International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

Admet Analysis, Biological Activity and Target Protein Prediction Studies of the Selected Phytocompounds Using Online Tools

[1]Asha. R, [2]Sundara Prabha. V and [3]Beschi Antony Rayan
[1]PG student , [2]Assistant Professor, Department of Botany, ST Hindu College, Affiliated to MS. University, Tirunelveli, Tamil Nadu. [3]Associate Professor, Department of Botany and Research Centre, St. Xavier's College (Autonomous), Affiliated to MS. University, Tirunelveli, Tamil Nadu, India

Submitted: 05-06-2022 Revised: 18-06-2022 Accepted: 27-06-2022

ABSTRACT:
A huge number of deaths have been reported in the recent past caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which turned it into a serious and momentous threat to public health worldwide. The causative factor of covid infection, a novel coronavirus (SARS-CoV-2) made it a pandemic situation. The outbreak of covid-19 in December 2019 was heavy and no antiviral drug was available for instant relief in the initial stage. The computational study of phytochemical compounds for their pharmacological quality may provide more promise for the designing of new antiviral drugswith the minimum side effects. Phytochemical compounds present in medicinal plants have widely been reported for their preventive efforts in the managing health condition of people affected. In this study five compounds viz., Quercetin, Helichrysetin, Hesperatin, Sinigrin and Sitosterol were selected and their probability to be the candidate drugs were analyzed. Biological activities were predicted through PASS, drug likeliness and toxicity through the software ADMET lab2.0 and the target prediction using SuperPred tools were carried out. The finding of the study showed encouraging results. Further studies will pave the way for drugs development in the future.

KEYWORDS: SAR-Co-V-2, ADMET lab 2.0, PASS, SuperPred.

I. INTRODUCTION

Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), which made anegative impact on the global health and economy COVID-19 is a family of positive ribonucleic acid (RNA) viruses is an infectious disease (Helmy et al., 2020). Coronaviruses are a group of viruses with the potential to infect several vertebrates, including humans, leading to Severe Acute Respiratory Syndrome (SARS) and a variety of other respiratory infections (Velthuis et al., 2010). In silico drug discovery is a useful approach because very large numbers (up to millions) of drug candidate compounds can be screened, which is not possible using experimental approaches. There are two main methods used for in silico drug discovery: ligand-based drug discovery (LBDD) and structure-based drug discovery (SBDD), which have various advantages and their own pitfalls too.

These biologically-active phytochemicals have pharmacological properties that have been proven to be safe and effective in the treatment of chronic disorders (Opstelten et al., 1995). In view of the wide range of biological activities displayed by flavonoids, researchers soon turned their heads towards these natural products in the search for weapons against the new viral diseases.Prediction of Activity Spectra for Substances (PASS), the web based application predicts the biological activity spectrum of a compound based on its structure. PASS software was assimilated from the bio-activities of more than 270,000 compound-ligand pairs (Mervin et al., 2015). A successful drug should exhibit a combination of biochemical behavior, pharmacokinetics and safety. ADMETlab 2.0, an online tool currently supports two computational modes: single-molecule evaluation and batch screening, allowing for the calculation of 88 ADMET-related parameters, including 17 physicochemical properties, 13 medicinal chemistry properties, 23 ADME properties, 27 toxicity endpoints and 8 toxicophore rules (751 sub-structures). The SuperPred web server comparesthe chemical similarity of drug-like compounds with the target molecules and the therapeutic approach based on the principle of similarity in their properties. The SuperPred web server basically works on two methods, one for the

International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

classification of drugs based on approved drugs classified by WHO (WHO, 2003) and another for the prediction of their targets based on compound–target interaction data.

II. MATERIALS AND METHODS

In this study, it was computationally screened and determined the physical properties, bioactivity scores and ADMET scores for the selected phytocompounds (Quercetin, Helichrysetin, Hesperetin, Sinigrin and Sitosterol).

Quercetin, with a chemical name of 2- (3,4- dihydroxyphenyl) - 3,5,7- trihydroxychromen- 4- one or 3,3′,4′,5,7- pentahydroxyflavone, is classified as a flavonol, one of the six subcategories of flavonoid compounds, and is the major polyphenolic flavonoid found in various vegetables and fruits, such as berries, lovage, capers, cilantro, dill, apples, and onions (Anand et al., 2016).

Hesperetin is a trihydroxyflavanone having the three hydroxy gropus located at the 3'-, 5- and 7-positions and an additional methoxy substituent at the 4'-position. Hesperetin is the most important orange antioxidant flavone, which is found in the fruit in a glycosylated form as hesperidin. (Nogata et al., 2000). Quercetin 3-β-D-glucoside and helichrysetin were reported to block the enzymatic activity of MERS-CoV 3CLpro. Helichrysetin is a natural product found in Alpinia blepharocalyx and Alpinia hainanensis; Sitosterol is present in the roots of Isatis indigotica, belonging to the family Cruciferae. Antiviral effects of I. indigotica root were found against influenza, hepatitis A and Japanese encephalitis (Lin et al., 2005). Sinigrin is a glucosinolate which belongs to the family of glucosides found in some plants of the family Brassicaceae such as Brussels sprouts, broccoli, and the seeds of black mustard (Brassica nigra). The chemical name of sinigrin is allylglucosinolate or 2-propenylglucosinolate. The structures and the physical properties of the phytochemicals were collected from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). The physical properties of the compound are presented in Table 1.

Table 1: Properties of phytochemical compounds

S.no	Name of the compound	Pub chem. ID	Mol. wt	Mol. formula
1	Quercetin	5280343	302.23g/mol	$C_{15}H_{10}O_7$
2	Helichrysetin	6253344	286.28g/mol	$C_{16}H_{14}O_5$
3	Hesperetin	72281	302.28g/mol	$C_{16}H_{14}O_6$
4	Sinigrin	23682211	397.5g/mol	$C_{10}H_{16}KNO_9S_2$
5	Sitosterol	222284	414.7g/mol	$C_{29}H_{50}O$

Prediction of activity spectra for substances (PASS) is an online server (http://www.way2drug.com/) that was used to predict probable pharmacological effects of compounds based on their structural information. This tool is based on the comparative study of <300 pharmacological activities and mechanism of action of different compounds. It gives us the probability of activity (Pa) and inactivity (Pi) values of a particular compound under study (Lagunin et al., 2000)

Evaluation of compounds for their drug likeliness

SuperPred (Nickel et al., 2014) is an on line server used to predict the target for each of the selected compounds. SuperPred, the online tool cancalculate the Tanimoto similarity by calculating the Tanimoto similarity between molecules and more than 3,00,000 known compounds in the database and thus can predict the potential targets of unknown molecules. The distributions of the property values were saved in the database for each Anatomical Therapeutic Chemical (ATC) group and class.ADMETlab 2.0 gives a convenient and user friendly interface for users. Two services, namely Evaluation and Screening, are designed to support bothsingle-molecule and batch evaluation. ADMET was performed virtually to know the pharmacokinetic or drug-likeness property and mutagenicity of phytochemical components.

International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

The drug-likeness of the compounds were predicted by adopting the Lipinski's Rule of 5. The rule was developed to set drug-likeness ground rules for new molecular entities (NMEs) (Lipinski, 2000). The Rule of 5 predicts molecules with not more than 5 H-bond donors, 10 H-bond acceptors, whose molecular weight should not be more than 500 Da, and the calculated Log P (CLog P) not exceeding 5. The compounds which deviate from these set of rules likely would be with poor absorption or permeation of the molecular entities. Hence, molecules will unlikely become orally bioavailable as a drug if their properties fall outside these boundaries.

The selected compounds's structure data in the SMILE format was taken from the PubChem database. The prediction was executed using the standard protocol which helps us to obtain the biological activity of the molecule.

III. RESULTS AND DISCUSSION

In this study, five naturally occurring phytocompounds with different properties were chosen from the previous work done by Remali and Aizat (2021). The physical properties, bioactivity, ADMET analysis and protein target prediction were done by different bioinformatics tools. In previous works, scientists have suggested that the selected compounds Sitosterol, Hesperatin, Helichrysetin, Quercetin and Sinigrin were having the capacity to minimize the infection caused by viral infections especially coronavirus.

In this computational study, the ability of the compound's potential drug likeliness was investigated using PASS, ADMETlab2.0 and SuperPred. Four of the five compounds showed positive results.

If Pa > 0.7, the molecule is likely to exhibit the activity in the experiment, but the chance of the substance being the analogue of a known pharmaceutical agent is also high. If 0.5<Pa <0.7, the substance is likely to exhibit the activity in the experiment, but the probability is less and the substance is unlike known pharmaceutical agents. If Pa <0.5 the substance is unlikely to exhibit the activity in the experiment. However if the presence of this activity is confirmed in the experiment the substance might be a new chemical entity.

Sitosterol compound has been predicted with many biological activities and some of the important activities are tabulated. They have DELTA14- sterol reductase inhibition, hypolipmic, antiviral, CYP2C substrate, acylcarnitine hydrolase inhibition activity as the most important ones. The bioactivity predictions in Hesperetin were found to be CYP1A1 inhibition, Chlordecone reductase inhibition and Membrane integrity agonistic. The biological activities like Anti-hypercholesterolemic, prostaglandin E2 9-reductase inhibition, CYP2C substrate, UGT1A4 substrate were few of the results predicted from Helichrysetin.

The compound Quercetin showed bioactivities such as peroxidase inhibition, HMOX expression enhancement, Antimutagenicity and MAP kinase stimulation. The Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) profile of the selected compounds were predicted using ADMET lab2.0. The values for all the parameters of physical property, absorption, distribution, metabolism and excretion fell within the non-toxic range and confirmed suitability of these compounds to be the lead molecules. The results of the ADMET analysis are presented..

From the results, it was confirmed that all the selected compounds obey the Lipinski's rule of five with few exceptions. The following observations were found in this analysis. Number of hydrogen bond donor was found to be less than 10; the hydrogen acceptor as less than 5. Molecular weight was found to be less than 500 for all the molecules. Number of rotatable bonds were less than 10 and LogP parameter less than 5.

Molecular weight was high for sitosterol and low in helichrysetin. The nHBA was more for sinigrin and less for sitosterol. The nHBD results showed maximum value in quercetin and minimum for sitosterol. TPSA was maximum for sinigrin and minimum for sitosterol among the compounds analysed. The logP value was high for sitosterol and low for sinigrin.In absorption, MDCK Permability was more in sitosterol and less in helichrysetin. P-gp inhibition value was highest for sitosterol and lowest in helichrysetin. PPB of distribution was more for helichrysetin and less for sinigrin. AMES Toxicity was more for helichrysetin and found to be less for sitosterol.

The tool available at (http:www.organicchemistry.org/prog/peo/) described the logS value which indicates solubility; the lesser the logS value, the higher the solubility which would enhance the absorption. Higher c logP value indicates lower hydrophilicity and thus poor absorption and permeation.

The compound with lower TPSA value indicates the compound's ability for drug likeliness property and also predicted that a molecule with better CNS penetration should have lower TPSA value (Blake, 2000 and Chico et al., 2009). Poor

International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

pharmacokinetics and toxicity in the biological system lead to failure in drug development. With the help of ADMET profile during the process of drug discovery one can remove incompatible compounds as well as play a significant role in reducing cost and efforts (Miza et al., 2015)

Sitosterol was predicted for its target protein and many results were obtained. Nuclear factor NF-Kappa-B p105 subunit is a ubiquitous transcription factor which helps in inflammatory and immune response, which helps in the regulation of many genes related to survival. (Giuliani et al 2018). Cathepsin D was the next target protein whose role is the degradation of many disease associated protein like α synuclein, amyloid precursor protein (Bunk et al 2021). Similar to this few target protein were predicted and tabulated for sitosterol. Their ChEMBL-ID uniprot ID, PDB visualization, TTDID, probability and Model accuracy are given in the prediction table 4a.

Hesperetin as analyzed in SuperPred and the predicted target proteins were Tyrosyl-DNA Phosphodiesterase protein. This encodes the gene which is involved in the repairing of statted topisomerase 1 DNA complexes. Arachidonate 12-lipoxygense, Cathepsin D were the other target proteins for Hesperetin.

Transthyretin protein transporsts vitamin A and a hormone called thyronine throughout the body. Archidonic acid 1, 2 lipoxygenase, Kruppel like factor5 were the target protein for helichrysetin. Cytochrome p450 19, Aldose reductase, Cytochrome p450 1B1, Serine\threonine protein kinase P1M1 were the target protein predicted by SuperPred for quercetin.

Serine\threonine kinase belongs to the group of calcium\ calmodulin regulated kinase (CAMK). It is widespread and ranges from hematopoietic and lymphoid system to prostate, testis or oral epithelial cell (Bachmann et al., 2005). Endoplasmic reticulum associated amyloid beta peptide binding protein, 15-hydroxyprostaglandin dehydrogenase, Adinosine A1 receptor, Bloom syndrome protein were some of the target protein predicted for sinigrin. BLM is a 3'→5' ATP dependent Rec Q DNA helicase that is one of the most essential genome stabilizers involved in the regulation of DNA replication, recombination and both homologous and non-homologous pathways of double strand break pair (Kaur et al., 2021).The analysis of the computational investigation results revealed encouraging results for all the selected phytocompounds. In future, an accurate toxicity analysis along with the combination of docking and lead optimization could yield better understanding of the compounds ability to get developed unit a potential drug. It helps in the discovery of many drugs from the medicinal plants with learnt side effect, time consuming and cost effective.

An absorption value below 30% indicates poor absorbance (Egan et al., 2000). All compounds selected displayed a value greater than 60% which shows good absorbency in the human intestine.

REFERENCE

[1]. Helmy, YA, Fawzy, M, Elaswad, A, Sobieh, A, Scott, P & Awad, SA 2020, ' The covid-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control', J Clin Med, vol.9, no.4, pp.1225.

[2]. Velthuis, AJWT, Arnold, JJ, Cameron, CE, Worm, S & Snijder, EJ 2010, 'The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent', Nucleic Acids Res, vol. 38, no.1, pp. 203–214.

[3]. Opstelten, DGE, Raamsman, MJB, Wolfs, K, Horzinek, MC & Rottier, PJM 1995, 'Envelope glycoprotein interactions in coronavirus assembly', J. Cell Biol, vol.131, pp. 339-349

[4]. Mervin, LH, Afzal, AM, Drakakis, G, Lewis, R, Engkvist, O & Bender, A 2015, 'Target prediction utilising negative bioactivity data covering large chemical space', J. Cheminform, vol.7, pp.1–16.

[5]. The selection and use of essential medicines. Report of the WHO expert committee, 2005 (including the 14th model list of essential medicines). (2003) World Health Organ. Tech. Rep. Ser., 1–119

[6]. Anand, DAV, Arulmoli, R & Parasuraman, S 2016, 'Overviews of biological importance of quercetin: A bioactive flavonoid', Pharmacognosy Reviews, vol.10, pp.84–89.

[7]. Nogata, Y, Sakamoto, K, Shiratsuchi, H, Ishii, T, Yano, M & Ohta H 2006, 'Flavonoid Composition of Fruit Tissues of Citrus Species', Biosci. Biotechnol. Biochem, vol.70, pp. 178–192.

[8]. Lin, CW, Tsai, FT, Tsai, CH , Lai, CC , Wan,L , Ho, TY, Hsieh, CC & Chao, DW 2005, 'Anti-SARS coronavirus 3C-like protease effects of Isatis indigotica root and plant-derived phenolic compounds', Antiviral Research, vol. 68, pp.36–42

[9]. Lagunin, A, Stepanchikova, A, Filimonov, D & Poroikov, V 2000, 'PASS: prediction

International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

activity spectra for biologically active substances', Bioinformatics, vol.16, no.8, pp. 747–48.

[10]. Lipinski, CA 2000, 'Drug-like properties and the causes of poor solubility and poor permeability', J Pharmacol Toxicol Methods, vol. 44, no. 1, pp. 235–249.

[11]. Nickel, J, Gohlke, BO, Erehman, J, Banerjee, P, Rong, WW, Goede, A, Dunkel, M & Preissner, R 2014, 'SuperPred: update on drug classification and target prediction', Nucleic Acids Res, vol. 42(Web Server issue), pp. W26-31

[12]. Remali, J & Aizat, WM 2021, 'A Review on Plant Bioactive Compounds and Their Modes of Action Against Coronavirus Infection', Front. Pharmacol, vol.11, no. 589044

[13]. Blake, JF 2000, 'Chemoinformatics-predicting the physiochemical properties of drug-like molecules', Current opinion in biotechnology, vol. 11, no. 1, pp. 104-107.

[14]. Chico, LK, Van Eldik & Watterson, DM 2009, 'Targetting protein kinases in central--- system disorders', Nature Reviews Discovery, vol. 8, no. 11, pp. 892-909.

[15]. Miza, MU, Ghori, NU, Narm, N, Adi, AR & S, Manzoor 2015, 'Pharmacoinformatics approach for investigation of alternative potential hepatitis C virus nonstructural protein 5B inhibitors', Drug Des DevTher, vol.9, pp. 1825-1941.

[16]. Giuliani, C, Bucci, I & Napolitano, G 2018, 'The role of the transcription factor nuclear factor-kappa B in thyroid autoimmunity and cancer's, Front.Endocrinol, vol.21, no.9, pp. 1-8.

[17]. Bunk, J, Huarcaya, SP, Drobny, A, Dobert, JP, Walther, L, John, SR, Arnold, P & Zunke, F 2021, 'Cathepsin D Variants Associated With Neurodegenerative Diseases Show Dysregulated Functionality and Modified α-Synuclein Degradation Properties', Front Cell Dev Biol., vol. 23, No. 9, PP. 1-16

[18]. Bachmann, C & Moroy, T 2005, 'The serine/threonine kinase Pim-1', Int J Biochem Cell Bill, vol. 37, no. 4, pp.726-30.

[19]. Kaur, E, Agarwal, R & Sengupta, S 2021, 'Functions of BLM Helicase in Cells: Is It Acting Like a Double-Edged Sword?', Front Genet., vol.12, PP. 1-11.

[20]. Egan, WJ, Merz, KM & Baldwin, JJ 2000, 'Prediction of drug absorption using multivariate statistics', J. Med. Chem, vol. 43, pp. 3867–3877

Table 4: List of protein target for the compounds using SuperPred

Table 4a: Sitosterol

S.No	Target Name	ChEMBL-ID	UniProt ID	PDB Visualization	TTD ID	Probability	Model accuracy
1	Nuclear factor NF-kappa-B p105 subunit	CHEMBL3251	P19838	1SVC	Not Available	97.23%	96.09%
2	Cathepsin D	CHEMBL2581	P07339	4OD9	T67102	96.37%	98.95%
3	Adenosine A1 receptor	CHEMBL226	P30542	5N2S	T92072	94.33%	95.93%
4	Cannabinoid CB2 receptor	CHEMBL253	P34972	6KPF	Not Available	93.89%	97.25%
5	Dual specifcity protein kinase CLK4	CHEMBL4203	Q9HAZ1	6FYV	Not Available	92.7%	94.45%

Table 4b: Hesperatin

S. No	Target Name	ChEMBL-ID	UniProt ID	PDB Visualization	TTD ID	Probability	Model accuracy
1	Tyrosyl-	CHEMBL10	Q9NUW	6N0D	Not	99.83%	71.22%

International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

	DNA phosphodiesterase 1	75138	8		Available		
2	Nuclear factor NF-kappa-B p105 subunit	CHEMBL3251	P19838	1SVC	Not Available	96.94%	96.09%
3	DNA-(apurinic or apyrimidinic site) lyase	CHEMBL5619	P27695	6BOW	T13348	96.84%	91.11%
4	Arachidonate 12-lipoxygenase	CHEMBL3687	P18054	3D3L	Not Available	96.76%	75.57%
5	Transcription intermediary factor 1-alpha	CHEMBL3108638	O15164	4YBM	Not Available	93.43%	95.56%

Table 4c: Helicrysetin

S.No	Target Name	ChEMBL-ID	UniProt ID	PDB Visualization	TTD ID	Probability	Model accuracy
1	Transthyretin	CHEMBL3194	P02766	6SUG	T86462	96.66%	90.71%
2	DNA-(apurinic or apyrimidinic site) lyase	CHEMBL5619	P27695	6BOW	T13348	96.08%	91.11%
3	Arachidonate 12-lipoxygenase	CHEMBL3687	P18054	3D3L	Not Available	94.58%	75.57%
4	Tyrosyl-DNA phosphodiesterase 1	CHEMBL1075138	Q9NUW8	6N0D	Not Available	93.91%	71.22%
5	Nuclear factor NF-kappa-B p105 subunit	CHEMBL3251	P19838	1SVC	Not Available	93.77%	96.09%

Table 4d: Quercetin

S.No	Target Name	ChEMBL-ID	UniProt ID	PDB Visualization	TTD ID	Probability	Model accuracy
1	Cytochrome P450 19A1	CHEMBL1978	P11511	3S79	T13260	12 nm	IC50

International Journal of Pharmaceutical Research and Applications
Volume 7, Issue 3 May-June 2022, pp: 2049-2055 www.ijprajournal.com ISSN: 2456-4494

2	Aldose reductase	CHEMBL1900	P15121	1US0	T26623	14.8 nm	IC50
3	Cytochrome P450 1B1	CHEMBL4878	Q16678	3PM0	Not Available	23 nm	Ki
4	Serine/threonine-protein kinase PIM1	CHEMBL2147	P11309	3A99	T50594	25 nm	Kd
5	ATP-binding cassette sub-family G member 2	CHEMBL5393	Q9UNQ0	6HBU	T56556	30 nm	EC50

Table 4e: Sinigrin

S.No	Target Name	ChEMBL-ID	UniProt ID	PDB Visualization	TTD ID	Probability	Model accuracy
1	Endoplasmic reticulum-associated amyloid beta-peptide-binding protein	CHEMBL4159	Q99714	2O23	Not Available	99.34%	70.16%
2	15-hydroxyprostaglandin dehydrogenase [NAD+]	CHEMBL1293255	P15428	2GDZ	Not Available	95.33%	83.57%
3	Adenosine A1 receptor	CHEMBL226	P30542	5N2S	T92072	94.49%	95.93%
4	Glutamate receptor ionotropic, AMPA 2	CHEMBL4016	P42262	2WJW	T42392	90.9%	86.92%
5	Transcription intermediary factor 1-alpha	CHEMBL3108638	O15164	4YBM	Not Available	88.78%	95.56%