Int J Clin Exp Med 2021;14(12):2620-2630
www.ijcem.com /ISSN:1940-5901/IJCEM0132009

Original Article

Identification of plant-based inhibitors by molecular docking against catalytic glucosyltransferase domains of TcdA and TcdB toxins from *clostridium difficile*

Khalid M Aljarallah[1,2], Sezanur Rahman[3]

[1]*College of Applied Medical Sciences, Majmaah University, University Campus, Majmaah, Riyadh Province, Saudi Arabia;* [2]*College of Applied Sciences, Almaarefa University, King Khalid Road, AlDere'yah, Riyadh;* [3]*International Centre for Diarrhoeal Disease Research, Bangladesh*

Received February 18, 2021; Accepted November 1, 2021; Epub December 15, 2021; Published December 30, 2021

Abstract: As a frequent cause of colitis and nosocomial infections, *Clostridium difficile* continues to be an important pathogen of concern. To the authors' knowledge, no specific medication is available to counteract the effect of *C. difficile* toxins. This study aimed to identify the most potential antagonists for the two major toxins of *C. difficile*: TcdA and TcdB. A library of plant-derived compounds (N=8,858) was screened for toxin inhibitors using a series of computational methods. In standard-precision docking, the binding affinity for TcdA inhibitors was lower (-13.265 to -12.132 kcal/mol) than the control UDP-glucose (-9.62 kcal/mol). Similarly, the binding affinity for TcdB inhibitors was lower (-6.995 to -5.395 kcal/mol) than the control apigenin (-4.377 kcal/mol). The Gibbs free energy (ΔG) was also lower for TcdA inhibitors (-116.08 to -67.31 kcal/mol) and TcdB inhibitors (-59.88 to -35.06 kcal/mol) as compared to their respective controls (-13.82 and -41.26 kcal/mol, respectively). Based on interactions of these toxins and drug candidacy, the top 10 compounds for each target, providing higher binding affinities for TcdA and TcdB as compared to their positive controls, were identified. An evaluation of the absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of these selected compounds suggests them to be good drug candidates against *C. difficile* toxins. Notably, one compound, TIP009928, was common in both selected inhibitor pools, and may be a good drug candidate against both TcdA and TcdB toxins. This finding is based on chemoinformatics and therefore requires validation with *in vivo* and *in vitro* experiments before these potential inhibitors could be considered as prospective therapeutics for *C. difficile* infections.

Keywords: *Clostridium difficile,* TcdA, TcdB, phytochemicals, molecular docking, small molecules

Introduction

Clostridium difficile is a gram-positive, spore-producing bacterium that causes diarrhea and intestinal inflammation, known as colitis, in patients who have undergone antibiotic therapy [1-3]. Globally, the *C. difficile* incidence rate was estimated by a meta-analysis study [3] to be 8.3 cases per 10,000 patient-days, while in the United States, the incidence was estimated to be 453,000, with 29,300 deaths [4], making *C. difficile* one of the most frequent and costly nosocomial emerging infections [5]. The pathogenic invasion of *C. difficile* depends on a variety of virulence factors, including toxins, spore formation, adherence, and motility [2, 6]. In a hostile environment, *C. difficile* produces toxins that target intestinal epithelial cells, and this is followed by endocytosis, the activation of catalytic domains, and tissue necrosis [2]. Therefore, the use of inappropriate antibiotics may increase toxin production [7]. Also, recently, the antibiotic resistance of *C. difficile* is increasing [8]. Further, its spore-forming ability helps this pathogen survive in unsuitable conditions and facilitates nosocomial and community transmission [2].

The *C. difficile* encodes five pathogen-related proteins [9]. Among them, two major homologous toxins are Toxins A (TcdA) and B (TcdB), both belonging to the family of clostridial toxins [10]. The lengths of the TcdA and TcdB proteins were 537 and 540 amino acid residues, respec-

tively, and both toxins contain glucosyl-transferase domains, which inactivate key signaling molecules within the host [2, 10]. For instance, once inside the cell, they deactivate the cytosolic Rho/Ras GTPases that disrupt the cytoskeleton and cellular junctions, leading to a loss of epithelial integrity [9]. Nevertheless, a multi-laboratory follow-up study revealed TcdB as the principal virulence factor, while TcdA was a minor influencer of intestinal inflammation in animal models [10, 11].

The infection caused by *C. difficile* becomes challenging to treat due to its increased resistance to multiple classes of antibiotics, including metronidazole and vancomycin [12], which necessitates the need for alternative therapeutics [1, 8, 13], including candidate vaccines [14, 15] and plant-based medicines [16, 17]. Since toxins cause *C. difficile* colitis, researchers are trying to identify small molecule inhibitors that specifically target these toxins to deactivate their glucosyltransferase activities [18]. In this case, computer-aided drug design and virtual screening may be an important approach to filter out potential drug candidates from large databases. Interestingly, plant-based chemicals may be a potential source of such inhibitors, as they are less likely to produce unwanted side-effects. In recent years, the computational approach has become a powerful tool in the field of drug discovery by reducing time, effort, and costs involved. For instance, several drugs, such as PRX-03140, PRX-00023 in Phase IIb, and Aggrastat, were successfully discovered through the virtual screening processes [19]. Furthermore, molecular docking and recognition can play significant roles in understanding the complex relationships in the biological system.

In this study, we intend to identify potential phytochemicals that can act against *C. difficile* toxins, and the identified phytochemicals will also be assessed for their drug-likeness by a computational approach.

Materials and methods

Collection and preparation of protein

The crystal (3D) structures of TcdA (PDB# 3SRZ; Res: 2.58 Å) and TcdB (PDB# 5UQT; Res: 2.75 Å) were collected from the Protein Data Bank (https://www.rcsb.org/) and processed for further computational analysis by the removing water molecules around the active site (>5Å). The Protein Preparation Wizard (Mestro v11.8, Schrodinger) was employed to optimize and minimize in terms of proper bonding orders, the addition of hydrogen (H) atoms and missing residues and the generation of tautomer/ionization states at the physiological pH of 7.0 [20]. Lastly, the crystal structure was minimized (RMSD cutoff =0.30 Å) by applying the OPLS_2005 force field [21].

Generation of receptor grid

A 3D cubic space (called the grid-box) around the active site of the protein, where the ligands are supposed to interact was identified for targeted molecular docking. Mestro's Glide v8.1 module was used to define the grid-box by using the co-crystal ligands of the TcdA and TcdB proteins. Furthermore, the cutoff value for the Van der Waals interactions was set to 1.0 Å, while 0.25 was kept as the threshold for the partial atomic charge. Moreover, no constraints or rotatable groups were defined.

Ligand selection and preparation

The TIPdb database (https://cwtung.kmu.edu.tw/tipdb/), which is manually curated, contains a variety of indigenous and endemic plant species of Taiwan [22]. All of the compounds from this database were downloaded in SDF format, containing the 3D structure, and an internal library was developed using the LigPrep v2.2 module (Maestro v11.8). This was followed by ligand preparation. The Epik v4.6 software [23] was applied to desalt the ligands and to generate ionization/tautomeric states at a physiological pH of 7.0±2.0. The 3D structures of the ligands were used to anticipate their chirality.

Molecular docking studies

The binding interactions between TcdA/TcdB and the ligand molecules were estimated with the Glide v8.1 module [24], where the co-crystal ligands, UDP-glucose analogue and apigenin, were considered as the controls for TcdA and TcdB, respectively. An extra-precision (XP) docking was applied to the top 10 identified candidates, that is, from the molecules showing greater binding affinities as compared to the controls in standard-precision (SP) dockings against toxin proteins. The binding affinity

was calculated with a standard-precision method, where results were provided by a score (G-score) assigned to each ligand. A lower G-score indicates a higher affinity for the molecule. The docked complexes were analyzed and visualized for binding interactions and associated residues with Discovery Studio 2017 (BIOVIA, USA).

Estimation of binding free energy

Docking interaction was then validated by calculating the binding Gibbs free energy (ΔG) of the docked complexes using the Prime/MM-GBSA method (Maestro v11.8). The optimization and minimization of the docked complexes were done by using the OPLS_2005 force field [25]. Then, ΔG was estimated via the combination of OPLS molecular mechanics energies, a polar solvation term, a nonpolar solvation terms, i.e., the solvent-accessible surface area and Van der Waals interactions. The simulations were executed with the GBSA continuum solvent model [26].

Prediction of ADMET properties

Based on the G-scores, we selected potential ligands and evaluated them on their absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties to identify the most important and most suitable drug leads [25]. The QikProp v5.8 module of Maestro v11.8 was used to evaluate the drug-likeness features of the potential ligands [27]. This module anticipates a total of 44 features that were considered while designing and finding drug leads. Finally, we compared those criteria with their respective optimal values found in 95% of known drugs [25]. Also, we applied Lipinski's rule of five [28] to filter out potential leads that are important in rational drug design [25].

Results

Proteins and ligand database

For docking studies, we optimized and minimized the TcdA and TcdB proteins energetically. An internal library, containing 8,858 phytochemicals with a total of 19,171 conformational states, was created in the LigPrep module. For ligand-based virtual screening, the whole library was docked into the generated grid-box area in the TcdA and TcdB active site.

Binding affinity and interactions

Molecular docking was performed to observe the efficiency of phytochemicals to inhibit the toxins, TcdA and TcdB, from *C. difficile* through competitive binding. In standard-precision (SP) docking against the TcdA protein, a total of 895 molecules showed greater binding affinity as compared to control UDP-glucose (-9.62 kcal/mol). After extra-precision (XP) docking, 10 candidate molecules with a greater binding affinity (-16.369 to -19.432 kcal/mol) were selected towards TcdA (**Table 1**). Similarly, a total of 883 molecules showed greater binding affinity in the SP docking for TcdB inhibitors. After XP docking, the top 10 potential inhibitors were selected with binding affinities ranging from -7.522 to -5.433 kcal/mol, which were lower than those of the control, apigenin, which had a binding affinity of -3.186 kcal/mol (**Table 1**).

Most of the molecules from the selected panel (**Table 1**) against TcdA had more active site residues than the control. There were two exceptions: TIP006135 and TIP009947. For instance, the native ligand produced a total of 23 contacts, including 20 hydrogen bonds (H-bonds), and interacted with 13 binding residues, whereas the TIP006588 molecule provided a total of 33 contacts (highest), including 24 H-bonds, that interacted with 15 binding residues. Notably, the binding patterns of the selected 10 molecules were similar to those of the native ligand. For example, the most common residues were Asp285 and Ile102, which were prevalent in all docked complexes (12), including the control, followed by Leu264 (11), Trp101 (11), Trp519 (10), Ser268 (10), Leu518 (9), Asp269 (8), Val100 (8), Asn516 (7), and Arg272 (6). Conversely, Lys141 and Gly469 were the least frequent residues, interacting only once with TIP006135 and TIP010278. Other less prevalent residues include Asp287, Gln513, and Ile382, which were found with three compounds while Ile465 was present with three compounds. Altogether, our results suggest 11 potential compounds which exhibit binding patterns that were similar to those of the native ligands, offering higher efficiency, and indicated the proper binding with the active site residues. **Figure 1** shows the interactions of the best three compounds with TcdA, along with the control ligand.

Table 1. Binding affinities, interactions, and residues involved between inhibitors and toxin (TcdA and TcdB)

Compounds	SP Docking (kcal/mol)	XP Docking (kcal/mol)	Total contacts (n)	H-bonds	Interacting residues
Docking result for TcdA toxin					
UDP-glucose	-9.62	-15.273	23	20 (15)	Arg272, Arg462, Asn383, Asp269, Asp285, Glu514, Ile102, Leu264, Pro470, Ser268, Trp101, Trp519, Val100 (13)
TIP006135	-12.534	-18.622	22	12 (8)	Ala265, Arg272, Asn138, Asp269, Asp285, Ile102, Leu264, Leu518, Lys141, Ser268, Trp101, Trp519, Val100 (13)
TIP006588	-12.629	-18.93	33	24 (15)	Ala377, Asn516, Asp269, Asp285, Asp522, Glu514, Ile102, Ile382, Ile465, Pro470, Ser517, Trp101, Trp519, Tyr283, Val100 (15)
TIP007177	-12.512	-18.634	31	18 (14)	Ala265, Arg272, Asn138, Asn516, Asp269, Asp285, Asp287, Gln513, Ile102, Leu264, Leu518, Ser268, Ser517, Trp101, Trp519 (15)
TIP007178	-12.429	-19.432	31	18 (14)	Ala265, Arg272, Asn138, Asn516, Asp269, Asp285, Asp287, Gln513, Ile102, Leu264, Leu518, Ser268, Ser517, Trp101, Trp519, Val100 (16)
TIP007429	-12.306	-18.76	24	13 (10)	Ala265, Arg272, Asn138, Asp269, Asp285, Ile102, Leu264, Leu518, Ser268, Trp101, Trp519, Val100 (12)
TIP009269	-12.47	-17.101	24	11 (6)	Ala265, Arg462, Asn383, Asp285, Ile102, Leu264, Leu518, Ser268, Trp101, Trp519 (10)
TIP009928	-12.458	-16.369	27	18 (12)	Ala265, Ala377, Arg462, Asn516, Asp269, Asp285, Glu514, Ile102, Ile382, Ile465, Leu264, Leu518, Pro470, Ser268, Trp101, Trp519, Val100 (17)
TIP009947	-12.132	-16.954	22	15 (11)	Ala265, Arg272, Arg462, Asn383, Asn516, Asp285, Gln513, Ile102, Leu264, Leu518, Ser268, Trp101 (12)
TIP010278	-13.265	-16.825	25	15 (8)	Ala265, Asn516, Asp269, Asp285, Gly469, Ile102, Leu264, Leu518, Pro470, Ser268, Trp101 (11)
TIP012332	-13.116	-17.793	26	14 (9)	Ala265, Asn138, Asn516, Asp285, Asp287, Ile102, Leu264, Leu518, Ser517, Trp101, Trp519,Val100 (12)
Docking result for TcdB toxin					
Apigenin	-4.377	-3.186	4	4 (3)	Arg445, Agi602, Gly444, Ala441, and Ile491 (5)
TIP006360	-6.995	-7.522	19	16 (11)	Ala439, Ala441, Arg445, Asn440, Asn490, Gly444, His492, Ile491, and Met436 (9)
TIP009928	-6.069	-5.545	14	12 (8)	Ala441, Arg445, Asn440, Glu449, and Ile491 (5)
TIP006585	-6.015	-6.101	14	11 (5)	Ala439, Ala441, Arg445, Glu449, and Ile491 (5)
TIP010071	-5.832	-7.315	16	13 (8)	Ala441, Arg445, Asn490, Glu449, and Gly444 (5)
TIP003016	-5.735	-5.827	11	8 (6)	Ala439, Ala441, Arg445, Gly444, Met436, and Met448 (6)
TIP006587	-5.548	-6.274	13	12 (7)	Ala441, Arg445, Asn440, Glu449, Glu472, Gly444, Lys452, and Met448 (8)
TIP011613	-5.545	-7.017	16	14 (9)	Ala439, Ala441, Arg445, Glu449, Gly444, Lys452, And Met448 (7)
TIP005703	-5.527	-5.433	4	3 (2)	Glu472, Asn467, Ala441, and Arg445 (4)
TIP006359	-5.467	-6.193	13	9(7)	Arg445, Asn443, Asn490, Glu449, Met448, and Met489 (6)
TIP007931	-5.395	-5.766	15	11 (8)	Ala441, Arg445, Asn490, Glu449, Gly444, and Met448 (6)

The calculation of interactions and involved residues were based on the XP docking and executed in Discovery Studio 2017. Abbreviations: SP, standard-precision docking; XP, extra-precision docking; H-bonds, hydrogen bonds; and CHB, conventional hydrogen bond; UDP-glucose, uracil-diphosphate glucose.

A similar result was observed for the TcdB protein, with much higher interactions for the potential candidates than the control, apigenin. For instance, the most potent inhibitor, TIP006360, interacted with TcdB through a total of 19 contacts (highest), while apigenin provided only 4 contacts. Moreover, most of them shared several active site residues with the native ligand. The residue, Arg445, was found to be involved in the interactions of all of the docked complexes. Other prevalent residues include Ala441 and Gly444. The number of residues involved in the apigenin-TcdB complex was four, while TIP006360 showed interactions to nine residues, which was followed by TIP006587 (8) and TIP011613 (7), as listed in **Table 1**. **Figure 2** shows the interactions of the best three compounds with TcdB, along with the control ligand.

Binding free energy and post-docking validation

The calculated binding energy (ΔG_{Bind}) for the TcdA inhibitors ranged from -116.08 to -67.31 kcal/mol, lower than the control (-13.82 kcal/

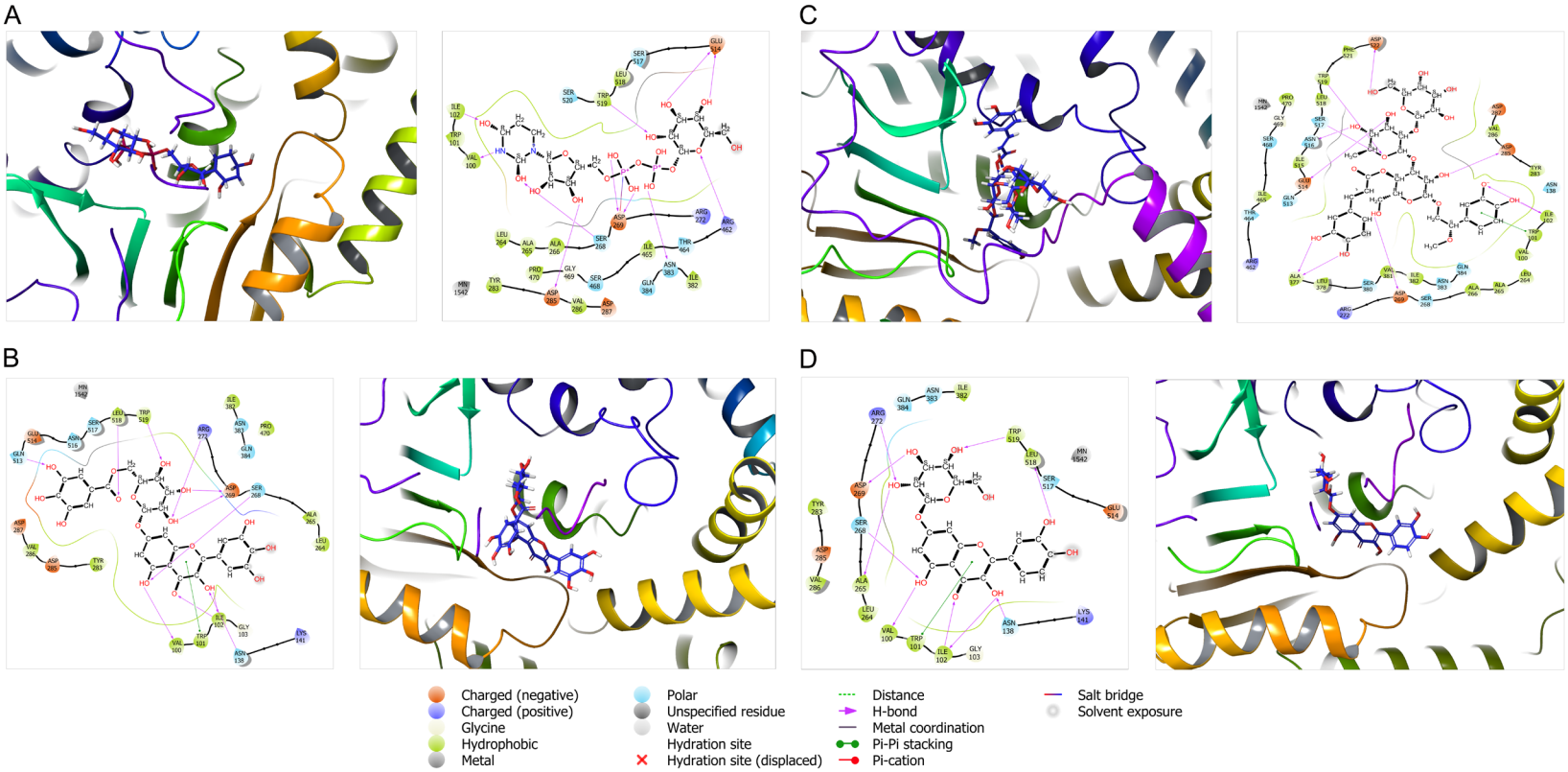

Figure 1. Molecular docking between TcdA and its inhibitors. The binding interactions between (A) TcdA protein and UDP-glucose analogue (control); (B) TcdA and compound TIP007178; (C) TcdA protein and compound TIP006588; and (D) TcdA and compound TIP007429. The docking illustration is prepared in Schrodinger's Maestro.

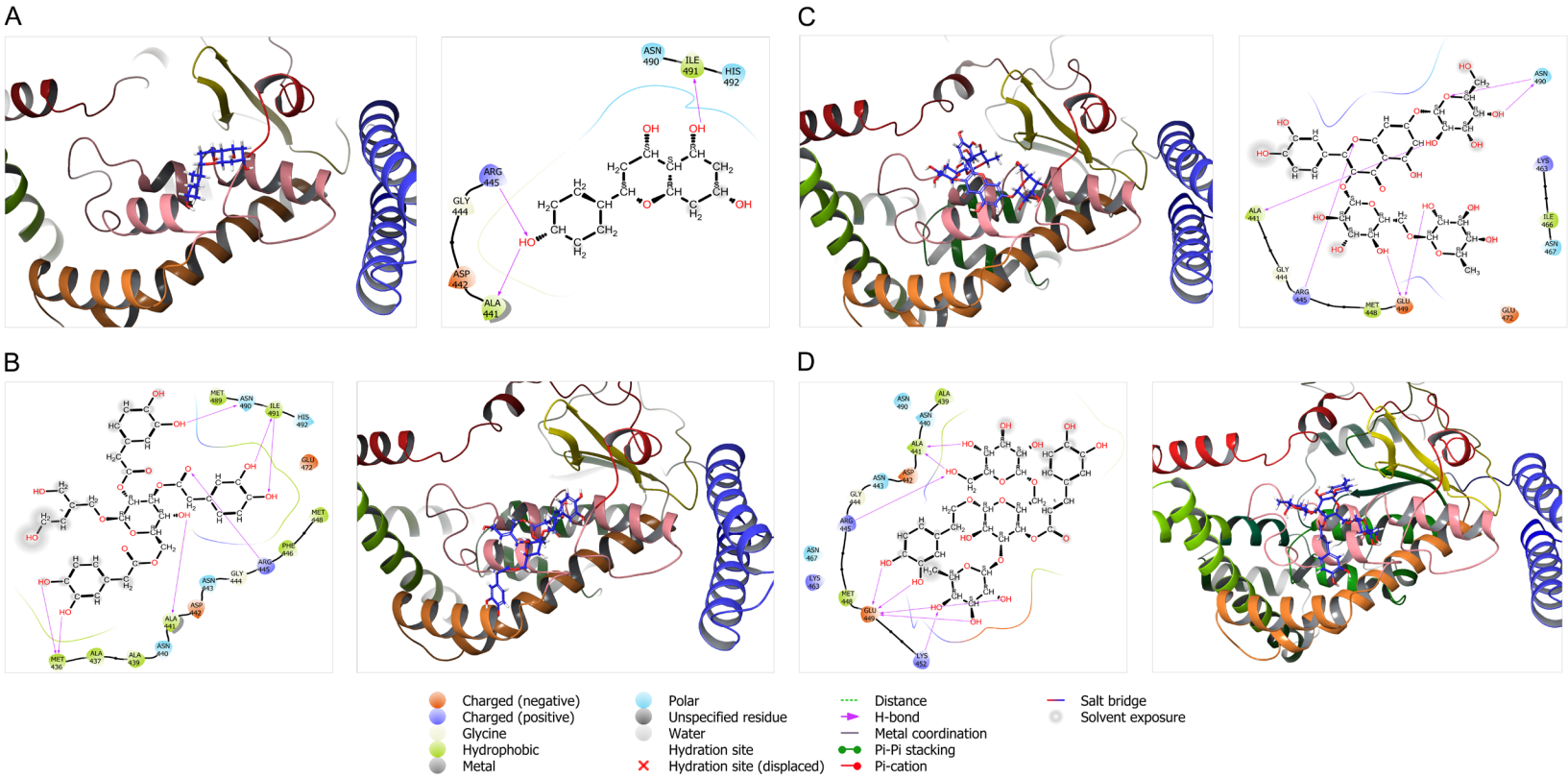

Figure 2. Molecular docking between TcdB and its inhibitors. The binding interactions between (A) TcdB protein and apigenin (control); (B) TcdB and compound TIP006360; (C) TcdB protein and compound TIP010071; and (D) TcdB and compound TIP011613. The docking illustration is prepared in Schrodinger's Maestro.

Table 2. The post-docking binding free energy of the top 10 inhibitors of TcdA and TcdB protein

Compounds	ΔG_{Bind}	ΔG_{Coul}	ΔG_{Coval}	ΔG_{Hbond}	ΔG_{Lipo}	ΔG_{SolvGB}	ΔG_{vdW}	SE_{Lig}
post-docking binding free energy of the top 10 inhibitors of TcdA								
UDP-glucose	-13.82	175.16	14.59	-6.1	-19.5	-129	-48.96	27.31
TIP009928	-100.55	-40.77	13.5	-4.37	-54.82	60.28	-71.61	3.7
TIP007429	-74.47	-40.93	16.26	-4.46	-35.09	43.97	-50.71	18.38
TIP012332	-67.31	-37.84	2.83	-4.48	-26.71	51.98	-51.84	9.59
TIP006135	-77.3	-26.02	10.24	-5.65	-35.29	36.23	-53.44	3.93
TIP009947	-93.31	-55.37	7.23	-4.67	-38.54	53.04	-51.91	13.32
TIP010278	-83.15	-47.86	20.53	-4.54	-53.15	56.55	-52.34	24.18
TIP009269	-79.82	-7.22	9.92	-7.45	-35.53	16.68	-52.36	8.77
TIP007177	-77.6	-19.22	11.31	-5.94	-40.12	47	-66.88	11.22
TIP007178	-82.29	-22.56	8.22	-7.38	-39.19	45.63	-63.68	14.75
TIP006588	-116.08	-32.92	13.39	-3.07	-79.11	55.83	-70.22	24.31
post-docking binding free energy of the top 10 inhibitors of TcdB								
Apigenin	-41.26	-16.79	2.49	-0.88	-23.68	16.89	-19.29	2.39
TIP006360	-44.43	-40.1	5.38	-2.5	-16.31	33.76	-22.98	22.03
TIP010071	-40.7	-11.78	-5.46	-3.06	-15.35	28.72	-33.6	16.97
TIP011613	-51.03	-27.59	2.01	-2.88	-25.63	35.08	-31.91	24.9
TIP006587	-54.7	-22.91	-0.43	-4.85	-26.15	40.19	-40.38	28.56
TIP006359	-51.34	-35.28	8.41	-3.65	-26.91	40.04	-32.97	32.22
TIP006585	-45.01	-19.95	-5.95	-3.4	-13.6	23.29	-24.07	24.64
TIP003016	-57.37	-28.86	6.91	-1.25	-26	25.7	-33.87	9.86
TIP007931	-49.28	13.87	10.04	-6.68	-19.56	-24.86	-22.09	10.56
TIP009928	-59.88	-45.59	-3.51	-3.61	-19.4	44.45	-31.73	16.12
TIP005703	-35.06	-66.18	2.13	-1.11	-18.33	62.84	-14.04	2.34
TIP006360	-44.43	-40.1	5.38	-2.5	-16.31	33.76	-22.98	22.03

The calculation of binding free energy and related energy statistics were based on the XP docking and executed in Maestro's Prime/MM-GBSA module. Abbreviations: ΔG_{Bind}, MM-GBSA free binding energy; ΔG_{Coul}, Coulomb energy of the complex; ΔG_{vdW}, van der Waals energy of the complex; ΔG_{Lipo}, lipophilic energy of the complex; ΔG_{SolvGB}, solvation energy of the complex; SE_{Lig}, strain energy of the ligands.

mol), while for the TcdB inhibitors it ranged from -35.06 to -59.88 kcal/mol, comparable with the control (-41.26 kcal/mol), as can be seen in **Table 2**. These results are completely in line with the docking studies. In brief, the ΔG of the H-bonds were also comparable (-3.07 to -7.45 kcal/mol) to those of the native ligand (-6.10 kcal/mol) for the TcdA inhibitor. Similarly, the ΔG of the covalent and Van der Walls interactions was close to those of the control ligand. Importantly, the strain energies of all ligands (3.70-24.31 kcal/mol) were lower than the 27.31 kcal/mol of the positive control (**Table 2**). Unlike the control, however, the solvation energy (ΔG_{SolvGB}) was positive for all potential compounds, which indicates their endothermic binding. Likewise, the ΔG of the Coulomb interactions of all compounds was negative, in contrast to the native ligand.

As for the TcdB inhibitors, the ΔG of these inhibitors was also consistent with their respective docking scores. For instance, the TIP009928 inhibitor revealed -59.88 kcal/mol of binding free energy, while it provided -5.545 kcal/mol of binding affinity.

Pharmacological profile and drug-likeness

The drug-like properties of potential compounds selected by SP and XP dockings were further considered in an absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties evaluation (**Table 3**). Most of the evaluated compounds of the TcdA inhibitor had molecular weights below 725 kDa, except for TIP009928, while only four compounds of the TcdB inhibitor were within the accepted range. Also, the solubility index of

Table 3. The drug-like properties of the 10 inhibitors of TcdA and TcdB protein

Properties (Optima)	Rotor (0-15)	CNS (-2-2)	MW (130-725)	SASA (300-1000)	dHB (0-6)	aHB (2-20)	logPC16 (4-18)	logPoct (8-35)	logPw (4-45)	logPo/w (-2-6.5)	logS (-6.5-0.5)	logBB (-3-1.2)	logKp (-8-1.0)	EA(eV) (-0.9-1.7)	logKhsa (-1.5-1.5)	VRO5 (0-4)
The drug-like properties of the top 10 inhibitors of TcdA protein																
TIP006135	12	-2	480.4	721.1	8	14.5	16.3	34.5	28.9	-2.1	-2.8	-4.8	-7.9	0.6	-1.1	2
TIP006588	27	-2	816.8	1087.0	12	30.5	26.2	57.0	48.7	-3.8	-1.4	-6.8	-7.1	0.6	-2.4	3
TIP007177	15	-2	616.5	917.2	9	16.3	21.9	41.2	33.4	-1.6	-4.3	-7.0	-9.1	0.8	-0.9	3
TIP007178	16	-2	632.5	927.7	10	17.1	22.5	43.7	35.5	-2.3	-4.1	-7.7	-9.9	0.8	-1.1	3
TIP007429	11	-2	464.4	710.5	7	13.8	15.8	32.2	26.8	-1.5	-3.0	-4.2	-6.9	0.7	-0.9	2
TIP009269	10	-2	478.4	661.0	6	13.1	15.2	29.8	24.9	-0.9	-2.7	-4.0	-7.6	0.9	-0.9	2
TIP009928	24	-2	756.7	1065.3	11	27.1	25.2	53.0	44.5	-2.9	-2.3	-6.7	-7.1	1.0	-2.0	3
TIP009947	11	-2	466.4	688.5	7	14.0	15.3	31.8	26.7	-1.5	-2.7	-4.0	-7.0	0.4	-0.9	2
TIP010278	16	-2	610.5	864.6	9	21.5	20.2	42.9	37.1	-2.9	-2.6	-5.6	-7.8	0.7	-1.5	3
TIP012332	11	-2	494.4	670.1	7	13.8	15.7	32.1	26.8	-1.3	-2.5	-4.3	-8.2	0.8	-1.0	2
The drug-like properties of the top 10 inhibitors of TcdB protein																
TIP003016	13	-2	581.0	790.2	8	18.9	18.0	39.4	32.9	-2.0	-2.7	-4.2	-6.8	0.8	-1.2	3
TIP005703	4	0	137.0	354.8	3	1.8	5.9	9.6	7.7	0.3	0.2	-0.2	-4.9	-0.3	-0.6	0
TIP006359	21	-2	699.0	905.8	7	16.6	21.8	40.7	28.4	1.3	-3.3	-4.6	-5.8	0.3	-0.4	3
TIP006360	24	-2	731.0	1078.1	9	19.0	26.0	47.6	34.7	0.1	-4.5	-7.3	-7.3	0.3	-0.9	3
TIP006585	26	-2	787.0	915.7	12	28.8	22.7	54.1	45.9	-3.4	-0.3	-5.2	-6.6	0.1	-1.9	3
TIP006587	26	-2	787.0	1062.7	12	28.8	25.8	56.2	47.3	-3.4	-1.6	-6.7	-7.1	0.5	-2.2	3
TIP007931	10	-2	492.0	753.2	5	13.1	16.1	29.9	23.6	-0.2	-3.9	-4.5	-7.4	0.8	-0.8	2
TIP009928	24	-2	757.0	1067.4	11	27.1	25.3	54.2	44.4	-2.6	-2.4	-6.4	-6.8	0.6	-1.9	3
TIP010071	21	-2	773.0	1017.8	12	29.1	24.7	55.6	48.6	-4.9	-1.9	-7.4	-9.2	1.2	-2.3	3
TIP011613	26	-2	787.0	1077.1	12	28.8	26.1	55.8	47.5	-3.8	-1.8	-7.4	-8.0	0.4	-2.3	3

Rotor, the number of rotatable bonds; CNS, central nervous system activity; MW, molecular weight; SASA, solvent accessible surface area; dHB, donor hydrogen bonds; aHB, acceptor hydrogen bonds; $logP_{C16}$, hexadecane/gas partition coefficient; $logP_{oct}$, octanol/gas partition coefficient; $logP_{w}$, water/gas partition coefficient; $logP_{o/w}$, octanol/water partition coefficient; logS, aqueous solubility; logBB, brain/blood partition coefficient; $logK_{p}$, skin permeability; $logK_{hsa}$, binding affinity to human serum albumin; V_{RO5}, number of violations of Lipinski's rule of five.

each compound was within the optimal range. Most importantly, all candidates were in violation of less than three Lipinski rules (RO5), which is in the acceptable range, since a maximum of four violations is allowed for a potential drug candidate.

Discussion

This study represents the computational analysis and molecular docking of TcdA and TcdB toxins of *C. difficile*, which has become an emerging infectious agent with antibiotic resistance. Through molecular docking, we analyzed 19,171 conformational states of 8,858 phytochemicals and identified 10 potential drug candidates to inhibit each toxin, separately. More importantly, one candidate (TIP009928) was common for both toxins.

Previously, *C. difficile* was treated with antibiotics and metronidazole. However, metronidazole has issues regarding the pharmacokinetics in intestinal infections, which made vancomycin the preferred choice for treatment. However, frequent use of vancomycin may lead to resistance. Therefore, Fidaxomicin becomes another choice [12]. However, not any of these treatments are without potential side effects [3]. In this study, therefore, a whole library of potential plant derivatives from a Taiwanese (TIPdb) database was investigated to specifically target *C. difficile* toxins along with the positive controls to provide a futuristic therapeutic candidate via a two-fold docking maneuver. Interestingly, most of the selected phytochemicals for toxin inhibition were polyphenols, specifically, quercetin and myricetin derivatives. They showed promising binding affinity towards their targets through various non-bonded interactions. Among the identified inhibitors, for instance, compound TIP007178, that showed the highest affinity for TcdA, is myricetin 7-(6”-galloylglucoside). Another myricetin-based inhibitor was TIP006135 (Myricetin 7-glucoside). Furthermore, three TcdA inhibitors (TIP007177, TIP007429, and TIP009269) and two TcdB inhibitors (TIP010071 and TIP003016) were quercetin derivatives. Two hemiterpene glucosides, hymenoside G (TIP006360) and hymenoside F (TIP006359), were also identified as potential TcdB inhibitors, and these inhibitors showed similar interactions to those of the native molecules.

The compound TIP009928 that can bind to and inactivate both TcdA and TcdB toxins is a phenylpropanoid glycoside known as Lavandulifolioside. Antibacterial activities of glycosides are well-reported [29]. It has been demonstrated that myricetin and quercetin possess antimicrobial activities against human pathogenic microorganisms due to their inhibitory effects on the cytoplasmic membrane, nucleic acid synthesis, and energy metabolism [30]. In some cases, the antibacterial effect of flavonoid glycosides is equal to or even higher than that of ciprofloxacin. Hence, they might be an effective replacement for antibiotics [29]. Other identified phytochemicals were mostly different types of glycosides, e.g., lavandulifolioside, lamiuside, plantagoside, hymenoside, and purpureaside.

Furthermore, we estimated the binding free energy to validate the docking results. The higher ΔG values of the docked complexes suggest that the identified phytochemicals are consistent and efficient in binding with their targets. For instance, inhibitor TIP009928 showed a strong binding affinity for TcdB and also provided the lowest binding free energies for TcdA (-100.55) and TcdB (-59.88) proteins, which were lower than their respective controls. Further, we evaluated these compounds for drug-ability and they showed optimal drug-like properties with some violations in Lipinski’s rule of five, though the number of violations lies within the acceptable limit. Therefore, these phytochemicals can be useful in developing therapeutic drugs to target the *C. difficile* toxins.

Conclusion

In a nutshell, we employed a virtual screening protocol to identify potential drug leads from phytochemical resources that can inhibit the catalytic domain of *C. difficile* toxins. The analysis was done based on the binding affinities of the compounds for TcdA and TcdB as compared to their known ligands and inhibitors. Interestingly, we found 19 potential phytochemicals that showed both higher binding affinities towards their respective targets and suitable drug-like properties. Therefore, we believe that these potential plant-based inhibitors can be used in a prospective drug development program for the treatment of *C. difficile*-

mediated colitis. Nevertheless, these findings are solely based on chemoinformatics. Hence, they are required to be validated with proper *in vivo* and *in vitro* experiments.

Acknowledgements

We extend our deepest gratitude towards Zulkar Nain (Research Assistant, Department of Biotechnology and Genetic Engineering, Islamic University, Bangladesh) for his help in the experimental procedure and manuscript revision. We also would like to thank Dr. Mohanasundaram (Department of Microbiology, Alagappa University, India) for his kind help in reviewing the first draft of this paper. The authors received no funding or external grants for this study.

Disclosure of conflict of interest

None.

Address correspondence to: Khalid M Aljarallah, College of Applied Medical Sciences, Majmaah University, University Campus, Majmaah, Riyadh Province, Saudi Arabia. Tel: +966-555223191; E-mail: k.aljarallah@mu.edu.sa

References

[1] Sandiford SK. Current developments in lantibiotic discovery for treating Clostridium difficile infection. Expert Opin Drug Discov 2019; 14: 71-79.

[2] Abt MC, McKenney PT and Pamer EG. Clostridium difficile colitis: pathogenesis and host defence. Nat Rev Microbiol 2016; 14: 609-620.

[3] Khorasani S, Dossa F, McKechnie T, Englesakis M, Brar MS and de Buck van Overstraeten A. Association between preoperative oral antibiotics and the incidence of postoperative clostridium difficile infection in adults undergoing elective colorectal resection: a systematic review and meta-analysis. Dis Colon Rectum 2020; 63: 545-561.

[4] Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, Farley MM, Holzbauer SM, Meek JI, Phipps EC, Wilson LE, Winston LG, Cohen JA, Limbago BM, Fridkin SK, Gerding DN and McDonald LC. Burden of Clostridium difficile infection in the United States. N Engl J Med 2015; 372: 825-834.

[5] Ma GK, Brensinger CM, Wu Q and Lewis JD. Increasing incidence of multiply recurrent Clostridium difficile infection in the United States: a cohort study. Ann Intern Med 2017; 167: 152-158.

[6] Carter GP, Chakravorty A, Pham Nguyen TA, Mileto S, Schreiber F, Li L, Howarth P, Clare S, Cunningham B, Sambol SP, Cheknis A, Figueroa I, Johnson S, Gerding D, Rood JI, Dougan G, Lawley TD and Lyras D. Defining the roles of TcdA and TcdB in localized gastrointestinal disease, systemic organ damage, and the host response during Clostridium difficile infections. mBio 2015; 6: e00551.

[7] Ballo O, Kreisel EM, Eladly F, Brunnberg U, Stratmann J, Hunyady P, Hogardt M, Wichelhaus TA, Kempf VAJ, Steffen B, Vehreschild JJ, Vehreschild MJGT, Finkelmeier F, Serve H and Brandts CH. Use of carbapenems and glycopeptides increases risk for clostridioides difficile infections in acute myeloid leukemia patients undergoing intensive induction chemotherapy. Ann Hematol 2020; 99: 2547-2553.

[8] Saha S, Kapoor S, Tariq R, Schuetz AN, Tosh PK, Pardi DS and Khanna S. Increasing antibiotic resistance in clostridioides difficile: a systematic review and meta-analysis. Anaerobe 2019; 58: 35-46.

[9] Hunt JJ and Ballard JD. Variations in virulence and molecular biology among emerging strains of clostridium difficile. Microbiol Mol Biol Rev 2013; 77: 567-581.

[10] Papatheodorou P, Zamboglou C, Genisyuerek S, Guttenberg G and Aktories K. Clostridial glucosylating toxins enter cells via clathrin-mediated endocytosis. PLoS One 2010; 5: e10673.

[11] Bilverstone TW, Garland M, Cave RJ, Kelly ML, Tholen M, Bouley DM, Kaye P, Minton NP, Bogyo M, Kuehne SA and Melnyk RA. The glucosyltransferase activity of C. difficile Toxin B is required for disease pathogenesis. PLoS Pathogens 2020; 16: e1008852.

[12] Stebel R, Vojtilová L and Husa P. Clostridium difficile infection: an update on treatment and prevention. Vnitr Lek 2020; 66: 58-62.

[13] Janvilisri T, Sorg JA, Scaria J and Sadowsky MJ. Alternative therapeutic approaches for multidrug resistant clostridium difficile. Front Microbiol 2019; 10: 1216.

[14] Cox AD, St Michael F, Aubry A, Strong PCR, Hayes AC and Logan SM. Comparison of polysaccharide glycoconjugates as candidate vaccines to combat clostridiodes (clostridium) difficile. Glycoconj J 2021; 38: 493-508.

[15] Lee Y, Lim WI, Bloom CI, Moore S, Chung E and Marzella N. Bezlotoxumab (Zinplava) for clostridium difficile infection: the first monoclonal antibody approved to prevent the recurrence of a bacterial infection. P T 2017; 42: 735-738.

[16] Aljarallah KM. Inhibition of Clostridium difficile by natural herbal extracts. J Taibah Univ Med Sci 2016; 11: 427-431.

[17] Roshan N, Riley TV, Knight DR, Steer JH and Hammer KA. Natural products show diverse mechanisms of action against Clostridium difficile. J Appl Microbiol 2019; 126: 468-479.
[18] Alvin JW and Lacy DB. Clostridium difficile toxin glucosyltransferase domains in complex with a non-hydrolyzable UDP-glucose analogue. J Struct Biol 2017; 198: 203-209.
[19] Clark DE. What has virtual screening ever done for drug discovery? Expert Opin Drug Discov 2008; 3: 841-851.
[20] Natarajan P, Priyadarshini V, Pradhan D, Manne M, Swargam S, Kanipakam H, Bhuma V and Amineni U. E-pharmacophore-based virtual screening to identify GSK-3β inhibitors. J Recept Signal Transduct Res 2016; 36: 445-458.
[21] Kaminski GA, Friesner RA, Tirado-Rives J and Jorgensen WL. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 2001; 105: 6474-6487.
[22] Tung CW, Lin YC, Chang HS, Wang CC, Chen IS, Jheng JL and Li JH. TIPdb-3D: the three-dimensional structure database of phytochemicals from Taiwan indigenous plants. Database (Oxford) 2014; 2014: bau055.
[23] Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR and Uchimaya M. Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. J Comput Aided Mol Des 2007; 21: 681-691.
[24] Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P and Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 2004; 47: 1739-1749.
[25] Saxena S, Durgam L and Guruprasad L. Multiple e-pharmacophore modelling pooled with high-throughput virtual screening, docking and molecular dynamics simulations to discover potential inhibitors of Plasmodium falciparum lactate dehydrogenase (PfLDH). J Biomol Struct Dyn 2019; 37: 1783-1799.
[26] Lyne PD, Lamb ML and Saeh JC. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. J Med Chem 2006; 49: 4805-4808.
[27] Ioakimidis L, Thoukydidis L, Mirza A, Naeem S and Reynisson J. Benchmarking the reliability of QikProp. Correlation between experimental and predicted values. Mol Inform 2008; 27: 445-456.
[28] Lipinski CA, Lombardo F, Dominy BW and Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 2001; 46: 3-26.
[29] Tagousop CN, Ekom SE, Ngnokam D and Voutquenne-Nazabadioko L. Antimicrobial activities of flavonoid glycosides from Graptophyllum grandulosum and their mechanism of antibacterial action. BMC Complement Altern Med 2018; 18: 252.
[30] Xie Y, Chen J, Xiao A and Liu L. Antibacterial activity of polyphenols: structure-activity relationship and influence of hyperglycemic condition. Molecules 2017; 22: 1913.