Volume 65, Issue 5, 2021
Journal of Scientific Research
of
The Banaras Hindu University

Molecular Docking Simulation Studies, Toxicity Study, Bioactivity Prediction, and Structure-Activity Relationship Reveals Rutin as a Potential Inhibitor of SARS-CoV-2 3CL pro

James H. Zothantluanga*[1]

[1]Department of Pharmaceutical Sciences, Faculty of Science and Engineering, Dibrugarh University, Dibrugarh 786004, Assam, India, jameshztta@gmail.com

Abstract: **Coronavirus disease 2019 (COVID-19) pandemic has infected billions and has killed millions of people. Despite the advancement in drugs and vaccines, the COVID-19 pandemic rages on. Bioactive compounds of medicinal plants with antiviral properties such as flavonoids might be useful as phytotherapy for COVID-19. The present study aims to evaluate 16 flavonoids and identify a potential inhibitor for the main protease of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 3CL pro) which is responsible for replication of the viral genome in a host cell. Among the 16 flavonoids, molecular docking simulation studies reveal that rutin has the highest binding affinity towards chain A, chain B, chain C, and chain D of SARS-CoV-2 3CL pro. Rutin formed conventional hydrogen bonds and non-covalent hydrophobic bonds with important amino acid residues at the active binding site of SARS-CoV-2 3CL pro. Also, rutin was computed to be non-toxic. Rutin had a better drug-likeness score and protease inhibition than the co-crystal inhibitor of SARS-CoV-2 3CL pro. The structure-activity relationship revealed important moieties of rutin that might be essential for the inhibition of SARS-CoV-2 3CL pro. The *in-silico* evidence suggests rutin may inhibit SARS-CoV-2 3CL pro and advocates its use as a phytotherapy for COVID-19.**

Index Terms: **COVID-19, Molecular docking, Rutin, SARS-CoV-2, Structure-activity relationship**

I. INTRODUCTION

At present, the world is facing an unprecedented physical, mental and emotional burden due to the ongoing coronavirus disease 2019 (COVID-19) pandemic that is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Hu et al., 2020). As of 11th May 2021, the COVID-19 pandemic has infected 1,58,651,638 people and has killed more than 3.29 million people (WHO, 2021a). Many drugs are currently being used to treat COVID-19 (Singh et al., 2021). Also, many vaccines are currently being used to treat COVID-19 infection (WHO, 2021b).

Despite the recent advancements in pharmacotherapy and vaccines for COVID-19, several countries remained severely affected by COVID-19. Problems such as adverse effects, toxicity, and drug interactions seem to be associated with repurposed drugs that are intended for the treatment of COVID-19 (Shende et al., 2020). Therefore, additional therapies for COVID-19 that are safe and effective seem to be the need of the hour (Shende et al., 2020). Interestingly, bioactive compounds from medicinal plants might be able to offer a safe and effective therapy for COVID-19 (Islam et al., 2020; Antonio et al., 2020). This is because bioactive compounds have been successfully used to treat many viral diseases (Narkhede et al., 2020; Kim et al., 2014; Tahir Ul Qamar et al., 2019). Thus, medicinal plants and their bioactive compounds offer promising phytotherapy for COVID-19 (Narkhede et al., 2020).

The main protease of SARS-CoV-2 is one of the most favorable drug targets to treat COVID-19 (Jin et al., 2020; Prajapat et al., 2020; Robson, 2020). This is because the main protease of SARS-CoV-2 is essential for the replication of viral RNA within the host cell (Jin et al., 2020). Once the viral RNA of SARS-CoV-2 enters a host cell, the polyproteins that are translated from the viral genome are cleaved by the main protease of SARS-CoV-2 which is followed by transcription and replication of the viral genome (Huang et al., 2020; Jin et al., 2020; Gildenhuys, 2020).

Flavonoids have been extensively studied for their antiviral properties (Zakaryan et al., 2017; Ninfali et al., 2020). In the current study, the inhibitory potential of 16 flavonoids against the main protease of SARS-CoV-2 was evaluated with *in-silico*

* Corresponding Author

techniques.

II. Materials and Methods

A. Retrieval of a target protein

The main protease of SARS-CoV-2 was used as the target protein in the study. The crystal structure of SARS-CoV-2 3CL pro with a protein data bank (PDB) ID of 6M2N was retrieved from the Research Collaboratory for Structural Bioinformatics (RCSB)-PDB website (RCSB-PDB, 2021). The main protease of SARS-CoV-2 was found to have four chains (chain A, chain B, chain C, and chain D). The co-crystal inhibitor (CCI) was also identified and retrieved from the PubChem database (PubChem, 2021).

B. Pre-processing of protein

The target protein was pre-processed with BIOVIA Discovery Studio Visualizer v20.1.0.19295 software. All the chains (A, B, C, and D) of the target protein were individually processed. Initially, water was removed from the target protein. Then, the active binding site was defined with the 'Define and edit binding site' feature of the Discovery Studio visualizer software. The 3-dimensional attributes (XYZ coordinates) of the active binding site were identified and saved for further use. Then, the CCI of the target protein was removed. Finally, polar hydrogen was added to the target protein with Discovery Studio visualizer software. The processed target protein was saved in PDB format for future use.

C. Preparation of compound library

The 3-dimensional (3D) structures of 16 flavonoids were retrieved from the PubChem database (PubChem, 2021). The structure of each flavonoid was saved in structure-data file format and their PubChem compound ID (CID) was also saved.

D. Energy minimization of ligands

Energy minimization of ligands was carried out on the PyRx 0.8 tool. To minimize the energy of the ligands used in the present study, the default parameters of the PyRx 0.8 tool was used (Force field = Universal force field; Optimization algorithm = Conjugate gradients; Total no. of steps = 200; No. of steps for update = 1; Stop if energy difference is less than = 0.1).

E. Detection and coordinates of the active binding site

In all the four chains (A, B, C, and D) of SARS-CoV-2 3CL pro, a CCI was found to be present at the active binding site. At first, chain A of the pre-processed target protein that was initially prepared with the Discovery Studio visualizer software was loaded onto the PyRx virtual screening platform. This particular protein was then converted into 'pdbqt.' Format with the PyRx software.

Then, the original target protein complexed with the CCI which did not undergo any processing was also loaded onto the PyRx virtual screening platform. The chains that made up protein complexed with the CCI were revealed by expanding the protein. Initially, chain B, chain C, and, chain D were removed from the scene. Then, chain A of the pre-processed target protein and chain A of the target protein complexed with the CCI was automatically made to superimpose by the PyRx tool.

Following this, the chain A of the target protein complexed with the CCI that was not pre-processed was expanded and the sequence of all the amino acids residues that made up the chain including the CCI (3WL) was revealed. The location of the CCI present at the active binding site of the target protein complexed with the CCI was detected by labelling the atoms of the CCI.

Then, the 3D affinity gird box was carefully made to align with the central part of the CCI to cover the entire active site residues by considering (i) the active binding site coordinates of the pre-processed target protein previously identified with the Discovery Studio visualizer software and (ii) the location of the CCI that was made visible on the protein. In this way, the entire amino acid residues present at the active binding site of the protein were covered by the 3D affinity grid box.

Finally, the chain A of the target protein complexed with the CCI that was not pre-processed was removed from the scene. However, the chain A of the pre-processed target protein was kept for MDSS. The active binding site of chain B, chain C, and, chain D of SARS-CoV-2 3CL pro was detected similarly. The size of the 3D affinity grid box was kept default at 25 Å for the entire simulation process. The coordinates of the active binding site of each chain are described below.

1) Chain A of SARS-CoV-2 3CL pro

In the Vina search space, the coordinates at the active binding site of the protein were x = -33.9203241113, y = -64.4918230137, z = 40.5838306859.

2) Chain B of SARS-CoV-2 3CL pro

In the Vina search space, the coordinates at the active binding site of the protein was x = -48.4728236411, y = -1.05170930781, z = -6.07895092506.

3) Chain C of SARS-CoV-2 3CL pro

In the Vina search space, the coordinates at the active binding site of the protein were x = -40.7113525071, y = -21.7108783003, z = 55.108989351.

4) Chain D of SARS-CoV-2 3CL pro

In the Vina search space, the coordinates at the active binding site of the protein were x = -61.8523694685, y = -35.3575321951, z = 22.9711975727.

F. Molecular docking simulation studies

In the present study, molecular docking simulation studies (MDSS) were carried out for 16 flavonoids and the CCI using AutoDock Vina (Trott & Olson, 2010) on a virtual screening tool known as PyRx 0.8 tool. Other than those specified, the standard operating procedure of PyRx tool was followed for the MDSS (Dallakyan & Olson, 2015).

The flavonoid which showed the highest binding affinity towards chain A, chain B, chain C, and chain D of SARS-CoV-2 3CL pro was selected for further analysis.

G. Visualization and analysis of ligand interactions

The ligand interactions formed by the flavonoid with the amino acid residues at the active binding site of SARS-CoV-2 3CL pro were analyzed with BIOVIA Discovery Studio Visualizer v20.1.0.19295 software. Visualization of the ligand interactions was done in a 2-dimensional (2D) and 3D view. The 3D binding pose of the flavonoid and the CCI at the active binding site of the target protein was generated using PyMOL

molecular graphics system, Version 2.4.1 Schrodinger, LLC software. The ligand interactions between the CCI and SARS-CoV-2 3CL pro were also visualized and analyzed using the same technique as described above.

H. Validation of molecular docking via re-docking, superimposing, and RMSD calculation

The molecular docking protocol used in the present study was validated by a re-docking method (Shivanika et al., 2020). To validate the docking protocol, the CCI was re-docked into the active binding site of each chain of the target protein using the same protocols that were previously used for docking the flavonoids to the target protein. After re-docking was completed, the 'pdbqt.' output of the re-docked ligand was opened with a text document and all the information on the ligand was copied. The 'pdbqt.' output of the target protein was opened with a text document and the information on the re-docked ligand was pasted at the end. This text document was saved for further use. Finally, the re-docked protein-CCI complex was converted to PDB file format from the 'pdbqt.' file format. This process was carried out for all the chains of the target protein.

At the same time, the original protein-CCI complex was processed with the Discovery Studio Visualizer software to obtain chain A complexed with the CCI, chain B complexed with the CCI, chain C complexed with the CCI and chain D complexed with the CCI. It may be noted that all the chains of the target protein are originally complexed with CCI. The original protein-CCI complex (chain A) and the re-docked protein-ligand complex (chain A) were opened simultaneously on the Discovery Studio visualizer software. The sequence alignment of the target proteins was created following which they were made to superimpose on each other. The root mean square deviation (RMSD) between the original protein-CCI complex and the re-docked protein-CCI complex was calculated. Also, the superimposed protein-CCI image was generated with the Discovery Studio Visualizer software. This process was repeated for all the chains of SARS-CoV-2 3CL pro.

I. Toxicity analysis and drug-likeness

The toxicity and drug-likeness of the flavonoid were evaluated with Data Warrior v.5.2.1 software. Toxicities such as mutagenicity, tumorigenic, reproductive effective and irritant of the flavonoid were analyzed. The toxicity and drug-likeness of the CCI were also analyzed.

J. Protease inhibition

The protease inhibition of the flavonoid with the highest binding affinity towards SARS-CoV-2 3CL pro was predicted with the Molinspiration Chemoinformatics web tool (Molinspiration, 2021). The protease inhibition of the CCI was also calculated.

K. Structure-activity relationship

The contribution of chemical groups, atoms, and sugar moiety of the flavonoid towards protease inhibition was studied. The structure of the flavonoid was modified at certain points using Marvin Sketch 20.10 software. The SMILES ID of each modified flavonoid structure was generated and its protease inhibition was evaluated with the Molinspiration Chemoinformatics web tool (Molinspiration, 2021). In this way, a preliminary structure-activity relationship (SAR) study was carried out for rutin.

III. Results and Discussion

A. Features of the target protein

The crystal structure of the SARS-CoV-2 3CL pro is available from the RCSB-PDB website. It is made up of four chains i.e. chain A, chain B, chain C and chain D. The sequence length of SARS-CoV-2 3CL pro is 306. All the chains are complexed with a co-crystal inhibitor *viz*. 3WL. 3WL is a flavonoid known as baicalein and bears a PubChem CID of 5281605. In the present study, 3WL was used as a standard drug whenever and wherever necessary.

B. Details of the flavonoid library

A total of 16 flavonoids were selected to be screened against SARS-CoV-2 3CL pro. The library of flavonoids to be used in the study included apigenin (5280443), aromadendrin (122850), eriodictyol (440735), fisetin (5281614), hesperetin (72281), isorhamnetin (5281654), kaempferol (5280863), luteolin (5280445), myricetin (5281672), naringenin (932), pachypodol (5281677), quercetin (5280343), rhamnazin (5320945), rutin (5280805), tangeretin (68077) and taxifolin (439533). The co-crystal inhibitor that is complexed with the target protein is also a flavonoid.

C. Molecular docking and binding affinities of flavonoids

Molecular docking is a computational technique that predicts possible interactions between a drug and a protein. It gives an idea of the inhibitory potential of a drug against a protein involved in a disease network (Meng et al., 2012). When MDSS is carried out on a PyRx tool, it provides a binding affinity value (kcal/mol) for each ligand so that the binding potential of ligands toward a protein can be ranked (Dallakyan & Olson, 2015).

The binding affinities of all the flavonoids towards chain A, chain B, chain C, and chain D of SARS-CoV-2 3CL pro are given in Table I. The binding affinity of the 3WL is also included in Table I. For each ligand, PyRx tool generates a total of 9 poses at the active binding site of the target protein. A more negative binding affinity value suggests a better binding between a compound and a protein (Dallakyan & Olson, 2015). A low binding affinity value also indicates the low energy requirement for protein-ligand binding (Azam & Abbasi, 2013). In all cases, the first pose is considered the best pose since it has the highest binding affinity towards the target protein. The ninth pose has the lowest binding affinity towards the target protein. Based on this, the first pose and its binding affinity value were considered for the study.

Table I. Binding affinities of the flavonoids and the CCI towards the active binding site of different chains of SARS-CoV-2

Compounds	**Binding affinity (kcal/mol)**			
	Chain A	**Chain B**	**Chain C**	**Chain D**
3WL	-7.3	-7.5	-7.4	-7.3
Apigenin	-7.3	-7.6	-7.1	-7.2

Aromadendrin	-7.5	-7.5	-7.0	-6.9
Eriodictyol	-7.6	-8.2	-7.4	-7.4
Fisetin	-7.7	-8.6	-7.1	-7.3
Hesperetin	-7.6	-8.0	-7.0	-7.1
Isorhamnetin	-7.5	-7.5	-7.2	-7.2
Kaempferol	-7.3	-7.4	-7.1	-6.8
Luteolin	-7.8	-8.6	-7.3	-7.4
Myricetin	-7.6	-8.4	-7.4	-7.6
Naringenin	-7.2	-7.6	-6.6	-7.1
Pachypodol	-7.0	-7.0	-7.3	-7.5
Quercetin	-7.9	-8.7	-7.4	-7.4
Rhamnazin	-7.4	-7.2	-7.1	-7.5
Rutin	-8.5	-9.3	-8.7	-9.7
Tangeretin	-7.1	-6.4	-6.4	-6.9
Taxifolin	-7.6	-8.1	-7.3	-7.1

Among all the flavonoids, rutin has the highest binding affinity towards chain A (-8.5 kcal/mol), chain B (-9.3 kcal/mol), chain C (-8.7 kcal/mol), and chain D (-9.3 kcal/mol) of SARS-CoV-2 3CL pro. This suggests that in comparison to other flavonoids, rutin will bind easily to the active binding site of different chains of the target protein with minimum energy involvement.

3WL is the CCI that is present at the active binding site of all the chains of the target protein. In comparison to 3WL, it was also observed that rutin had a better binding affinity towards all the chains of SARS-CoV-2 3CL pro. Thus, the MDSS study revealed that rutin has the best potential to bind and interact with the amino acid residues at the active binding site of the target protein. Based on this evidence, rutin was selected for further analysis.

D. Molecular interactions between different chains of SARS-CoV-2 3CL pro and rutin

A good understanding of the interaction between a protein and a ligand is important in the field of drug development (Du et al., 2016). The 2D and 3D ligand interactions along with the 3D binding pose of rutin and 3WL at the active binding site of chain A, chain B, chain C, and chain D of SARS-CoV-2 3CL pro are given in Fig. 1, 2, 3, and 4 respectively. The ligand interactions of rutin and 3WL with the amino acid residues at the active binding site of different chains of SARS-CoV-2 3CL pro are given in Table II. Generally, rutin was found to interact with more amino acid residues at the active binding site than 3WL.

Table II. Interacting active site residues of SARS-CoV-2 3CL pro with 3WL and rutin

SARS-CoV-2 3CL pro	**Ligand interactions**	
	3WL	**Rutin**
Chain A	GLY143, GLU166, MET165, CYS145, ASN142, HIS163	MET49, GLU166, ASP187, ARG188, GLN189, THR190, GLN192, HIS41, CYS44
Chain B	LEU141, SER144, CYS145, GLU166, HIS41, CYS44, MET49	HIS41, THR24, CYS44, MET49, CYS145, ASN142
Chain C	ASN142, GLY143, SER144, GLU166, CYS145, MET165	LEU141, SER144, CYS145, MET165, GLU166
Chain D	GLY143, SER144, HIS163, GLU166, CYS145, MET165	THR26, HIS41, TYR54, LEU141, ASN142, GLU166, ASP187, MET49, CYS44, HIS164, MET165, ARG188

In chain A, rutin interacted with 9 amino acid residues (MET49, GLU166, ASP187, ARG188, GLN189, THR190, GLN192, HIS41, and CYS44) while 3WL interacted with 6 amino acid residues (GLY143, GLU166, MET165, CYS145, ASN142, and HIS163) (Fig. 1, Table II). Both rutin and 3WL interacted with GLU166 (Table II). The hydroxyl (-OH) group at position 13 of rutin interacted with MET49 and ASP187 through a conventional hydrogen bond. The -OH at position 19 of rutin interacted with GLU166 through a conventional hydrogen bond. The -OH group at positions 28 and 29 of rutin interacted with ARG188 and GLN192 through conventional hydrogen bonds respectively. The -OH group at position 30 of rutin interacted with GLN189 and THR190 through a conventional hydrogen bond. The A-ring of the C_6-C_3-C_6 scaffold of rutin formed pi-sigma interaction with MET49, pi-pi interaction with HIS41, and pi-sulfur interaction with CYS44. The C-ring of the C_6-C_3-C_6 scaffold of rutin also formed pi-alkyl interaction with MET49 and pi-pi interaction with HIS41.

In chain B, rutin (HIS41, THR24, CYS44, MET49, CYS145, ASN142) and 3WL (LEU141, SER144, CYS145, GLU166, HIS41, CYS44, MET49) interacted with 6 and 7 amino acid residues respectively (Fig. 2, Table II). Both rutin and 3WL interacted with CYS44, HIS41, MET49, and CYS145 (Table II). The -OH group at position 13 and the oxygen atom at position 43 of rutin interacted with HIS41 through a conventional hydrogen bond. The -OH group at position 40 of rutin interacted with THR24 through a conventional hydrogen bond. The A-ring and C-ring of the C_6-C_3-C_6 scaffold of rutin formed pi-pi interaction with HIS41. The A-ring of the C_6-C_3-C_6 scaffold of rutin formed pi-alkyl interaction with MET49 while the B-ring and C-ring of rutin formed pi-alkyl interaction with CYS145. The A-ring of rutin formed pi-sulfur interaction with CYS44.

In chain C, rutin interacted with 5 amino acid residues (LEU141, SER144, CYS145, MET165, GLU166) while 3WL interacted with 6 amino acid residues (ASN142, GLY143, SER144, GLU166, CYS145, MET165) (Fig. 3, Table II). Both rutin and 3WL interacted with SER144, CYS145, MET165, and GLU166 (Table II). The -OH group at position 19 of rutin interacted with SER144 through a conventional hydrogen bond. The -OH group at position 20 of rutin interacted with LEU141 through a conventional hydrogen bond. The B-ring and C-ring of the C_6-C_3-C_6 scaffold of rutin formed pi-alkyl interaction with CYS145 and MET165 respectively.

In addition to hydrogen bonds, non-covalent bonds such as hydrophobic interactions and electrostatic interactions are also considered important for the stabilization of protein-ligand complexes (de Freitas & Schapira, 2017; Kumar & Nussinov, 2002). Rutin showed non-covalent interactions (pi-pi, pi-alkyl, and pi-sulfur) with different amino acid residues at the active binding site of different chains of SARS-CoV-2 3CL pro (Fig. 1,

Fig. 2, Fig. 3 and Fig. 4). Thus, rutin might be able to inhibit SARS-CoV-2 3CL pro as it was found to have a good binding affinity towards all the different chains. In comparison to 3WL, rutin showed better molecular interactions with different amino acid residues at the active binding site of SARS-CoV-2 3CL pro. Based on the above findings, rutin was subjected to further analysis.

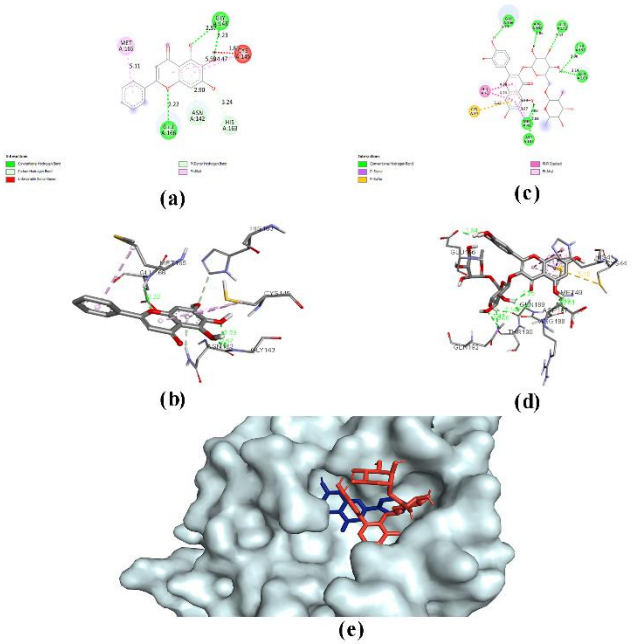


Fig. 1. Chain A: 2D ligand interactions (a, c), 3D ligand interactions (b, d), and 3D binding pose (e) (blue=3WL, red=Rutin) of 3WL (a, b, e) and rutin (c, d, e) at the active binding site of SARS-CoV-2 3CL pro.

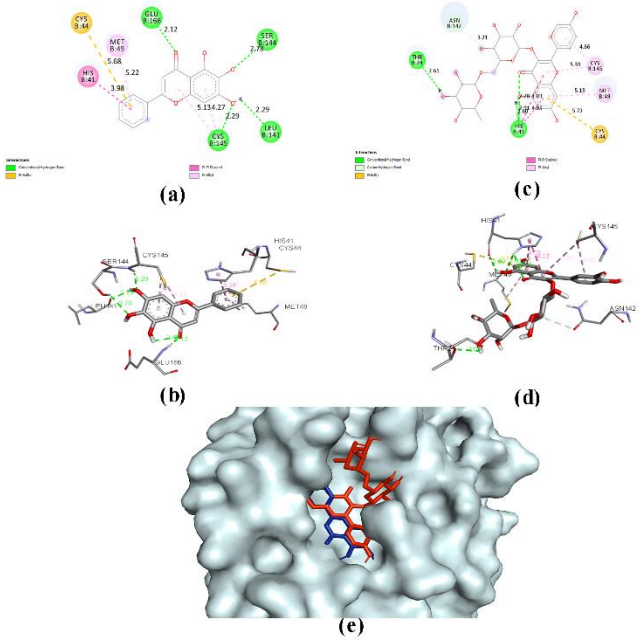


Fig. 2. Chain B: 2D ligand interactions (a, c), 3D ligand interactions (b, d), and 3D binding pose (e) (blue=3WL, red=Rutin) of 3WL (a, b, e) and rutin (c, d, e) at the active binding site of SARS-CoV-2 3CL pro.

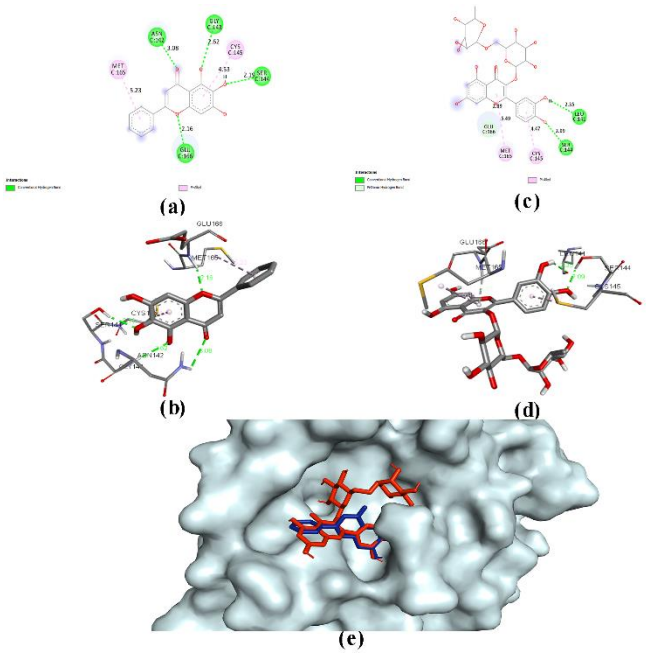


Fig. 3. Chain C: 2D ligand interactions (a, c), 3D ligand interactions (b, d), and 3D binding pose (e) (blue=3WL, red=Rutin) of 3WL (a, b, e) and rutin (c, d, e) at the active binding site of SARS-CoV-2 3CL pro.

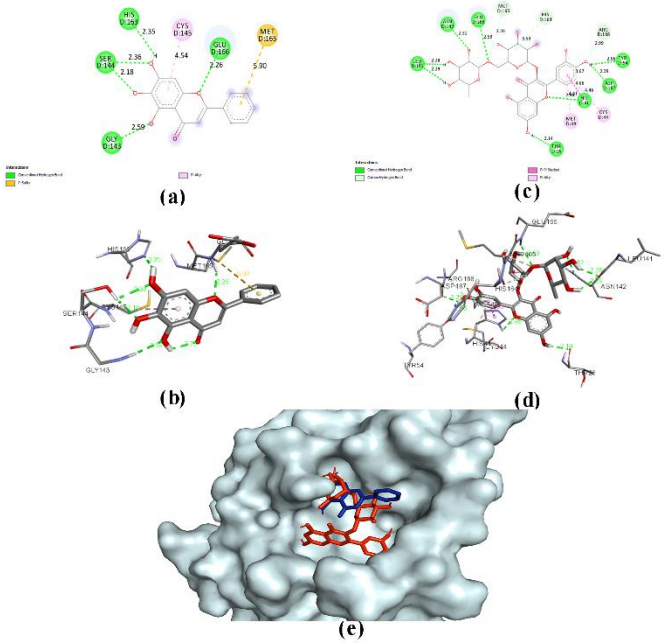


Fig. 4. Chain D: 2D ligand interactions (a, c), 3D ligand interactions (b, d), and 3D binding pose (e) (blue=3WL, red=Rutin) of 3WL (a, b, e) and rutin (c, d, e) at the active binding site of SARS-CoV-2 3CL pro.

E. Validation of docking

The purpose of this study was to evaluate the efficiency of the docking protocol used in the present study. The docking protocol was validated through re-docking of the CCI to the active binding site of the target protein using the same parameters that were used to dock the flavonoids. The re-docked protein-ligand complex was then superimposed to the original protein-ligand complex. The superimposed image of the protein-ligand complex is given in Fig. 5.

The RMSD of the re-docked protein-ligand complex was calculated by using the original protein-ligand complex as the reference protein. Several researchers had also validated their docking protocol with this method (Shivanika et al. 2020). A low RMSD value is preferred over a higher value to provide validation to the docking protocol.

In the present study, an RMSD value of 0.00 was obtained for all the docking procedures used for docking the CCI towards the active binding site of chain A, chain B, chain C, and chain D. Re-docking showed that the CCI had a binding affinity of -7.3 kcal/mol, -7.5 kcal/mol, -7.4 kcal/mol, and -7.3 kcal/mol towards chain A, chain B, chain C, and chain D respectively. The binding affinity values obtained from re-docking are similar to the value previously obtained in the first docking (Table I). Also, the docking protocol used for the present study allowed the binding of CCI into the active binding site of the target protein as evidence by re-docking (Fig. 5). This proves the efficiency and validity of the docking protocol followed in the present study.

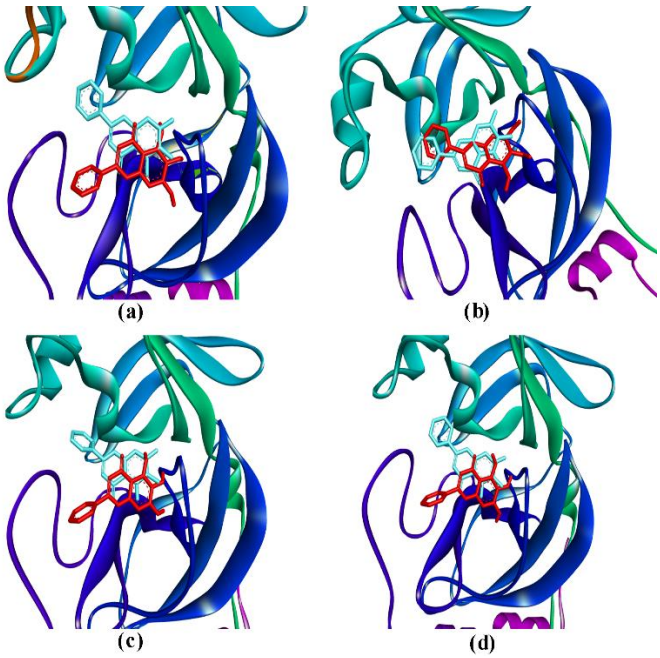


Fig. 5. Superimposition of re-docked protein-3WL complex (3WL=red color) with the original co-crystallized complex (3WL=blue color) for chain A (a), chain B (b), chain C (c) and chain D (d) of SARS-CoV-2 3CL pro.

F. Toxicity analysis and drug-likeness

Toxicity is one of the main reasons behind the withdrawal of many drugs from the market (Guengerich, 2011). Drug-likeness is a parameter that indicates the possibility of a molecule to become an orally bioavailable drug (Daina et al., 2017). Therefore, it may be considered logical to assess the potential toxicity and drug-likeness of a compound at the initial stage of a study. The potential mutagenicity, tumorigenic, reproductive effective and irritant of rutin and 3WL are given in Table III. The drug-likeness of rutin and 3WL are also included in Table III. It was observed that rutin was free from all possible toxicities. Also, the drug-likeness of rutin was higher than 3WL. Based on these findings, rutin was subjected to further analysis.

Table III. Toxicity and drug-likeness of 3WL and rutin

Property	**3WL**	**Rutin**
Mutagenicity	None	None
Tumorigenic	None	None
Reproductive effective	None	None
Irritant	None	None
Drug-likeness	0.28194	1.9337

G. Protease inhibition

The protease inhibition of rutin was predicted and it was compared to that of 3WL (Table IV). SARS-CoV-2 3CL pro is a viral protease (Prajapat et al., 2020). Although the predicted protease inhibition might not necessarily correlate to the inhibition of the main protease of SARS-CoV-2, it provides a preliminary idea of the potential bioactivity of rutin. With a bioactivity score of -0.07, rutin has a better protease inhibition than 3WL (-0.35). Based on these findings, the structure-activity relationship of rutin concerning its protease inhibition was studied.

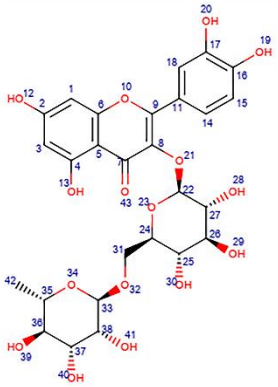
Table IV. Protease inhibitory potential of 3WL and rutin

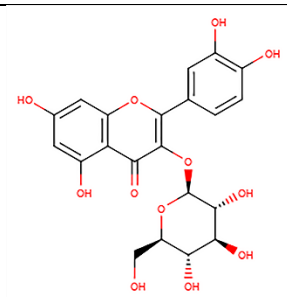
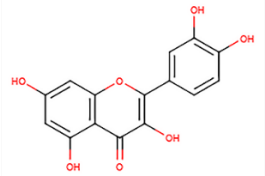
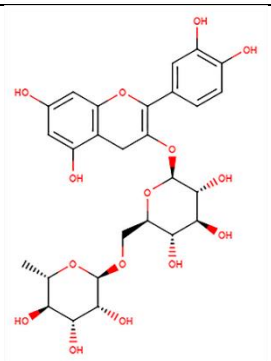
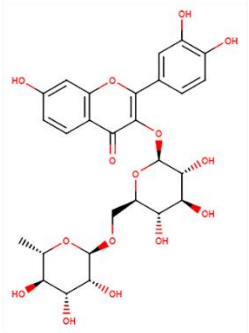
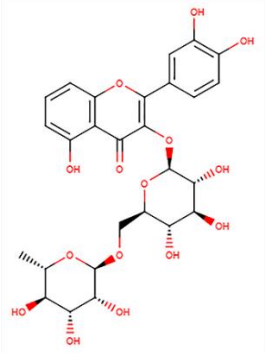
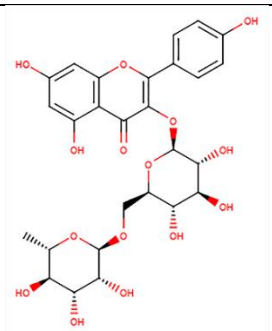
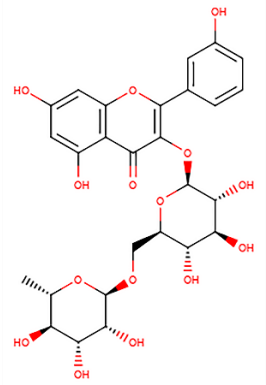
Compound	**Protease inhibition**
3WL	-0.35
Rutin	-0.07

H. Structure-activity relationship

Preliminary SAR studies can be used to identify potential chemical groups that might be responsible for eliciting the bioactivity of a compound (McKinney et al., 2000). The SAR of rutin concerning its protease inhibition was studied with the Molinspiration Chemoinformatics web tool. The contribution of different atoms, groups, and the sugar moiety towards protease inhibition of rutin are given as numerical value in Table V.

Table V. SAR of rutin

Structure No.	**Structure**	**Protease inhibition**
1 (Rutin)		-0.07

2		-0.06
3		-0.25
4		0.05
5		-0.08
6		-0.03
7		-0.04
8		-0.05

From structures no. 2 and 3 (Table V), the sugar moiety of rutin might be necessary for its protease inhibition activity. Also, the -OH groups of the sugar moiety of rutin were significantly involved in the formation of conventional hydrogen bonds with the amino acid residues at the active binding site of different chains of SARS-CoV-2 3CL pro (Fig. 1, 2, and 4). The removal of an oxygen atom at position 43 attached to the C-ring of rutin significantly increases the protease inhibition activity (Table V, Structure no. 4). Also, the oxygen atom at position 43 attached to the C-ring of rutin was not involved in any molecular interactions (Fig. 1, 3, and 4). In structures no. 5 to 8 (Table V), the removal of -OH groups from different positions does not seem to have a positive effect on the protease inhibition activity of rutin. The aglycone moiety (C_6-C_3-C_6 scaffold) was actively involved in the formation of different types of protein-ligand interactions (Fig. 1, 2, 3, and 4).

Based on these findings, it may be assumed that the sugar moiety and the aglycone moiety of rutin are essential for binding to the active binding site of SARS-CoV-2 3CL pro. The sugar moiety of rutin may also be essential for the protease inhibition activity of rutin (Table V). On the other hand, the oxygen atom at position 43 attached to the C-ring of rutin does not seem to contribute to the formation and stabilization of the protein-ligand complex (Fig. 1, 2, 3, and 4).

Interestingly, molecular docking and molecular dynamics simulation studies also showed rutin as a potential inhibitor of SARS-CoV-2 3CL pro (Cherrak et al., 2020). In the study, chain A of SARS-CoV-2 3CL pro was used. Cherrak et al. (2020) showed that the sugar moiety of rutin formed two hydrogen bonds with GLU166. Cherrak et al. (2020) also showed that the A-ring of rutin formed hydrophobic interactions with MET49. Similarly, in our study, rutin formed a conventional hydrogen bond with GLU166. Rutin also formed conventional hydrogen

bonds and non-covalent interactions with MET49 of chain A of SARS-CoV-2 3CL pro. In addition, our study also provides data for chain B, chain C, and chain D.

Rutin is a citrus flavonoid (Kreft et al., 1997) that is abundantly found in apples, tea, buckwheat, and passionflowers (Harborne, 1986). Rutin exerts several pharmacological activities such as anticarcinogenic, antioxidant, vasoprotective, cytoprotective, cardioprotective, and neuroprotective activities (Ganeshpurkar & Saluja, 2017). An analog of rutin known as sodium rutin sulfate showed anti-retroviral activity against various strains of human immunodeficiency virus 1 (Tao et al., 2007). Rutin also showed antiviral activity against vesicular stomatitis virus (Wacke & Eilmes, 1978), canine distemper virus (Carvalho et al., 2013), and avian influenza (H5N1) virus (Ibrahim et al., 2013). The literature review revealed the potent anti-viral activity of rutin against different pathogens of viral origin. Considering the anti-viral activity of rutin against other RNA viruses, there is a possibility that rutin might elicit anti-viral activity against SARS-CoV-2 by inhibiting its main protease (3CL pro) which is a key component for viral replication.

Conclusion

The present *in-silico* study revealed rutin as a potential inhibitor of SARS-CoV-2 3CL pro. Rutin showed a good binding affinity at the active binding pocket of all the four chains of SARS-CoV-2 3CL pro. Rutin also showed better ligand interactions with SARS-CoV-2 3CL pro than 3WL. In many cases, rutin and 3WL showed similar molecular interactions with the same amino acid residues at the active binding site of SARS-CoV-2 3CL pro. Rutin formed more conventional hydrogen bonding than 3WL. Rutin is non-toxic and had a better drug-likeness score than 3WL. Also, rutin was predicted to be a better protease inhibitor than 3WL. SAR revealed that the aglycone moiety and the sugar moiety of rutin might be essential for its protease inhibition activity. SAR also revealed that the oxygen atom at position 43 of rutin does not contribute to protein-ligand binding and protease inhibition activity. Therefore, the present *in-silico* study advocates rutin as a potential inhibitor of SARS-CoV-2 3CL pro and as promising phytotherapy for COVID-19. The C_6-C_3-C_6 scaffold of rutin may also be used to design and develop potent anti-viral agents against SARS-CoV-2.

Acknowledgments

The author acknowledges the Department of Pharmaceutical Sciences, Dibrugarh University for providing the necessary facilities to carry out the *in-silico* work.

References

Antonio, A.D.S., Wiedemann, L.S.M., & Veiga-Junior, V.F. (2020). Natural products' role against COVID-19. RSC Advances, 10, 23379-23393.

Azam, S.S., & Abbasi, S.W. (2013). Molecular docking studies for the identification of melatoninergic inhibitors for acetylserotonin-O-transferase using different docking routines. Theoretical biology and medical modeling, 10, 63.

Carvalho, O.V., Botelho, C.V., Ferreira, C.G., Ferreira, H.C., Santos, M.R., Diaz, M.A., Oliveira, T.T., Soares-Martins, J.A., Almeida, M.R., & Silva, A. (2013). *In vitro* inhibition of canine distemper virus by flavonoids and phenolic acids: implications of structural differences for antiviral design. Research in Veterinary Science, 95(2), 717-724.

Cherrak, S.A., Merzouk, H., Mokhtari-Soulimane, N. (2020). Potential bioactive glycosylated flavonoids as SARS-CoV-2 main protease inhibitors: A molecular docking and simulation studies. PLoS ONE, 15(10), e0240653.

Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S.M., & Savidge, T.C. (2016). Regulation of protein-ligand binding affinity by hydrogen bond pairing. Science Advances, 2, e1501240.

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports, 7, 42717.

Dallakyan, S., & Olson, A.J. (2015). Small-molecule library screening by docking with PyRx. Methods in Molecular Biology, 1263, 243-250.

de Freitas, R.F., & Schapira, M. (2017). A systematic analysis of atomic protein-ligand interactions in the PDB. Medicinal Chemistry Communications, 8, 1970-1981.

Du, X., Li, Y., Xia, Y.L., Ai, S.M., Liang, J., Sang, P., Ji, X.L., & Liu, S.Q. (2016). Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. International Journal of Molecular Sciences, 17(2), 144.

Ganeshpurkar, A., & Saluja, A.K. (2016). The Pharmacological Potential of Rutin. Saudi Pharmaceutical Journal, 25(2), 149-164.

Gildenhuys, S. (2020). Expanding our understanding of the role polyprotein conformation plays in the coronavirus life cycle. Biochemical Journal, 477, 1479-1482.

Guengerich, F.P. (2011). Mechanisms of drug toxicity and relevance to pharmaceutical development. Drug Metabolism and Pharmacokinetics, 26, 3-14.

Harborne, J.B. (1986). Nature, distribution and function of plant flavonoids. Progress in Clinical and Biological Research, 213, 15-24.

Hu, B., Guo, H., Zhou, P., & Shi, Z.L. (2021). Characteristics of SARS-CoV-2 and COVID-19. Nature Reviews Microbiology, 19(3), 141-154.

Huang, Y., Yang, C., Xu, X., Xu, W., & Liu, S. (2020). Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacologica Sinica, 41, 1141-1149.

Ibrahim A.K., Youssef A.I., Arafa A.S., & Ahmed S.A. (2013). Anti-H5N1 virus flavonoids from *Capparis sinaica* Veill. Natural Product Research, 27(22), 2149-2153.

Islam, M.T., Sarkar, C., El-Kersh, D.M., Jamaddar, S., Uddin, S.J., Shilpi, J.A., & Mubarak, M.S. (2020). Natural products and their derivative against coronavirus: A review of the non-

clinical and pre-clinical data. Phytotherapy Research, 34, 2471-2492.

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. (2020). Structure of M_{pro} from SARS-CoV-2 and discovery of its inhibitors. Nature, 582, 289-293.

Kim, D.W., Seo, K.H., Curtis-Long, M.J., Oh, K.Y., Oh, J.W., Cho, J.K., Lee, K.H., & Park, K.H. (2014). Phenolic phytochemical displaying SARS-CoV papain-like protease inhibition from the seeds of *Psoralea corylifolia*. Journal of Enzyme Inhibition and Medicinal Chemistry, 29, 59-63.

Kumar, S., & Nussinov, R. (2002). Close-range electrostatic interactions in proteins. Chembiochem, 3, 604-617.

Kreft, S., Knapp, M., & Kreft, I. (1997). Extraction of rutin from buckwheat (*Fagopyrum esculentum* Moench) seeds and determination by capillary electrophoresis. Journal of Agricultural and Food Chemistry, 47(11), 4649-4652.

McKinney, J.D., Richard, A., Waller, C., Newman, M.C., & Gerberick, F. (2000). The Practice of Structure Activity Relationships (SAR) in Toxicology. Toxicological Sciences, 56(1), 8-17.

Meng, X.Y., Zhang, H.X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. Current computer-aided drug design, 7(2), 146-157.

Molinspiration. (2021). Calculation of molecular properties and bioactivity score. Retrieved May 7, 2021, from https://www.molinspiration.com/cgi-bin/properties

Narkhede, R.R., Pise, A.V., Cheke, R.S., Shinde, S.D. (2020). Recognition of natural products as potential inhibitors of COVID-19 main protease (Mpro): *In-Silico* evidences. Natural Products and Bioprospecting, 10, 297-306.

Ninfali, P., Antonelli, A., Magnani, M., & Scarpa, E.S. (2020). Antiviral Properties of Flavonoids and Delivery Strategies. Nutrients, 12(9), 2534.

PubChem. (2021). Retrieved May 5, 2021, from https://pubchem.ncbi.nlm.nih.gov/

Prajapat, M., Sarma, P., Shekhar, N., Avti, P., Sinha, S., Kaur, H., Kumar, S., Bhattacharyya, A., Kumar, H., Bansal, S., & Medhi, B. (2020). Drug targets for corona virus: A systematic review. Indian Journal of Pharmacology, 52, 56-65.

Robson, B. (2020). Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus. Computers in Biology and Medicine. 119, 103670.

RCSB-PDB. (2021). SARS-CoV-2 3CL protease (3CL pro) in complex with a novel inhibitor. Retrieved May 5, 2021, from https://www.rcsb.org/structure/6M2N

Shivanika, C., Deepak Kumar, S., Ragunathan, V., Tiwari, P., Sumitha, A., Brindha Devi, P. (2020). Molecular docking, validation, dynamics simulations, and pharmacokinetic prediction of natural compounds against the SARS-CoV-2 main-protease. Journal of Biomolecular Structure and Dynamics, 8:1-27.

Shende, P., Khanolkar, B., Gaud, R.S. (2020). Drug repurposing: new strategies for addressing COVID-19 outbreak. Expert Review of Anti Infective Therapy, 3, 1-18.

Singh, S.P., Pritam, M., Pandey, B., & Yadav, T.P. (2021). Microstructure, pathophysiology, and potential therapeutics of COVID-19: A comprehensive review. Journal of Medical Virology, 93(1), 275-299.

Tahir Ul Qamar, M., Maryam, A., Muneer, I., Xing, F., Ashfaq, U.A., Khan, F.A., Anwar, F., Geesi, M.H., Khalid, R.R., Rauf, S.A., & Siddiqi, A.R. (2019). Computational screening of medicinal plant phytochemicals to discover potent pan-serotype inhibitors against dengue virus. Scientific Reports, 9, 1433.

Tao J., Hu Q., Yang J., Li R., Li X., Lu C., Chen C., Wang L., Shattock R., & Ben K. (2007). *In vitro* anti-HIV and -HSV activity and safety of sodium rutin sulfate as a microbicide candidate. Antiviral Research, 75(3), 227-233.

Trott, O., & Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31, 455-461.

Wacker A., & Eilmes H.G. (1978). Antiviral activity of plant components. 1st communication: Flavonoids. Arzneimittelforschung, 28(3), 347-350.

World Health Organization. (2021a). WHO Coronavirus (COVID-19) Dashboard. Retrieved May 11, 2021, from https://covid19.who.int/

World Health Organization. (2021b). Coronavirus disease (COVID-19) pandemic. Retrieved May 11, 2021, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Zakaryan, H., Arabyan, E., Oo, A., & Zandi, K. (2017). Flavonoids: promising natural compounds against viral infections. Archives of Virology, 162(9), 2539-2551.

***