Current Chemistry Letters 11 (2022) 309–320

Contents lists available at GrowingScience

Current Chemistry Letters

homepage: www.GrowingScience.com

Computational investigation of Betalain derivatives as natural inhibitor against food borne bacteria

Fahima Siddikey[a], Md Abul Hasan Roni[b], Ajoy Kumer[c], Unesco Chakma[d] and Mohammed Mahbubul Matin[e*]

[a]*Department of Food Engineering & Tea Technology, Shahjalal University of Science & Technology, Sylhet, Bangladesh*
[b]*Department of Science and Humanities, Bangladesh Army International University of Science & Technology, Bangladesh*
[c]*Department of Chemistry, European University of Bangladesh, Gabtoli, Dhaka, 1216, Bangladesh*
[d]*Department of Electrical and Electronics Engineering, European University of Bangladesh, Gabtoli, Dhaka, 1216, Bangladesh*
[e]*Bioorganic and Medicinal Chemistry Laboratory, Department of Chemistry, Faculty of Science, University of Chittagong, Chittagong, 4331, Bangladesh*

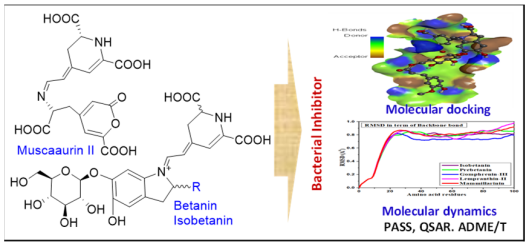
C H R O N I C L E

Article history:
Received August 25, 2021
Received in revised form
October 26, 2021
Accepted March 11, 2022
Available online
March 11, 2022

Keywords:
ADMET
Betalains
Drug-likeness
Molecular Docking
QSAR
Toxicity
pIC50

A B S T R A C T

Natural organic pigments such as carotenoids, betalains, anthocyanins, and carminic acid are notably found as safer food preservatives compared to other harmful synthetic chemicals. Due to glycosylation and acylation, betalains exhibit a broad-spectrum antimicrobial functionality with protection against degenerative diseases. Thus, betalains have been investigated as a potential bacterial inhibitor for food preservative applications. Initially, 36 betalain derivatives have been taken for primary screening using molecular docking. Afterward, the top ten ligands are taken for further study and analysis. The results of Prediction of Activity Spectrum of Substances (PASS) assured the antibacterial capabilities of betalains, and Lipinski's rule-of-five ensures the acceptability of the selected ligands as antibacterial inhibitors. The bacterial pathogens, such as *C. botulinum* (3FIE), *E. coli* (2ZWK), and *S. typhi* (3UU2) are selected for molecular docking by these betalain pigments. Furthermore, ADMET investigations and QSAR studies are performed to check insights into the bacterial inhibition process. Most active and common binding sides were observed at GLY159, ASN165, and SER166 for *C. botulinum*, at ASP8, LYS40, and TRP50 for *E. coli*; and at ARG37, GLN5, and ARG74 for *S. typhi*. The present study clearly shows an excellent insight towards the invention of plant-based new organic inhibitors to face the challenges of bacterial-resistant common food preservatives.



Graphical Abstract

* Corresponding author. Tel.: +88 01716 839689; Fax: +88 031 2606014
E-mail address: mahbubchem@cu.ac.bd mmmatin2004@yahoo.co.in (M. Matin)

doi: 10.5267/j.ccl.2022.3.003

1. Introduction

Consumption of leafy green vegetables or ready-to-eat (RTE) vegetables in raw, unprocessed, or cooked forms is an ancient and worldwide practice phenomenon for human foods. Some of them, such as spinach, lettuce, tomato, basil, and many others, are used raw for salads and garnished in food for daily life without processing or even cooking. Thus, green and fresh plant bodies or vegetables are crucial requirements to get the apex level of nutrients, vitamins, and fibers.[1,2] In the last two decades, fresh leafy (FL) and RTE foods have been getting vast popularity despite the health concerns. Due to high demand RTE market has rapidly and meticulously proliferated.[3,4] However, many contaminations of these vegetables, which are consumed raw, have illustrated the cardinal problematic concern in recent years. There are few studies on the FL and RTE vegetable products and their bacterial attack. A study conducted in Italy has assured the presence of pathogenic microorganisms in TRE.[5] During the pre-harvest and harvesting processes, pathogenic bacteria, such as *E. coli*, *Salmonella*, *Clostridium*, and *Listeria*, can attack edible parts of plants, and they can even persist after minimal processing of food.[6] Several studies have demonstrated that plant roots, body, and soil contacting parts can be infected by living pathogens in soil even edible parts of plants.[7] Widespread consumption of FL vegetables, for instance, spinach and lettuce were associated with *E. coli* outbreaks in Canada in 2000, South Wales in 2005, North America in 2006, the United Kingdom in 2009, and Germany in 2011. In 2012 and 2018, numerous people died due to food-borne illnesses in the United States of America.[8,9]

Many natural pigments and flavonoids are capable of inhibiting FL vegetable pathogens, which are highly responsive to decompose the vegetable and plants. Of them, betalains, natural pigments primarily derived from plants, are commonly found in red beetroot, amaranths, prickly pear, red pitahaya. Betalains have a color range from red to violet to yellow to orange. These are also found on *Amanita muscaria*, a fungus of a higher class.[10] Betalains have pharmacological properties in addition to their use as a food coloring. Numerous studies have been conducted on their antioxidant, anticancer, antimalarial, antilipidemic, and antimicrobial properties.[11-13] Moreover; betanin (the most studied betalain) exerts an anti-inflammatory effect aiding plant resistance to pathogenic activity. Their original nutritional betacyanin has significant growth inhibition in tumor cells of the stomach, breast, lung, colon, and central nervous system.[14] Presence of betalains in fresh and also processed fruit and vegetables also contribute to similar functions consumers' health.[14]

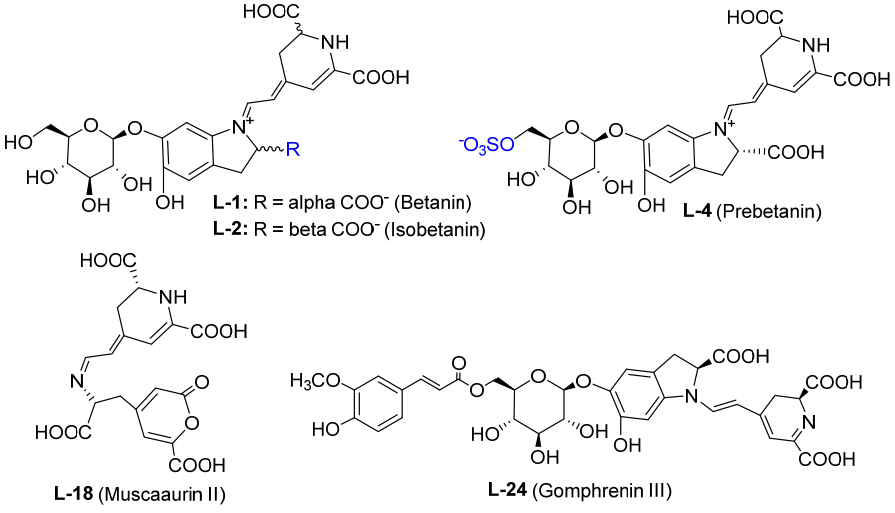


Fig. 1. Structures of some potential antibacterial betalains.

So far, 75 betalain derivatives have been discovered from 17 plants. Among them, 36 betalains (**L-1** to **L-36**) were chosen for this study based on their availability in standard data banks. From the Prediction of Activity Spectra for Substances (PASS) (e.g. antibacterial, antifungal, antiviral, antioxidant, and anti-carcinogenic properties) these betalains showed substantial antibacterial properties. Hence, molecular docking against important food and water-borne three pathogenic bacteria (*E. coli*, *S. typhi*, and *C. botulinum*) is conducted.[15] The highly active ten compounds (**L-1**, **L-2**, **L-4**, **L-18**, **L-24**, **L-27**, **L-28**, **L-30**, **L-31**, and **L-34**; some of them are shown in **Fig. 1**) are subjected for further study for chemical reactivity descriptors, frontier molecular orbitals, drug-likeness, ADME/T, QSAR, and molecular dynamics and discussed here.[16,17]

2. Results and Discussion

2.1. Optimized structures of betalains

The optimized molecular structure represents the most important fact for determining the molecular reactivity, chemical reactivity, and biological activity. Through the use of computational tools, the DFT method was used to simulate the studied betalain compounds, and their optimized chemical structures are listed below in **Fig. 2**.[18]

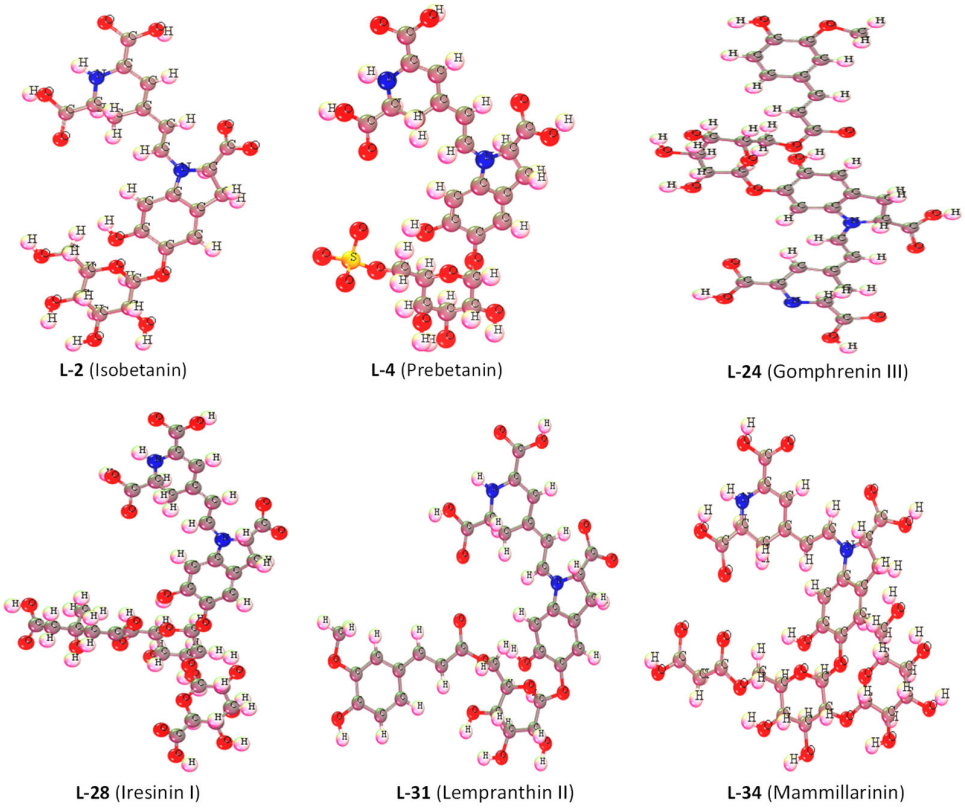


Fig. 2. Optimized structures of ligands (betalains)

2.2. Prediction of Activity Spectrum for Substances (PASS)

In the PASS program, if a compound is to be considered a potential medication its Pa value must be greater than its Pi value. If Pa>0.7, the probability of identifying the activity empirically is very high. When Pa values are between 0.5 and 0.7 (0.5<Pa<0.7), biological activity is significantly reduced.[19] For this study, antibacterial, antiviral, antifungal, anti-carcinogenic, and antioxidant properties of betalains are analyzed and are mentioned in **Table 1**.

Table 1. PASS calculated data of betalains.

S/N	Com-pound No.	Predicted activities									
		Antibacterial		Antifungal		Anti-carcinogenic		Antiviral		Antioxidant	
		Pa	Pi	Pa	Pi	Pa	Pi	Pa	Pi	Pa	Pi
01	**L-1**	0.486	0.018	0.219	0.125	0.265	0.074	0.305	0.086	0.326	0.019
02	**L-2**	0.486	0.018	0.219	0.125	0.265	0.074	0.305	0.086	0.326	0.019
03	**L-4**	0.479	0.018	0.363	0.059	0.210	0.110	0.407	0.084	0.287	0.025
04	**L-18**	0.313	0.055	NA	NA	NA	NA	0.409	0.082	NA	NA
05	**L-24**	0.325	0.051	0.307	0.078	0.441	0.025	0.212	0.076	0.475	0.008
06	**L-27**	0.416	0.026	0.359	0.060	0.392	0.032	0.300	0.089	0.489	0.007
07	**L-28**	0.492	0.017	0.439	0.042	0.387	0.033	0.331	0.192	0.365	0.015
08	**L-30**	0.555	0.012	0.353	0.062	0.318	0.052	0.589	0.007	0.323	0.020
09	**L-31**	0.457	0.021	0.275	0.093	0.395	0.031	0.203	0.085	0.392	0.013
10	**L-34**	0.491	0.017	0.208	0.132	0.451	0.024	0.286	0.282	0.408	0.011

NA = not available

Pa values of Betalain compounds showed 0.159<Pa<0.555 for antibacterial, 0.219<Pa<0.462 for antifungal, 0.210<Pa<0.441 for anti-carcinogenic, 0.203<Pa<0.589 for antiviral and 0.130<Pa0.475 for antioxidant properties. Overall predicted results indicated that the selected compounds might have better antibacterial properties compared to their antifungal and anti-carcinogenic and antioxidant properties (**Table 1**). As betalains possessed significantly greater antibacterial properties than the remaining properties, this study further aims to demonstrate whether betalains are viable drug candidates against bacteria.

2.3. Molecular docking and binding affinity analysis

Molecular docking is commonly utilized in current drug design to investigate drug-receptor interactions, and the docking result is expressed in terms of the drug's affinity for the pathogen's active site, as well as the number of hydrogen bonds, hydrophobic bonds, polar and non-polar bonds, and the polar and non-polar links that connect them.[20] Based on the docking score top 10 compounds out of 36 (**Table 2** and S1) are represented here.

For the *E. coli* pathogen, compounds L-10, L-12, L-13, L-15, L-19, L-21, and L-22 had binding affinity less than -6.0 kcal/mol. All other compounds had binding affinity greater than -6.0 kcal/mol, with lempranthin II (**L-31**) having the greatest binding affinity of -7.9 kcal/mol. Betalain compounds had a stronger binding affinity against *S. typhi* since the binding affinity of all 36 compounds was larger than -6.0 kcal/mol. Lempranthin II (**L-31**) has the highest binding affinity of -10.0 kcal/mol out of all of them.

In the case of *C. botulinum*, it was established that betalain compounds are perhaps the most active, with 32 compounds out of 36 having a binding affinity greater than -8.0. **L-29** had the highest binding affinity in this case (-10.5 kcal/mol).

From the study, it can be summarized that compound **L-31** has the highest binding energy with *E. coli*, *S. typhi*, and *C. botulinum* proteins. Supplementary Tables (S8, S9, and S10) displayed the amino acid residues of *S. typhi*, *E. coli*, and *C. botulinum* for hydrogen bond and hydrophobic bond interaction with bond distance, respectively.

Table 2. Molecular docking score (kcal/mol) against three bacterial pathogens.

S/N	Ligand No.	*Escherichia coli*			*Salmonella typhe*			*Clostridium botulinum*		
		Binding affinity	No. of H bond	No. of hydrophobic bond	Binding affinity	No. of H bond	No. of hydrophobic bond	Binding affinity	No. of H bond	No. of hydrophobic bond
01	**L-1**	-7.1	6	1	-8.6	10	No	-9.5	10	No
02	**L-2**	-7.1	4	1	-9.0	7	No	-9.8	8	No
03	**L-4**	-7.8	7	1	-9.2	11	No	-9.9	10	No
04	**L-18**	-6.9	3	No	-7.9	5	No	-9.4	9	1
05	**L-24**	-7.4	6	2	-9.5	10	1	-10.1	10	1
06	**L-27**	-7.0	11	No	-9.3	11	No	-8.9	14	1
07	**L-28**	-6.6	10	No	-9.1	10	1	-9.8	13	1
08	**L-30**	-6.9	5	No	-8.7	11	1	-9.1	7	No
09	**L-31**	-7.9	7	3	-10	11	1	-9.6	10	1
10	**L-34**	-7.0	6	No	-8.8	11	No	-9.3	13	No

*Based on combined binding affinity only ten betalains are selected

2.4. Chemical reactivity and descriptors

HOMO denotes the highest occupied molecular orbital, while LUMO denotes the lowest unoccupied molecular orbital, both of which are considered substantial orbitals of frontier molecular orbitals (FMOs).[21] FMOs energy-related data are presented in **Table 3**.

Table 3. Frontier molecular orbitals and Reactivity descriptor analysis.

Comp-ound	LUMO, eV	HOMO, eV	HOMO LUMO gap, eV	Ionization potential (I), eV	Electron affinity (A), eV	Chemical potential (μ), eV	Hardness (η), eV	Electrons activity (χ), eV	Electrophilicity (ω), eV	Softness (S), eV
L-1	-3.459	-7.993	4.534	7.993	3.459	-5.726	2.267	5.726	7.231	0.441
L-2	-3.459	-7.993	4.534	7.993	3.459	-5.726	2.267	5.726	7.231	0.441
L-4	-4.488	-7.966	3.478	7.966	4.488	-6.227	1.739	6.227	11.14	0.575
L-18	-2.064	-8.364	6.300	8.364	2.064	-5.214	3.150	5.214	4.315	0.317
L-24	-1.933	-7.560	5.627	7.56	1.933	-4.746	2.813	4.746	4.003	0.355
L-27	-1.953	-8.034	6.081	8.034	1.953	-4.993	3.040	4.993	4.100	0.328
L-28	-3.503	-8.034	4.531	8.034	3.503	-5.7685	2.265	5.768	7.343	0.441
L-31	-2.812	-8.070	5.258	8.07	2.812	-5.441	2.629	5.441	5.630	0.380
L-34	-1.094	-8.330	7.236	8.33	1.094	-4.712	3.618	4.712	3.068	0.276

The lower the energy gap between the two levels, the higher the chemical reactivity and the lower the chemical stability. The energy is listed in Table 3 conveys the chemical reactivity, thermal stability, part of the electrophilic or nucleophilic region, and biological dissociation or attachment with protein.[21] It is revealed that the HOMO-LUMO energy gaps are about -6.00 to -9.00 eV for best fitting organic compounds (**Table 3**).[22,23] In addition, the softness for these ligands is obtained

from 0.574 to 0.317 eV while sample **L-4** shows the highest softness value.

2.5. Frontier molecular orbital of HOMO and LUMO

Frontier molecular orbital of HOMO and LUMO has been depicted below (**Fig. 3**) by different colors for well understanding. Also, a specific color map makes all other molecules available and accessible. In HOMO, the deep radish color represents positive nodes and the yellow color represents negative nodes. That is why the HOMO is the part of biologically active molecules where the electrophilic attractive group can be attached to form the chemical bonding, and the LUMO is the positive part where the nucleophilic group can be added.[24-26] The negative part of the orbital is lime green, while the positive part is olive green.

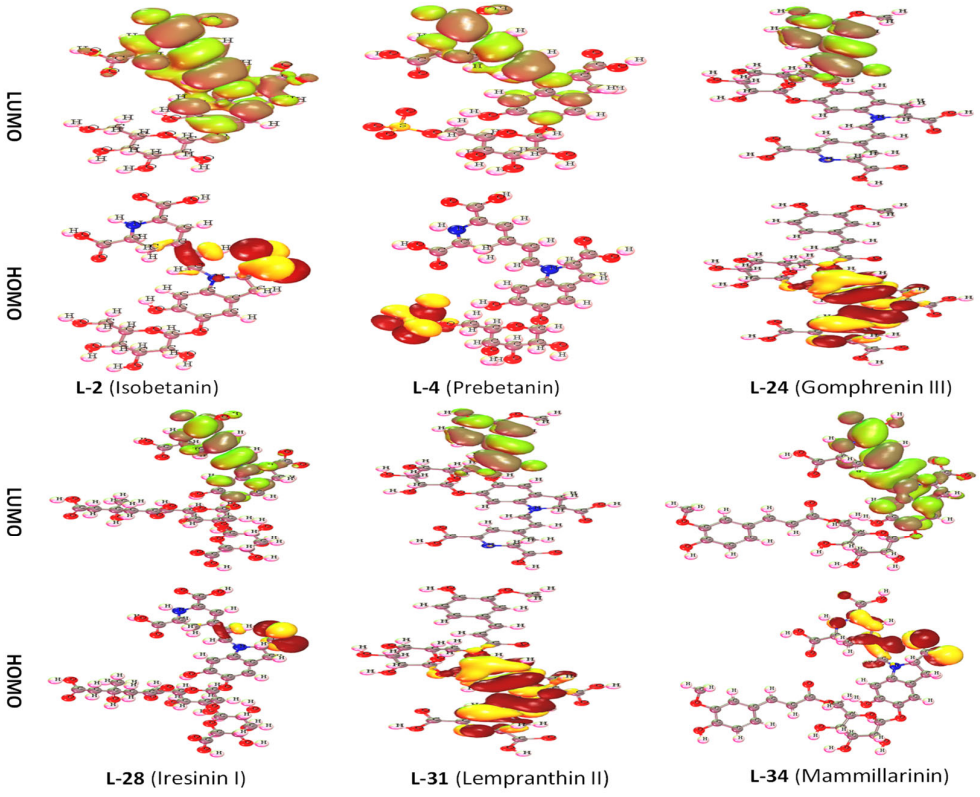


Fig. 3. Frontier molecular orbitals diagram for HOMO and LUMO

2.6. Drug-like properties of betalains

The drug-like properties of potential compounds are predicted using Lipinski's rule of drug-likeness or the rule-of-five. According to this rule, a compound exhibits optimal drug-like behavior if it possesses at least four of the five characteristics including, a molecular weight of fewer than 500 Daltons, five or fewer hydrogen bond donors, ten or fewer hydrogen bond acceptors, the lipophilicity of less than five, and molar refractivity between 40 and 130.[27,28] All these data are presented in **Table 4** and shows that ligands **L-5** to **L-21** and **L-35** follow Lipinski rule-of-five. Ligand **L-5**, **L-7**, **L-8**, **L-14**, **L-15**, and **L-21** show a strong bioavailability score as we know 0.55.[29]

Table 4. Data of Lipinski rule, pharmacokinetics, and drug-likeness.

Ligand	NBR	HBA	HBD	TPSA A^2	Consensus Log Po/w	Log KP (Skin permeation) cm/s	Lipinski rule		MW	Bioavailability score	GI absorption
							Result	Violation			
L-1	7	13	8	249.38	-3.14	-10.73	No	3	550.47	0.11	Low
L-2	7	13	8	249.38	-3.14	-10.73	No	3	550.47	0.11	Low
L-4	9	16	8	301.13	-3.58	-11.89	No	3	503.30	0.11	Low
L-18	8	11	5	203.80	-0.74	-9.08	Yes	1	420.33	0.11	Low
L-24	13	17	8	282.64	-0.20	-10.40	No	3	726.64	0.11	Low
L-27	11	20	11	343.33	-3.53	-12.58	No	3	726.59	0.11	Low
L-28	16	23	12	409.20	-4.92	-14.34	No	3	870.72	0.11	Low
L-30	9	16	8	301.13	-3.47	-11.49	No	3	630.53	0.11	Low
L-31	12	16	8	284.91	-1.78	-10.58	No	3	726.64	0.11	Low
L-34	14	21	12	369.07	-4.85	-13.58	No	3	799.66	0.11	Low

2.7. ADME properties of betalains

Absorption, distribution, metabolism, and excretion (ADME) are all terms used to describe how drugs move into, though, and out of the body- known as the pharmacokinetics of a drug. For evaluating pharmacokinetic properties of betalain compounds total of eight properties, including human intestinal absorption, Caco-2 permeability, blood-brain barrier, P-glycoprotein substrate and inhibitor, renal organic cation transporter, sub-cellular localization, CYP450 2C9 substrate, and CYP450 1A2 inhibitor were taken in our study. admetSAR was used for predicting these eight properties.[17,30]

All 36 betalain compounds showed negative blood-brain barrier & human intestinal absorption. So, they cannot pass through the blood-brain barrier and don't have good gastrointestinal absorption. **L-25** showed the highest intestinal absorption, whereas **L-32** showed low intestinal absorption.

P-glycoprotein inhibition can cause chemical species absorption, permeability, and retention to be disrupted. According to our findings, none of these compounds were found as inhibitors. Except for **L-5** and **L-6**, there were no CYP450 1A2 inhibitors among the 36 compounds. Aside from that, none of the 36 compounds tested were CYP450 2C9 substrates. These eight mentioned properties of ten compounds, which showed the highest binding affinity, are listed in **Table 5**.

Table 5. Pharmacokinetic parameters of the best betalain compounds.

S/N	Human Intestinal Absorption	Caco-2 Permeability	Blood Brain Barrier	P- glycoprotein inhibitor	P- glycoprotein substrate	Renal Organic Cation Transporter	Sub-cellular localization	CYP450 2C9 Substrate	CYP450 1A2 Inhibitor
L-1	0.9530	0.6622	no	no	yes	0.7801	Nucleus	no	no
L-2	0.9530	0.6622	no	no	yes	0.7801	Nucleus	no	no
L-4	0.9397	0.6103	no	no	yes	0.8303	Lysosome	no	no
L-18	0.6978	0.7203	no	no	yes	0.8133	Mitochondria	no	no
L-24	0.8331	0.6075	no	no	yes	0.6709	Mitochondria	no	no
L-27	0.7470	0.6589	no	no	yes	0.6857	Mitochondria	no	no
L-28	0.9333	0.6317	no	no	yes	0.8396	Mitochondria	no	no
L-30	0.9632	0.6166	no	no	yes	0.8469	Lysosome	no	no
L-31	0.9722	0.6255	no	no	yes	0.7705	Nucleus	no	no
L-34	0.9671	0.6691	no	no	yes	0.7621	Nucleus	no	no

2.8. Aquatic and non-aquatic toxicity of betalains

Green chemistry is a chemical science approach that focuses on minimizing or eliminating the usage of hazardous compounds. When developing a drug, it is vital to ensure that the medicine will not be detrimental to the environment and free of aquatic and non-aquatic toxicity. Hence, an environmental impact study of aquatic and non-aquatic attributes was theoretically conducted using admetSAR, a web-based tool.[17,31] Eight ADMET parameters are investigated for evaluating aquatic and non-aquatic toxicity, including AMES toxicity, carcinogenicity, water-solubility, plasma protein binding, acute oral toxicity, acute oral toxicity in rats, fish toxicity, and *T. pyriformis* toxicity (**Table 6**).

Table 6. Aquatic and non-aquatic toxicity of betalain compounds.

T/S	AMES toxicity	Carcinogenicity	Water solubility, Log S	Plasma protein binding	Acute Oral Toxicity, kg/mol	Oral Rat Acute Toxicity (LD50) (mol/kg)	Fish Toxicity pLC50 mg/L	*T. pyriformis* toxicity (log µg/L)
L-1	no	no	-2.5688	0.754	2.401	2.4833	0.8206	0.5557
L-2	no	no	-2.5688	0.754	2.400	2.4833	0.8206	0.5557
L-4	no	no	-3.2566	0.858	2.717	2.5384	1.1675	0.5224
L-18	no	no	-2.8326	0.890	1.960	2.4120	1.1993	0.4165
L-24	no	no	-2.3983	1.019	2.227	2.5418	0.5650	0.6453
L-27	no	no	-2.4206	0.749	2.620	2.4698	0.6531	0.5696
L-28	no	no	-2.9600	0.446	2.766	2.6770	0.7932	0.6514
L-30	no	no	-3.1954	0.749	2.661	2.5275	1.2160	0.4867
L-31	no	no	-2.5555	1.038	2.254	2.5326	0.7415	0.6259
L-34	no	no	-2.5366	0.668	2.736	2.4900	0.9649	0.5099

The Ames toxicity test is used to determine whether a chemical substance is capable of causing mutagenesis. According to our findings, ligands **L-5** and **L-6** have the potential to cause AMES toxicity. Similarly, carcinogenic toxicity is studied, and it is determined that no betalains are carcinogenic.

According to a prior study,[32] substances displayed low acute toxicity in rats when their molecular weight was between 2.1398 and 2.5469 mol/kg. According to this study, betalains likewise have low acute toxicity in rats, with a median fatal dose (LD_{50}) ranging from 2.1117 to 2.5418 mol/kg. By comparing our findings to those of various other research,[32] it can be concluded that our study established values for water solubility, plasma protein binding, acute oral toxicity, fish toxicity, and *T. pyriformis* toxicity, which are safer to use. Aquatic and non-aquatic toxic properties of ten Ligands, which showed the highest binding affinity, are listed in **Table 6**.

2.9. QSAR and pIC_{50} analysis of betalains

The following multiple linear regression (MLR) equation is used for calculating pIC_{50} values.

pIC_{50} (Activity) = −2.768483965 + 0.133928895 × (Chiv5) + 1.59986423 × (bcutm1) + (− 0.02309681) × (MRVSA9) + (− 0.002946101) × (MRVSA6) + (0.00671218) × (PEOEVSA5) + (− 0.15963415) × (GATSv4) + (0.207949857) × (J) + (0.082568569) × (Diametert)

MLR equation is used to investigate the association between a single dependent variable (pIC_{50}) and a set of independent variables (descriptors Chiv5, bcutm1, MRVSA9, MRVSA6, PEOEVSA5, GATSv4, J, and Diametert).[28] **Table 7** shows the QSAR data and estimated pIC_{50} values. It is seen from the data that the values of pIC_{50} range between 3.98 and 5.66, which are considered standard levels. It means that all 36 compounds have the potential to be used against bacteria.

Table 7. QSAR and pIC_{50} of betalain compounds.

S/L	L-1	L-2	L-4	L-18	L-24	L-27	L-28	L-30	L-31	L-34
pIC_{50}	4.798	4.798	5.66	4.24	4.72	5.11	5.02	5.05	5.28	5.18

2.10. Different interactions of betalains with bacterial proteins

The hydrogen bonding and hydrophobic bonding against *C. botulinum* (3FIE), *E. coli* (2ZWK), and *S. typhi* (3UU2) proteins are calculated with the bond distance (supplementary Tables S8, S9, and S10). The hydrogen bond distance is lower than the hydrophobic bond distance. **Fig. 4** represents the various docking poses for the **L-34** ligand with *C. botulinum*, *E. coli*, and *S. typhi* bacterial proteases, and hydrogen bonding modes (similar results for all other ligands attached in supplementary files).

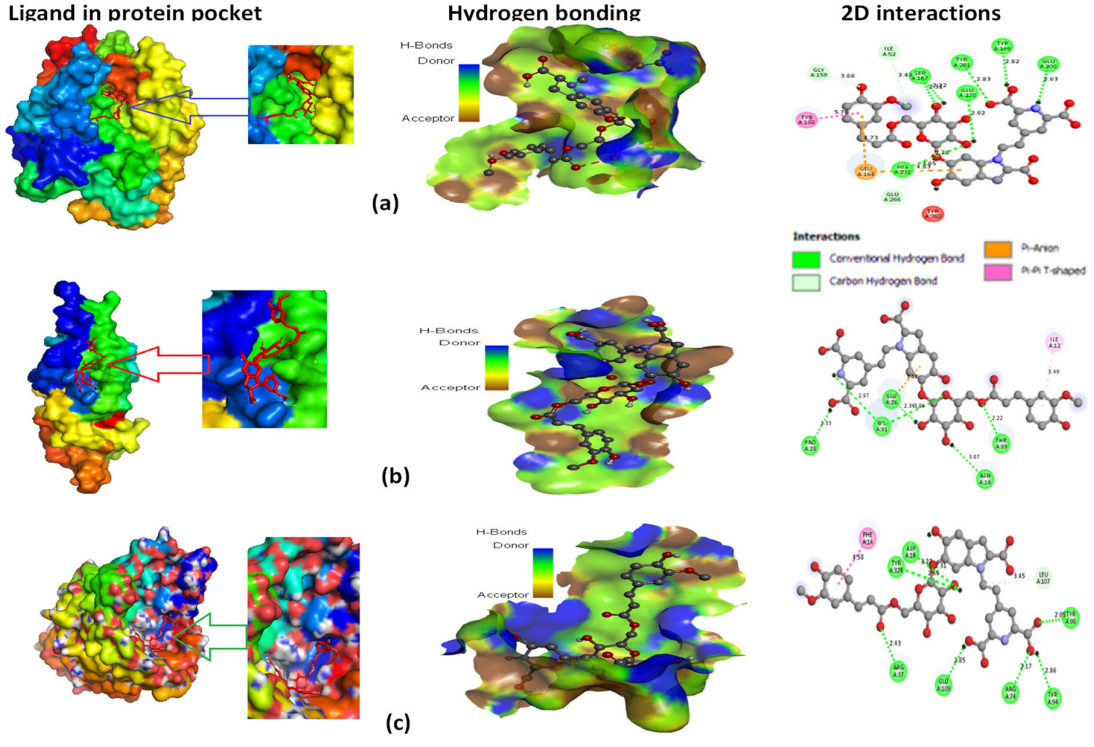


Fig. 4. Various poses after docking against (a) *C. botulinum*, (b) *E. coli*, and (c) *S. typhi* for **L-34**

Compound **L-32** was found to form twelve hydrogen bonds and no hydrophobic bonds with *S. typhi* protein (3UU2). Hydrogen bonds are formed at positions ARG37 (2.40557 A), GLN55 (3.05904), ARG74 (3.04629), ARG74 (2.96977), GLY111 (2.05138), ASP113 (2.37028), GLN123 (1.97551), ARG124 (2.15953), ARG132 (1.92716), GLY111 (2.67761), ASP105 (2.06046), and TYR98 (2.70845).

In the case of *E. coli* protein (2ZWK), eight hydrogen bonds were formed by compound **L-34** at the position ASP8 (2.68086), LYS40 (2.50927), TRP50 (2.24932), ASP8(2.53093), GLY44(2.43073), ASP45 (2.63056), LEU39 (2.25202), GLY46(3.25372) and a hydrophobic bond was formed at the position TYR48 (5.09098).

The study also found that compound **L-32** was found to form six hydrogen bonds and no hydrophobic bonds with *C. botulinum* protein (3FIE). Hydrogen bond was formed at position GLY159 (2.26503), ASN165 (2.57693), SER166 (2.26306), ARG171 (2.13738), THR194 (2.11572), and ASN165 (3.3639).

2.11. Electrostatic potential (ESP) charge distribution

A map illustrating the attractive or repulsive force experienced by a fixed charged particle (typically a point positive charge, i.e., a proton) at various points in space that are equidistant from a molecular surface. **Fig. 5** shows a 3D mapped electrostatic potential charge distribution, with blue representing a negative charge and red representing a positive charge. It is evident from the ESP that the negative charge region is greater than the positive charge region, indicating that the electrophilic groups in these molecules are more attractive.

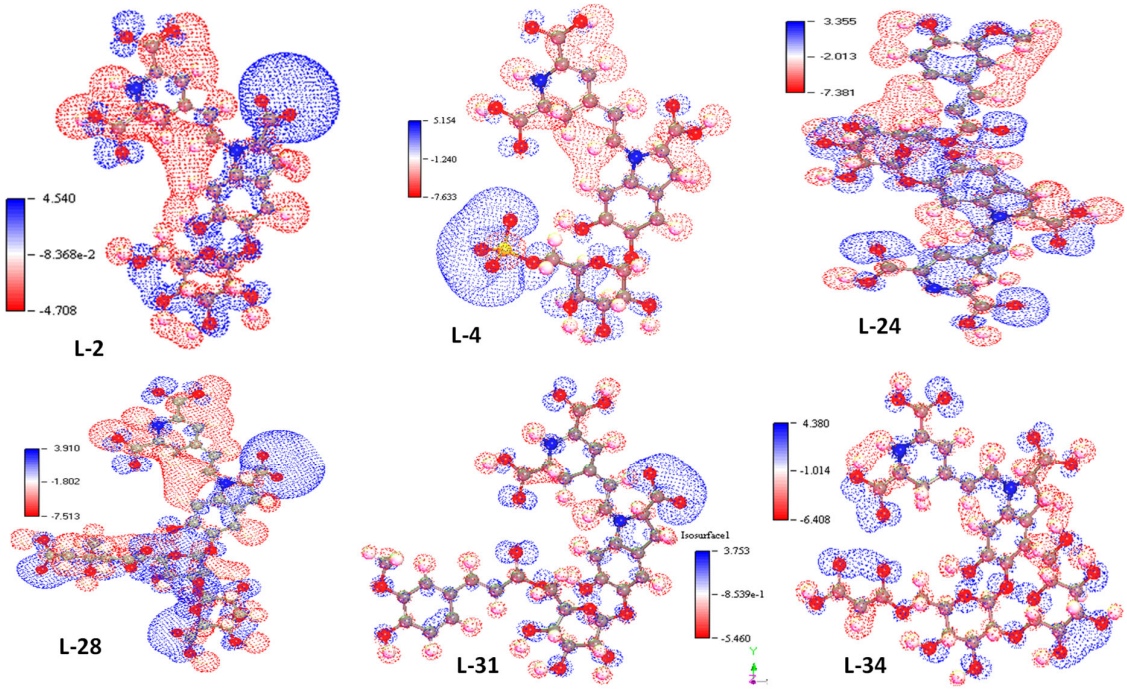


Fig. 5. 3D mapped of electrostatic potential charge distribution

2.12. Molecular dynamics simulations

RMSD stands for root-mean-square deviation and is a standard measure of structural distance. In other words, the RMSD between two sets of atomic coordinates tells us how much the protein conformation has changed. On the other hand, RMSF stands for root mean square fluctuation and is a measure of how much a residue moves (fluctuates) during a simulation. Molecular dynamics is a method for assessing docking accuracy in terms of RMSD and RMSF (RMSF).[33] Protein-ligand RMSD, ligand-protein interaction, hydrogen bonding, and ligand RMSF were all used to assess the stability of docked complexes. In this study, the RMSD was calculated for 0-100 ns and protein amino acid residue interaction.

To proceed, the RMSD is shown in **Fig. 6a-i** in terms of time and amino acid residue dependence where an innovative relationship is found for the first three figures. The RMSD is less than 2.0 shown in **Fig. 6b** (without bond or interaction) indicating its best-fitting nature with protein. But the RMSD changes after the formation of a backbone or hydrogen bond as it increases to 2.6. With an RMSD of about 2.6, hydrogen bonds have little effect on molecular docking affinity and stability, but hydrophobic bonds in protein-ligand interactions play a major role in docking score and stability. The docked complexes are more stable when their RMSF is low. The RMSF is approximately 3 when there is no bonding or interaction,

such as ligand-protein interaction. The hydrogen bond scores 2.6, indicating a low response to stability. The RMSF of derivatives **L-2** and **L-24** was around 0.7, indicating the highest stability of the docked complexes.

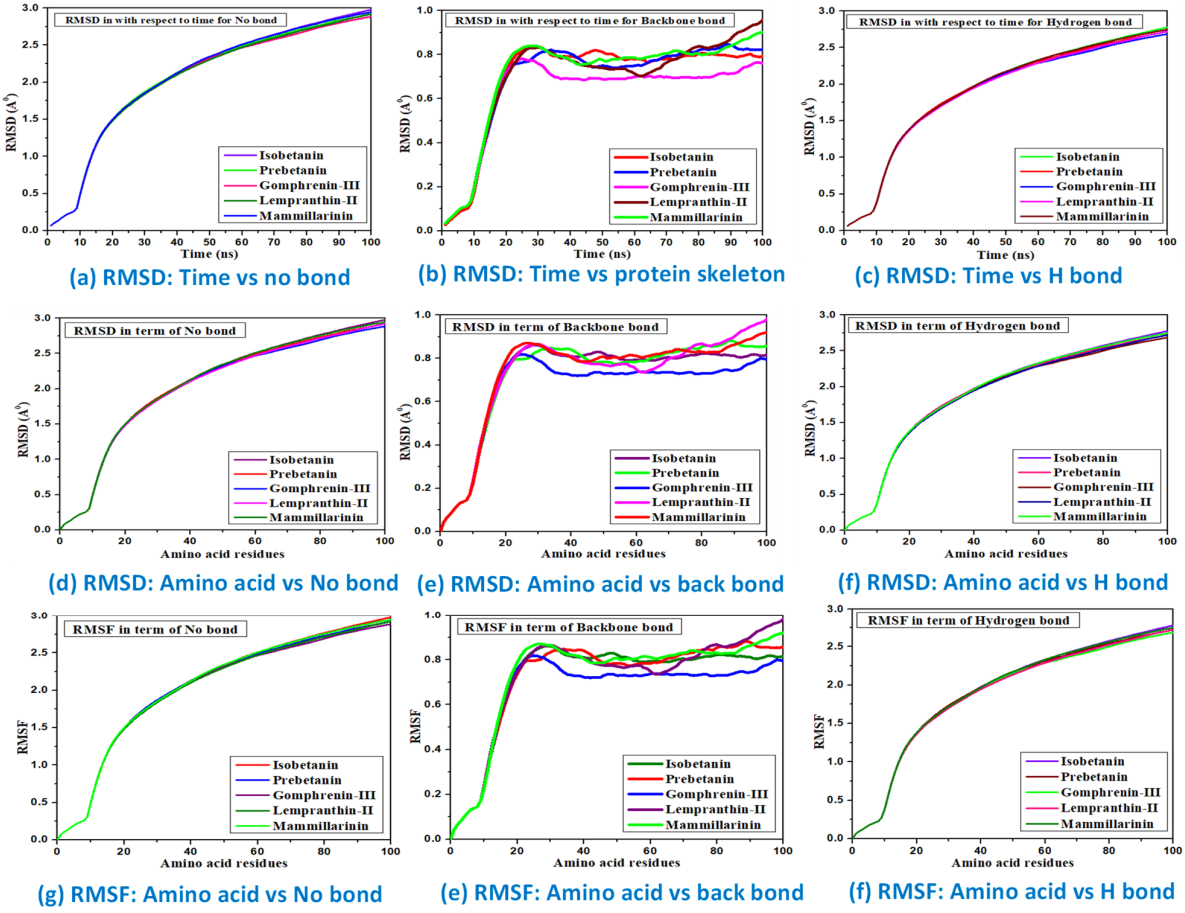


Fig. 6. Various pictures of RMSD and RMSF for *C. botulinum* protein (3FIE)

3. Conclusions

A detailed structural and antibacterial *in silico* study of the potential 36 natural betalains are described here. From PASS analyses and molecular docking best ten highly potential betalains are selected for probable future investigations. Drug-likeness, toxicity, ligand-protein interaction, QSAR, and molecular dynamics simulation indicated their standard drug nature. Although different ADME attributes of these compounds indicated a lower value, advantageously all these natural pigments are non-carcinogenic and are less hazardous in both aquatic and non-aquatic environments. As a result, these findings may aid in the development of antibacterial agents. In short, many of the food and vegetable decompositions due to bacterial invasion could be prevented by using spraying betalains type natural organic pigments, which are comparatively safe.

Acknowledgements

We are indebted to the Ministry of Science and Technology, Bangladesh for a financial grant to accomplish this work (2021-2022: Phys/624).

4. Computational details

4.1. Optimization and ligand preparation

Each of the structures of betalains (**L-1** to **L-36**) was downloaded from MDL/SDF format on the PubChem database server (https://pubchem.ncbi.nlm.nih.gov). Then, using the B3LYP/6-31G basis set in Gaussian16 software, each of the pigment structures was optimized.[34] These ligands were then prepared using the AutoDock tools. With the help of following equations, the global reactivity descriptors, such as global softness (*S*), electron affinity (*A*), ionization potential (*I*),

electronegativity (χ), global hardness (η), global electrophilicity index (ω), and chemical potential (μ) can be determined.[35-37]

$E_{(gap)} = E_{LUMO} - E_{HOMO}$ (1)
$I = -E_{HOMO}$ (2)
$A = -E_{LUMO}$ (3)
$\mu = -(I+A)/2$ (4)
$\eta = (I-A)/2$ (5)
$S = 1/\eta$ (6)
$\chi = (I+A)/2$ (7)
$\omega = \mu^2/2\eta$ (8)

4.2. Optimization and ligand preparation

The PASS (prediction of activity spectra for substances) website (http://www.way2drug.com/passonline) accurately predicts over 4000 pharmacological effects and biochemical mechanisms from a substance's structural formula. It is nearly impossible to find out the biological activity of each potential drug-like compound for research. The PASS study enables the rapid identification of drug candidates by predicting their biological potential.[38] PASS prediction studies employ Pa (probability of activity) and Pi (probability of inactivity) values to forecast whether tested biological substances belong to the active or inactive subclass. Pa and Pi have values ranging from 0.0000 to 1.0000.

4.3. Ligand-based pharmacokinetics and ADME/T properties calculation

Drug-likeness is a qualitative notion used in drug design for how "drug-like" a material is to variables like bioavailability. It is estimated from the molecular structure before synthesizing and testing the chemical. Besides, the movement of drugs in the body is known as their pharmacokinetics. SwissADME (http://www.swissadme.ch/) is used for finding the drug likeliness and pharmacokinetics properties of betalains.[39] ADMET (absorption, distribution, metabolism, excretion toxicological) properties of potential drug candidates are evaluated by the admetSAR (http://lmmd.ecust.edu.cn/admetsar2/), an online database.[31,40]

4.4. Protein preparation and docking study

Docking is a virtual screening method in molecular modeling used to predict ligand-protein interactions and the ligand's binding orientation to target proteins. Moreover, it accurately predicts the binding affinity and activity of small molecules.[41,42] For the docking protein of *E. coli* (strain O157:H7), *S. typhi*, and *C. botulinum* with protein ID: 2ZWK, 3UU2, and 3FIE, respectively were taken from RCSB protein data bank [https://www.rcsb.org/]. Following that, all three proteins were purified and saved in the PDB format using PyMOL version 4.6.[43] Then, using PyRx and AutoDock, three proteins were docked against all 36 ligands. Finally, the discovery studio was used for the visualization of protein-ligand interaction.[44]

4.5. Molecular dynamics

Since *C. botulinum* had a higher docking score against betalain compounds, MD simulations of the top 5 betalains were run interactively with live view on a high configuration laptop computer. To validate docking results for the optimal ligand-protein interaction up to 100 ns for holo-form (drug-protein) using the AMBER14 force field MD simulation was used.[45] The total system was equilibrated with 0.9 percent NaCl at 298 K temperature in the presence of a water solvent. During the simulation, a cubic cell was propagated within 20 on each side of the process and periodic boundary circumstance, and the RMSD and RMSF were analyzed using the NAMD software after simulation.

4.6. Quantitative structure-activity relationship (QSAR)

Based on mathematical and statistical correlations, quantitative structure-activity relationships (QSAR) are used to build connections between the physicochemical features of chemical compounds and their biological functions.[46-48] QSAR is characterized in computer-aided drug design by comparing it to drug-related eight descriptors obtained from ChemDes (http://www.scbdd.com/chemdes/), a free web-based platform for the calculation of molecular descriptors.[49]

References

1 Abadias M., Usall J., Anguera M., Solsona C., and Viñas I. (**2008**) Microbiological quality of fresh, minimally-processed fruit and vegetables, and sprouts from retail establishments. *Int. J. Food Microbiol*., 123 121-129.
2 Jeddi M. Z., Yunesian M., Gorji M. E., Noori N., Pourmand M. R., and Khaniki G. R. (**2014**) Microbial evaluation of fresh, minimally-processed vegetables and bagged sprouts from chain supermarkets. *J. Health Popul. Nutr.*, 32 391-399.
3 Jung Y., Jang H., and Karl R. M. (**2014**) Effect of the food production chain from farm practices to vegetable processing

on outbreak incidence. *Microb. Biotechnol*., 7(6) 517-527. OI: 10.1111/1751-7915.12178
4 Kenny M. (**2002**) Quality and safety of fresh fruits and vegetables along the production chain. *Food Nutr. Agric.,* 78-86.
5 Losio M. N., Pavoni E., Bilei S., Bertasi B., Bove D., Capuano F., Farneti S., Blasi G., Comin D., Cardamone C. et al. (**2015**) Microbiological survey of raw and ready-to-eat leafy green vegetables marketed in Italy. *Int. J. Food Microbiol*., 210 88-91.
6 Erickson M. C. (**2012**) Internalization of fresh produce by foodborne pathogens. *Annual Rev. Food Sci. Tech.*, 3 (1) 283-310.
7 Delaquis P., Bach S., and Dinu L. –D. (**2007**) Behavior of Escherichia coli O157: H7 in leafy vegetables. *J. Food Prot.*, 70 1966-1974.
8 Hirneisen K. A.., Sharma M., and Kniel K. E. (**2012**) Human enteric pathogen internalization by root uptake into food crops. *Foodborne Pathog. Dis.*, 9 396-405.
9 Jongman M., and Korsten L. (**2018**) Irrigation water quality and microbial safety of leafy greens in different vegetable production systems: A review. *Food Rev. Int.*, 34 308-328.
10 Azeredo H. M. C. (**2009**) Betalains: properties, sources, applications, and stability–a review. *Int. J. Food Sci. Tech.*, 44(12) 2365-2376.
11 Gengatharan A., Dykes G. A., and Choo W. S. (**2015**) Betalains: Natural plant pigments with potential application in functional foods. *LWT-Food Sci. Tech.*, 64 645-649.
12 Stintzing F. C., and Carle R. (**2004**) Functional properties of anthocyanins and betalains in plants, food, and in human nutrition. *Trends Food Sci. Tech.*, 15 19-38.
13 Madadi E., Mazloum-Ravasan S., Yu J. S., Ha J. W., Hamishehkar H., and Hyun K. K. (**2020**) Therapeutic application of betalains: A review. *Plants*, 9(9) 1219.
14 Rodríguez-Ramiro I., Ramos S., Bravo L., Goya L., and Martín M. Á. (**2012**) Procyanidin B2 induces Nrf2 translocation and glutathione S-transferase P1 expression via ERKs and p38-MAPK pathways and protect human colonic cells against oxidative stress. *Eur. J. Nutr.*, 51 881–892.
15 Ooms F. (**2000**) Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. *Cur. Med. Chem.*, 7(2) 141-158.
16 Uzzaman M., Hasan M. K., Mahmud S., Fatema K., and Matin M. M. (**2021**) Structure-based design of new diclofenac: Physicochemical, spectral, moleculardocking, dynamics simulation and ADMET studies. Informatics Med. Unlocked, 25 100677.
17 Yang H., Lou C., Sun L., Li J., Cai Y., Wang Z., Li W., Liu G., and Tang Y. (**2019**) admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics*, 35(6) 1067-1069.
18 Kumer A., Chakma U., Matin M. M., Akash S., Chando A., and Howlader D. (**2021**) The computational screening of inhibitor for black fungus and white fungus by D-glucofuranose derivatives using in silico and SAR study. *Org. Commun.*, 14(4) 305-322.
19 Matin M. M., Ibrahim M., Anisa T. R., and Rahman M. R. (**2022**) Synthesis, characterization, in silico optimization, and conformational studies of methyl 4-O-palmitoyl-α-L-rhamnopyranosides. *Malaysian J. Sci.*, 41(1) 91-105.
20 Al-Karmalawy A. A., Dahab M. A., Metwaly A. M., Elhady S. S., Elkaeed E. B., Eissa I. H., and Darwish K. M. (**2021**) Molecular docking and dynamics simulation revealed the potential inhibitory activity of ACEIs against SARS-CoV-2 targeting the hACE2 receptor. *Front. Chem.*, 9 661230.
21 Kumer A., and Khan M. W. (**2021**) Synthesis, characterization, antimicrobial activity and computational exploirations of ortho toludinium carboxylate ionic liquids. *J. Mol. Struct.*, 1245 131087.
22 Nath A., and Kumer A. (**2021**) Investigating the binding affinity, molecular dynamics, and ADMET properties of 2,3-dihydrobenzofuran derivatives as an inhibitor of fungi, bacteria, and virus protein. *Beni-Suef Univ. J. Basic Appl. Sci.*, 10 1-13.
23 Hoque M. M., Kumer A., Hussen M. S., and Khan M. W. (**2021**) Theoretical evaluation of 5, 6-diaroylisoindoline-1,3-dione as potential carcinogenic kinase PAK1 inhibitor: DFT calculation, molecular docking study and ADMET prediction. *Int. Adv. Biol. Biomed. Res.*, 9 77-104.
24 Ahmed M. K., and Kumer A. (**2021**) Facile fabrication of polymer network using click chemistry and their computational study. *Royal Soc. Open Sci.*, 8 202056.
25 Islam M. J., Kumer A., Paul S., and Sarker M. N. (**2020**) The activity of alkyl groups in morpholinium cation on chemical reactivity, and biological properties of morpholinium tetrafluroborate ionic liquid using the DFT method. *Chem. Method.*, 4 130-142.
26 Kumer A., Paul S., Sarker M. N., and Islam M. J. (**2019**) The prediction of thermo physical, vibrational spectroscopy, chemical reactivity, biological properties of morpholinium borate, phosphate, chloride and bromide ionic liquid: A DFT study. *Int. J. New Chem*., 6 236-253.
27 Zhang M. -Q., and Wilkinson B., (**2007**) Drug discovery beyond the ‘rule-of-five’. *Curr. Opin. Biotech.*, 18 478-488.
28 Rahman M. A., Matin M. M., Kumer A., Chakma U., and Rahman M. R. (**2022**) Modified D-glucofuranoses as new black fungus protease inhibitors: Computational screening, docking, dynamics, and QSAR study. *Phys. Chem. Res.*, 10(2) 189-203.
29 Martin Y. C. (**2005**) A bioavailability score. *J. Med. Chem.*, 48 3164-3170.
30 Islam N., Islam M. D., Rahman M. R., and Matin M. M. (**2021**) Octyl 6-O-hexanoyl-β-D-glucopyranosides: Synthesis,

PASS, antibacterial, in silico ADMET, and DFT studies. *Curr. Chem. Lett.*, 10(4) 413-426.
31 Cheng F., Li W., Zhou Y., Shen J., Wu Z., Liu G., Lee P. W., and Tang Y. (**2012**) admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. *J. Chem. Inf. Model.*, 52(11) 3099–3105.
32 Rakib A., Nain Z., Sami S. A., Mahmud S., Islam A., Ahmed S., Siddiqui A. B. F., Babu S. M. O. F., Hossain P., Shahriar A., Nainu F., Emran T. B., and Simal-Gandara J. (**2021**) A molecular modelling approach for identifying antiviral selenium-containing heterocyclic compounds that inhibit the main protease of SARS-CoV-2: An in silico investigation. *Brief, Bioinform*., 22 1476-1498,
33 Trott O., and Olson A. J. (**2010**) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.*, 31 455-461.
34 Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., et al., (**2009**) Gaussian 09W. Gaussian, Inc., Wallingford CT, USA.
35 Ayers P. W., and Parr R. G. (**2000**) Variational principles for describing chemical reactions: the Fukui function and chemical hardness revisited. *J. Am. Che. Soc.*, 122 2010-2018.
36 Kohn W., Becke A. D., Parr R. G. (**1996**) Density functional theory of electronic structure. *J. Phys. Chem.*, 100 12974-12980.
37 Parr R.G., von Szentpaly L. and Liu S. (**1999**) Electrophilicity Index. *J. Am. Che. Soc.*, 21 1922–1924.
38 Filimonov D. A., Lagunin A. A., Gloriozova T. A., Rudik A. V., Druzhilovskii D. S., Pogodin P. V., and Poroikov V. V. (**2014**) Prediction of the biological activity spectra of organic compounds using the PASS online web resource. *Chem. Heter. Comp.*, 50 444-457.
39 Daina A., Michielin O., and Zoete V. (**2017**) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.*, 7 42717.
40 Matin M. M., Islam N., Siddika A., and Bhattacharjee S. C. (**2021**) Regioselective synthesis of some rhamnopyranoside esters for PASS predication, and ADMET studies. *J. Turk. Chem. Soc. Sect. A: Chem.*, 8(1) 363-374.
41 Vijesh A. M., Isloor A, M., Telkar S., Arulmoli T., and Fun H. -K (**2013**) Molecular docking studies of some new imidazole derivatives for antimicrobial properties. *Arabian J. Chem.*, 6 197-204.
42 Ristovski J. T., Matin M. M., Kong R., Kusturica M. P., and Zhang H. (**2022**) In vitro testing and computational analysis of specific phytochemicals with antiviral activities considering their possible applications against COVID-19. *South African J. Botany*, Accepted manuscript (DOI: 10.1016/j.sajb.2022.02.009).
43 Pasi M., Matteo T., Arrigoni A., and Papaleo E. (**2012**) xPyder: a PyMOL plugin to analyze coupled residues and their networks in protein structures. *J. Chem. Inf. Medel.*, 52(7) 1865-1874.
44 Discovery Studio 2.1 (**2008**) (Accelrys, San Diego, CA, USA)
45 Skjevik A. A., Madej B. D., Dickson C. J., Teigen K., Walker R. C., and Gould I. R. (**2015**) All-atom lipid bilayer self-assembly with the AMBER and CHARMM lipid force fields. *Chem. Commun*., 51 4402-4405.
46 Verma J., Khedkar V. M., and Coutinho E. C. (**2010**) 3D-QSAR in drug design- A review. *Curr. Top. Med. Chem.*, 10(1) 95-115.
47 Matin M. M., Bhattacharjee S. C., Chakraborty P., and Alam M. S. (2019) Synthesis, PASS predication, in vitro antimicrobial evaluation and pharmacokinetic study of novel n-octyl glucopyranoside esters. *Carbohydr. Res.*, 485 107812.
48 Peter S. C., Dhanjal J. K., Malik V., Radhakrishnan N., Jayakanthan M., and Sundar D. (**2019**) Quantitative structure-activity relationship (QSAR): modeling approaches to biological applications. 2 661-676.
49 Dong J., Cao D. S., Miao H. Y., Liu S., Deng B. C., Yun Y. H., Wang N. N., Lu A. P., Zeng W. B., and Chen A. F. (**2015**) ChemDes: An integrated web-based platform for molecular descriptor and fingerprint computation. *J. Cheminform.*, 7 1-10.