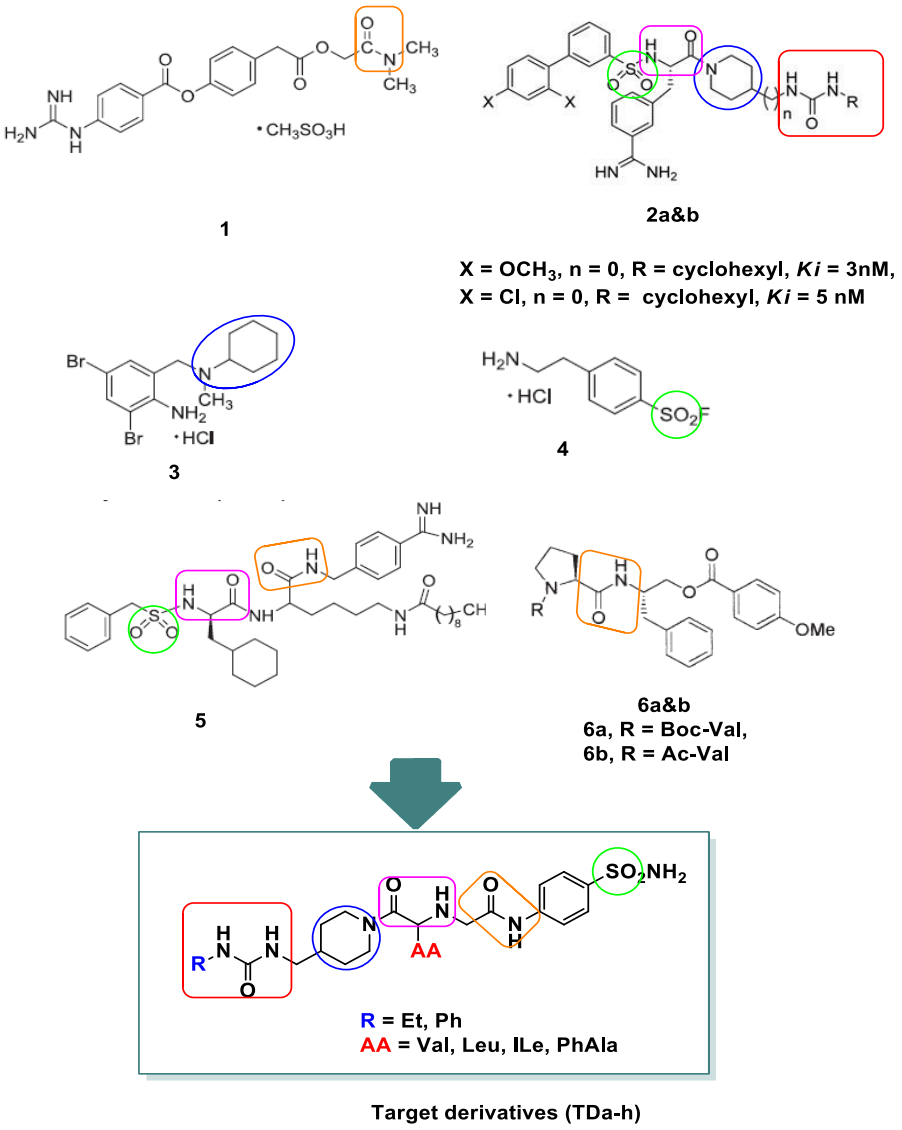

Figure 3: Structure of reported seine protease inhibitors (**1-6**) and a design for the target novel derivatives (TDa-h).

MATERIALS AND METHODS

Molecular docking calculations

All the designed compounds were built using ChemDraw Ulta 12.0. They were energy minimized by using MOE 2014.0901 system. The docking experiment was carried out using TMPRSS2 enzyme (PDB ID: 5CE1) linked to its ligand [23]. Downloaded enzyme and its ligand were prepared for docking by removing water molecules and incomplete residues then hydrogen added and the optimized protein was saved as pdb file, which in turn used for molecular docking studies. Ligand derivative was docked and the best pose and binding score was selected. The lead compound, camostat mesylate as well as the designed compounds was prepared for the docking experiment. The prepared data base was used for docking. The results of the docking scores and H-bonding interactions for the best poses were recorded in Table 1.

Drug-likeness properties

Molinspiration (2018.02 version) [24] was used to predict physicochemical properties and drug-likeness of the designed compounds. Molecular weight (MW), number of hydrogen-bond acceptor (n-ON), number of hydrogen-bond donors (n-OHNH), partition coefficient (logP), number of rotable bonds (n-rotb), topological polar surface area (TPSA) and molecular volum (MV) were calculated. Both acceptable values and predicted results were summarized in Table 2.

In silico ADME properties

In silico ADME properties were calculated using PreADME online surver. Human intestinal absorption (HIA), cell permeability of CaCo-2 cell, skin permeability (SP), blood brain barrier (BBB) and plasma protein binding (PPB) were calculated and predicted results were listed in Table 3**.**

Metabolic prediction

Metabolism prediction for the designed compounds was examined using online preADMET server. The most important metabolic parameter tested was cytochrome P450 (CYP) isoforms, CYP2C19, CYP2C9, CYP2D6 and CYP3A4.

Toxicity prediction

AMES test, rodent carcinogenicity assay and hERG-inhibition were examined to predict toxicity properties for the designed compounds by using preADMET online server [24].

RESULTS AND DISCUSSION

Molecular docking study

To test the ability of the rationalized derivatives as anti-SARS CoV-2 agents, molecular docking study was done using Molecular Operating Environment (MOE 2014.0901) software. The crystal structure of target protein, human serine protease hepsin in complex with its inhibitor, 2-[6-(1-hydroxycyclohexyl)pyridin-2-yl]-1H-indole-5-carboximidamide, was downloaded from Protein Data Bank (PDB ID: 5CE1) in a resolution of 2.50 Å [23].

By docking ligand compound inside enzyme active site, it was found that it forms four hydrogen bonds with His203, Cys349, Gly380 and Asp347 amino acids and docking score = -7.79 Kcal/mol.

The lead compound, camostat mesylate, as TMPRSS2 inhibitor, was drown and docked inside the enzyme. It showed binding score of -7.5383 Kcal/mol and formed 3 H-bonding interactions with Val339, Arg208 and Cys188 amino acids (Figures 4-9).

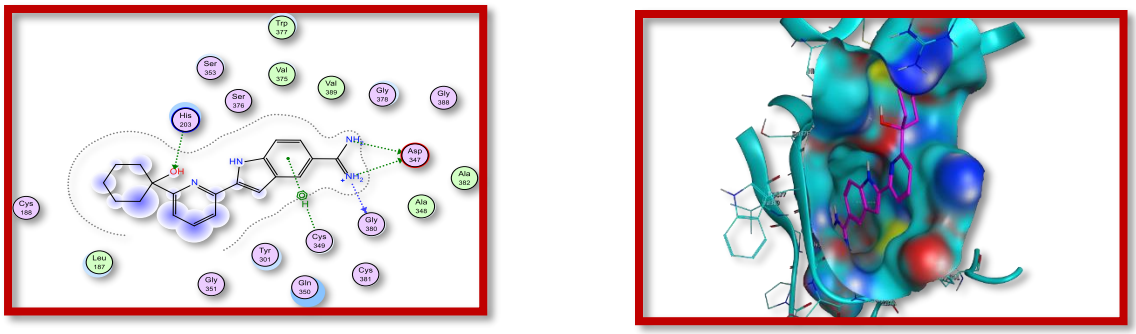

Figure 4: The co-crystallized ligand at the binding site of the target serine protease protein, it forms three H-bonding interactions with Asp347, Gly380 and His203 amino acids, (A) 3D interactions, (B) 2D interactions.

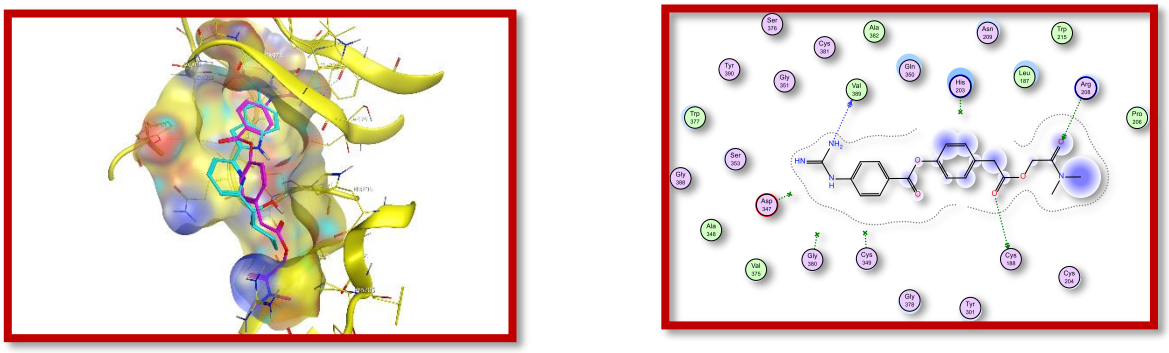

Figure 5: TMPRSS2 inhibitor camostat at the binding site of the target serine protease protein, it forms three H-bonding interactions with Arg208, Cys188 and Val339 amino acids, (A) 3D interactions, cyan is the ligand and magenta represent camostat, (B) 2D interactions.

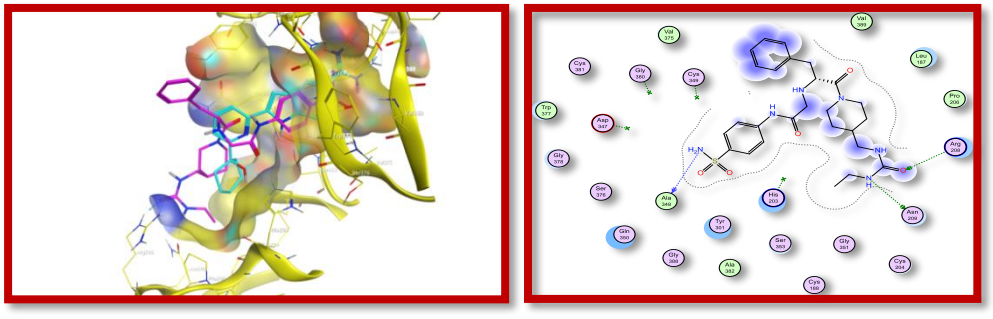

Figure 6: Compound TDd at the binding site of the target serine protease protein, it forms three H-bonding interactions with Ala348, Asn209 and Arg208 amino acids, (A) 3D interactions, cyan is the ligand and magenta represents TDd, (B) 2D interactions.

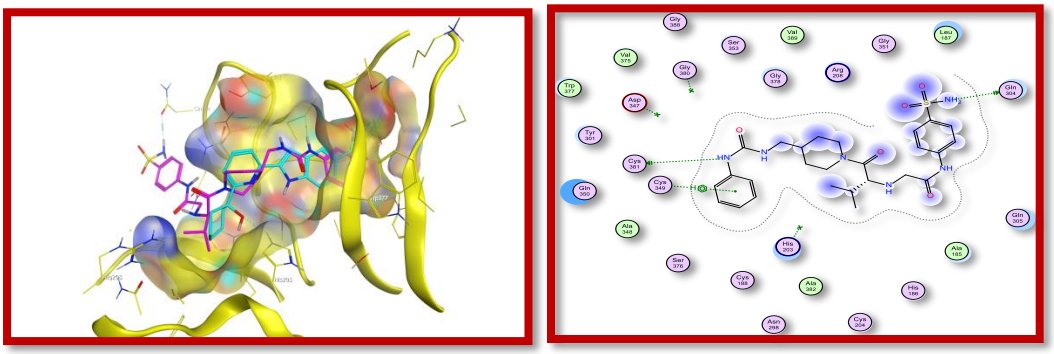

Figure 7: Compound TDe at the binding site of the target serine protease protein, it forms two H-bonding interactions with Cys381 and Cys349 amino acids, (A) 3D interactions, cyan is the ligand and magenta represent TDe, (B) 2D interactions.

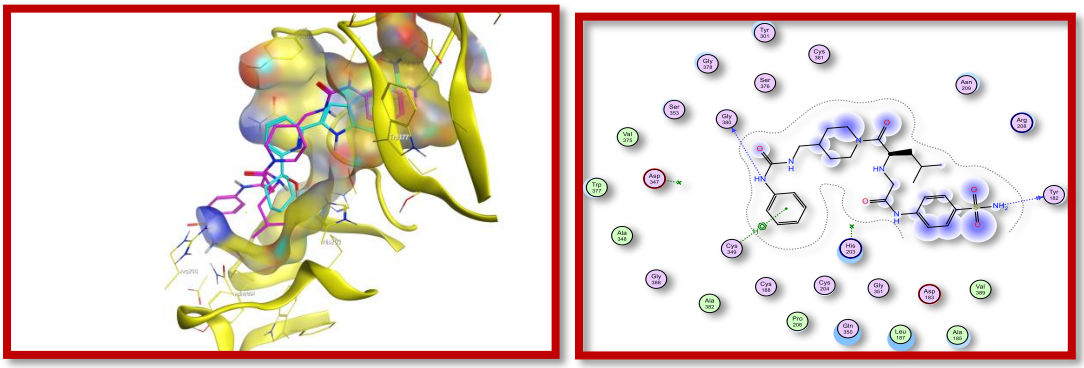

Figure 8: Compound TDf at the binding site of the target serine protease protein, it forms two H-bonding interactions with Thr182 and Gly380 amino acids, (A) 3D interactions, cyan is the ligand and magenta represents TDf, (B) 2D interactions.

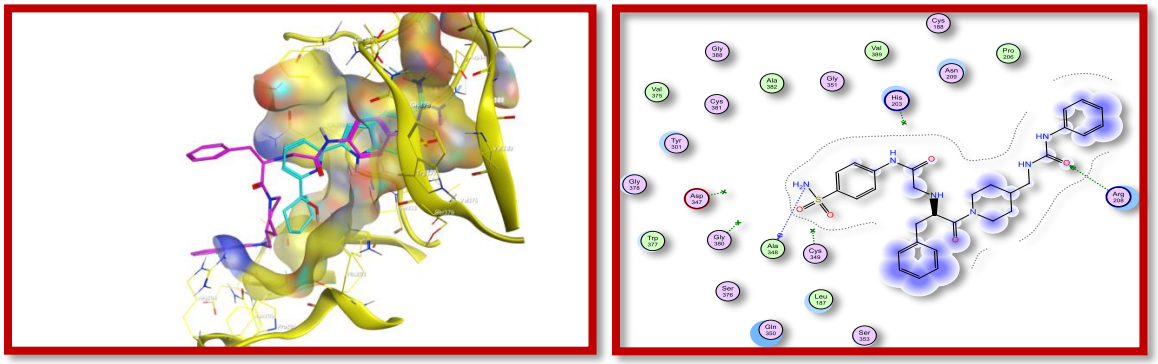

Figure 9: Compound TDh at the binding site of the target serine protease protein, it forms two H-bonding interactions with Arg208 and Ala343 amino acids, (A) 3D interactions, cyan is the ligand and magenta represents TDh, (B) 2D interactions.

Table 1: Results of molecular docking study of designed compounds TDa-h and ligand inside serine protease active site.

Compd. No.	Binding Energy (K cal/mol)	No. of H-bonds	Interacting amino acids
TDa	-6.87224436	1	Ala343
TDb	-7.2998848	3	Ala343
			Cys381
			Gln350
TDc	-6.98205328	1	Ala343
TDd	-6.91680193	3	Ala343

			Asn209
			Arg208
TDe	-8.216	3	Cys381
			Cys349
			Gln304
TDf	-8.04319	2	Gly380
			Tyr182
TDg	-7.70803881	1	Trp377
TDh	-7.01473379	2	Ala343
			Arg208
Camostat	-7.5383	3	Val339
			Arg208
			Cys188
Ligand	-7.7973	4	His203, Cys349, Gly380, Asp347

Drug-likeness properties

To design a drug candidate, it is better to predict its molecular properties beside pharmacokinetic and pharmacodynamics properties. Applying Lipiniski's rule of five helps in prediction of considering the bioactive compound as a drug candidate.

Lipiniski's rule of five possess different physicochemical parameters such as, molecular weight (MW), number of H-bond acceptor (n-ON) and number of H-bond donor (n-OHNH).

Not only is the rule of five the way to state a molecule drug-likeness, but also its cellular permeability, distribution and excretion. These parameters can be predicted through measuring topology polar surface area (TPSA), number of rotable bonds (n-rotb), molecular volum (MV) as well as lipophilic indicator logP "octanol/water" partition coefficient.

Tabulated results (Table 2) showed that most of the designed compounds were in accordance with Lipinski's rule with exception of slightly high in molecular weight as in compounds TDb and TDc (MW = 510), compounds TDd and TDe (MW = 544), compounds TDf and TDg (MW = 558) and TDh (MW=592). All of the compounds had acceptable values for number of H-bond bond acceptors, H-bond bond donors and partition coefficient represented by logP which indicating good membrane permeability. All the analogues had molecular volume within the acceptable range except for compounds TDf, TDg and TDh. Derivatives TDa and TDe had good molecular flexibility with number of rotable bonds of 10. Slight increase in TPSA was observed for all compounds to be 162.72Å instead of 160Å. Finally, the best drug-likeness physicochemical properties were observed in compound TDa.

Table 2: Drug-likeness and physicochemical properties of compounds TDa-h.

Compd.	**MW (kDa)**	***n*-ON**	***n*-OHNH**	**Log*P* (o/w)**	***n*-rotb**	**MV (g/mol)**	**TPSA (Å)**
Acceptable value	≤ 500	≤ 10	≤ 5	≤ 5	≤ 10	< 500	< 160
TDa	496.62	7	5	0.76	10	453.04	162.72
TDb	510.65	7	5	0.91	11	469.85	162.72
TDc	510.65	7	5	0.97	11	469.85	162.72
TDd	544.67	7	5	1.25	11	491.31	162.72
TDe	544.67	7	5	1.5	10	491.09	162.72
TDf	558.69	7	5	1.91	11	507.89	162.72
TDg	558.69	7	5	1.84	11	507.89	162.72
TDh	592.71	7	5	2.02	11	529.35	162.72

MW: Molecular weight; n-ON: Number of hydrogen-bond acceptor; n-OHNH: Number of hydrogen-bond donors; LogP (o/w): Octanol-water partition coefficient; n–rotb: Number of rotable bonds; MV: Molecular volume; TPSA: Topological polar surface area.

***In silico* ADME properties**

Pharmacokinetic parameters are absorption, distribution, metabolism and excretion.

The *in-silico* ADME properties of the designed compounds (Table 3) showed satisfactory results. All the compounds showed good human intestinal absorption (HIA) ranged from 71–88 %. The designed analogs had very low permeability for in-vitro Caco-2 cells ranged from 0.63 to 0.88 and very low skin permeability. Predicted blood-brain barrier indicating whether the compounds can pass across it to be CNS active or can't pass across it and be CNS inactive compounds with no CNS side effects. All the compounds had very low absorption across BBB and couldn't penetrate CNS with values ranged from 0.03–0.04. Only unbound drugs are available for diffusion and transport across cell membrane and in turn interact with their pharmacological targets. The designed derivatives showed weakly bound to plasma protein in a range of 20.87 to 64.13%.

Table 3: Predicted in silico ADME properties for target compounds TDa-h.

Compd.	**Absorption**			**Distribution**	**Excretion**
No.	[a]HIA	[b]CaCo-2	[c]SkinP (logKp)	[d]BBB	[e]PPB
Acceptable value	70-100% (good absorption)	> 70 (high permeability)	> 90 (high permeability)	> 0.40 (CNS active compound)	> 90%
TDa	71.13	0.63	-2.84	0.03	20.87
TDb	73.6	0.64	-2.74	0.03	26.34
TDc	73.6	0.64	-2.73	0.03	26.52
TDd	84.39	0.7	-2.5	0.04	31.68
TDe	84.51	0.81	-2.47	0.04	39.05
TDf	85.22	0.81	-2.41	0.04	45.02
TDg	85.22	0.82	-2.41	0.04	47.35
TDh	88.36	0.88	-2.3	0.04	64.13

[a]Human intestinal absorbtion (%), [b]*in vitro* CaCo cell permeability (nm/sec), [c]Skin permeability, [e]*in-vitro* blood brain barrier penetration (C. brain/C. blood), [e]*in-vitro* plasma protein binding (%).

Metabolic prediction

The most important metabolic parameter is cytochrome P450 (CYP) and its isoforms. *In silico* data indicated that all of the compounds could inhibit CYP3A4. Most of them were inhibitor for CYP2C19 except compounds TDa and TDc. Only compounds TDd and TDh showed CYP2C9 inhinition. None of the tested compounds could inhibit CYP2D6, (Table 4).

Table 4: Metabolic prediction of target compounds TDa-h.

Compound no	**CYP2C19 inhibitor**	**CYP2C9 inhibitor**	**CYP2D6 inhibitor**	**CYP3A4 inhibitor**
TDa	non	non	non	inhibitor
TDb	inhibitor	non	non	inhibitor
TDc	non	non	non	inhibitor
TDd	inhibitor	inhibitor	non	inhibitor
TDe	inhibitor	non	non	inhibitor
TDf	inhibitor	non	non	inhibitor
TDg	inhibitor	non	non	inhibitor
TDh	inhibitor	inhibitor	non	inhibitor

Toxicity prediction

AMES test and carcino-Mouse/Rate tests were used to predict mutagenicity and carcinogenicity for the tested compound, respectively. Moreover, cardiac toxicity of the target synthesized compounds was checked by measuring their hERG-inhibition, (Table 5).

All of the compounds showed non-mutagenic behavior in AMES test except compounds TDc and TDg. All of them had negative carcinogenic effect in both mouse and rat. Additionally, their cardiotoxic effect was not clear. From the above results, it was justified that the target derivatives may have good characters as lead drugs.

Table 5: Toxicity prediction of target compounds TDa-h.

Compound no	AMES	Carcino-Mouse	Carcino-Rat	hERG-inhibition
TDa	non-mutagen	negative	negative	ambiguous
TDb	non-mutagen	negative	negative	ambiguous
TDc	mutagen	negative	negative	ambiguous
TDd	non-mutagen	negative	negative	ambiguous
TDe	non-mutagen	negative	negative	ambiguous
TDf	non-mutagen	negative	negative	ambiguous
TDg	mutagen	negative	negative	ambiguous
TDh	non-mutagen	negative	negative	ambiguous

CONCLUSION

Searching for effective therapeutic agents for treatment of COVID-19 is considered as a major challenge that faces scientific researchers all over the world. Taking into consideration that camostat mesylate is the most acceptable and FDA approved drug as serine protease inhibitor, so designing of new agents mimic it in their chemical structure may be a privilege stone for finding therapeutic agents that help in treatment of this pandemic disease.

All the designed compounds were subjected to molecular docking experiment inside TMPRSS2 active site, four derivatives TDd, TDe, TDf and TDh shared the same amino acids observed from docking of ligand and/or camostat derivatives.

Using compoutational analysis, it eas observed that compound TDa, exhibited good drug-likeness properties (Lipiniski's rule). Most of the compounds showed in silico ADMET predicted values. Hence, the designed compounds could be a promising drug leads for future development of anti-SARS-Cov-2 drugs aiming to stop coronaviruses spread.

Conflict of interest

The authors have no conflict of interest to declare.

REFERENCES

[1] A.R. Fehr, R. Channappanavar and S. Perlman, Annual Review of Medicine. **2017.** 68: p. 387.
[2] J.S.M. Peiris, C.M. Chu, V.C.C. Cheng et al., Lancet. **2003.** 361: p. 1767.
[3] R.J. de Groot, S.C. Baker, R.S. Baric et al., J. Vir. **2013.** 87: p. 7790.
[4] M. Hoffmann, H. Kleine-Weber, S. Schroeder et al., Cell, **2020.** 181(2): p. 271.
[5] C.L. Huang, Y.M. Wang, X.W. Li et al., China. Lancet. **2020.** 395(10223): p. 497.
[6] C. Wang, P.W. Horby, F.G. Hayden et al., Lancet, **2020.** 395(10223): p. 470.
[7] P. Zhou, X.L. Yang, X.G. Wang et al., Nature, **2020.** 579 (7798): p. 270.
[8] E. de Wit, N. van Doremalen, D. Falzarano et al., Nature Reviews Microbiology. **2016.** 14(8): p. 523.
[9] D. Bestle, M.R. Heindl, H. Limburg et al., **2020.**
[10] M. Hoffmann, H. Kleine-Weber, N. Krüger et al., **2020.**
[11] D. Wrapp, N.S. Wang, K.S. Corbett et al., Science, **2020.** 367(6483): p. 1260.
[12] A.C. Walls, Y.J. Park, M.A. Tortorici et al., Cell, **2020.** 181(2): p. 281.
[13] Y.S. Wan, J. Shang, R. Graham et al., J. Vir. **2020.** 94 (7).
[14] W.H. Li, M.J. Moore, N. Vasilieva et al., Nature, **2003.** 426 (6965): p. 450.
[15] J.K. Millet and G.R.Whittaker, Virus Research, **2015.** 202: p. 120.
[16] J. Reguera, G. Mudgal, C. Santiago et al., Virus Research, **2014.** 194: p. 3.
[17] L.W. Shen, H.J. Mao, Y.L. Wu et al., Biochimie, **2017.** 142: p. 1.
[18] Y. Zhou, P. Vedantham, K. Lu et al., Antiviral Research, **2015.** 116: p. 76.
[19] J. Sturzebecher, D. Prasa, J. Hauptmann et al., J. Med. Che. **1997.** 40 (19): p. 3091.
[20] M. Yamamoto, S. Matsuyama, X. Li et al, Antimicrobial Agents and Chemotherapy, **2016.** 60 (11): p. 6532.
[21] E. Bottcher-Friebertshauser, C. Freuer, F. Sielaff et al., J. Vir, **2010.** 84 (11): p. 5605.
[22] B. Walker and J.F. Lynas, Cellular and Molecular Life Sciences, **2001.** 58 (4): p. 596.
[23] https://www.rcsb.org/structure/5CE1.
[24] J.C. Cole, J.W.M. Nissink and R.Taylor. Virtual screening in drug discovery. **2005.**