Dear Author

Here are the proofs of your article.

- You can submit your corrections online, via e-mail or by fax.
- For online submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and email the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the journal title, article number, and your name when sending your response via e-mail or fax.
- Check the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- Check the questions that may have arisen during copy editing and insert your answers/corrections.
- Check that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the Edited manuscript.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please do not make changes that involve only matters of style. We have generally introduced forms that follow the journal’s style.
- Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections within 48 hours, we will send you a reminder.
- Your article will be published Online First approximately one week after receipt of your corrected proofs. This is the official first publication citable with the DOI. Further changes are, therefore, not possible.
- The printed version will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL:

http://dx.doi.org/10.1007/s00894-019-3979-6

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information, go to:


Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us, if you would like to have these documents returned.
Metadata of the article that will be visualized in OnlineFirst
Graphical abstract

Structure-based identification of potent VEGFR-2 inhibitors from in vivo metabolites of a herbal ingredient
Abstract

Vascular endothelial growth factor receptor-2 (VEGFR-2) is one of the regulatory elements of angiogenesis that is expressed highly in various diseases and is also essential for solid tumor growth. The present study was aimed at identifying potent inhibitors of VEGFR-2 by considering herbal secondary metabolites; as natural molecules, the latter are less toxic than synthetic derivatives. A structure-based virtual screening protocol consisting of molecular docking, MM-GBSA and ADME/T analysis was initially used to screen a library of in vivo metabolites of the herbal ingredient. Using a fixed cutoff value, four potent virtual hits were identified from molecular docking, ADME/T and binding affinity calculations, which were considered further for molecular dynamics (MD) simulation to broadly describe the binding mechanism to VEGFR-2. The results suggested that these molecules have high affinity for the catalytic region of VEGFR-2, and form strong hydrophobic and polar interactions with the amino acids involved in the binding of ATP and linker regions of the catalytic site. Subsequently, the stability of the docked complexes and binding mechanisms were evaluated by MD simulations, and the energy of binding was calculated through MM-PBSA analysis. The results uncovered two virtual hits, designated ZINC14762520 and ZINC36470466, as VEGFR-2 inhibitors, and suggested that they bind in
an ATP-competitive manner. These virtual hits will offer a suitable starting point for the further design of their various analogs, allowing a rational search for more effective inhibitors in the future.

Keywords separated by ' - '

VEGFR-2 - Angiogenesis - Molecular docking - Molecular dynamics - MM-PBSA

Foot note information

The online version of this article (https://doi.org/10.1007/s00894-019-3979-6) contains supplementary material, which is available to authorized users.

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Electronic supplementary material

ESM 1

(DOC 6881 kb)

ESM 2

(XLS 62 kb)
Structure-based identification of potent VEGFR-2 inhibitors from in vivo metabolites of a herbal ingredient

Raju Dash1,2,3 · Md. Junaid1 · Sarmistha Mitra4 · Md Arifuzzaman3,5 · S. M. Zahid Hosen1

Received: 30 October 2017 / Accepted: 5 March 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract
Vascular endothelial growth factor receptor-2 (VEGFR-2) is one of the regulatory elements of angiogenesis that is expressed highly in various diseases and is also essential for solid tumor growth. The present study was aimed at identifying potent inhibitors of VEGFR-2 by considering herbal secondary metabolites; as natural molecules, the latter are less toxic than synthetic derivatives. A structure-based virtual screening protocol consisting of molecular docking, MM-GBSA and ADME/T analysis was initially used to screen a library of in vivo metabolites of the herbal ingredient. Using a fixed cutoff value, four potent virtual hits were identified from molecular docking, ADME/T and binding affinity calculations, which were considered further for molecular dynamics (MD) simulation to broadly describe the binding mechanism to VEGFR-2. The results suggested that these molecules have high affinity for the catalytic region of VEGFR-2, and form strong hydrophobic and polar interactions with the amino acids involved in the binding of ATP and linker regions of the catalytic site. Subsequently, the stability of the docked complexes and binding mechanisms were evaluated by MD simulations, and the energy of binding was calculated through MM-PBSA analysis. The results uncovered two virtual hits, designated ZINC14762520 and ZINC36470466, as VEGFR-2 inhibitors, and suggested that they bind in an ATP-competitive manner. These virtual hits will offer a suitable starting point for the further design of their various analogs, allowing a rational search for more effective inhibitors in the future.

Keywords VEGFR-2 · Angiogenesis · Molecular docking · Molecular dynamics · MM-PBSA

Abbreviations
ADME/T Absorption distribution metabolism excretion and toxicity
AMBER Assisted model building with energy refinement
HTVS High through put virtual screening
KDR Kinase insert domain receptor

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00894-019-3979-6) contains supplementary material, which is available to authorized users.

S. M. Zahid Hosen
smzahidhosen@bcsir.gov.bd

Raju Dash
rajudash.bgctub@gmail.com

Md. Junaid
md.junaid@northsouth.edu

Sarmistha Mitra
samisthacu@gmail.com

Md Arifuzzaman
larif67@yahoo.com

1 Molecular Modeling and Drug Design Laboratory, Pharmacology Research Division, Bangladesh Council of Scientific and Industrial Research (BCSIR), Chittagong 4220, Bangladesh
2 Department of Anatomy, Dongguk University Graduate School of Medicine, Gyeongju 38066, Republic of Korea
3 Department of Biochemistry and Biotechnology, University of Science and Technology Chittagong, Chittagong 4202, Bangladesh
4 Plasma Bioscience Research Center, Plasma-bio display, Kwangwoon University, Seoul 01897, Republic of Korea
5 Department of Natural Sciences, LAGCC, City University of New York (CUNY), New York, NY, USA
Introduction

The process by which new blood vessel is formed—dubbed angiogenesis—has recently gained significant attention in anticancer drug development. The process of angiogenesis is pivotal in solid tumor growth, diabetes and other age-related diseases [1, 2]. In solid tumor, blood vessels are formed excessively to confer sufficient oxygen and nutrient supply to malignant cells in order to establish the process of metastasis [3]. Evidence has also shown that vascular endothelial growth factor (VEGF) and its receptors (VEGFRs) are the main modulators of angiogenesis [4, 5].

The members of VEGFR including VEGFRs 1–3 are targeted by VEGF, which binds to the extracellular domain of VEGFRs to activate downstream signaling pathways [3]. VEGFR-2 (KDR) directly regulates endothelial cell proliferation and migration [6–8]. Upon binding of VEGF-A, the tyrosine residues at the carboxy-terminal of the receptor are autophosphorylated, thus activating cell signaling, vasculogenesis, and angiogenesis [9, 10]. Aggressive expression of VEGFR-2 activates downstream pathways of cell survival, which makes this an attractive target for developing novel anticancer agents.

A number of VEGFR inhibitors are currently in phase I or phase II clinical trials [11], while other inhibitors, such as sorafenib (Nexavar®, BAY 43–9006), sunitinib malate (Sutent®, SU11248), and bevacizumab (Avastin®), are already approved for cancer treatment by the Food and Drug Administration (FDA) [12]. Pharmacologically, they act by blocking the VEGFR pathway; however, their prolonged use can lead to serious side-effects, such as bleeding, gastrointestinal perforation and hypertension [13, 14]. Therefore, the search for novel and potent VEGFR inhibitors that could effectively block angiogenesis remains a major challenge in anti-angiogenic drug development.

Due to their high efficacy and lower toxicity, compounds from natural sources have recently gained more attention as anti-angiogenesis agents [15]. Nevertheless, they still suffer from poor bioavailability [16]. Consequently, in vivo metabolites of some herbal compounds were found to show better pharmacokinetic, potency and safety profiles in comparison to their parent compounds [17]. For this reason, to find potent VEGFR-2 inhibitors with optimum pharmacokinetic profiles, this study applied a virtual in silico screening approach using a database of in vivo metabolites of herbal ingredients. The virtual screening procedure was integrated with a three-step filtering approach, including Glide XP docking, ADME/T and MM-PBSA analysis, to sort out the top hits from the database (Fig. S1, see Supplementary file 1). After that, each docked protein–ligand complex was subjected to 100-ns MD simulations along with binding energy calculations through MM-PBSA analysis, in order to elucidate the VEGFR-2 binding mechanisms of identified compounds. The study identified two virtual hits, ZINC14762520 and ZINC36470466, that could be potent VEGFR-2 inhibitors, and also suggested a basis for further rational design of their various analogs to find more effective inhibitors in the future.

Materials and methods

Database processing

A dataset of in vivo secondary metabolites of herbal ingredients was retrieved from the ZINC server named Herbal Ingredients In-Vivo Metabolism (HIM) (last updated: 12 August 2013), giving a set of 663 in-vivo metabolites of herbal active ingredients. Although the ZINC database consists of molecules ready to dock in three dimensional (3D) formats, to make it more rationale, we further prepared the dataset using Schrödinger Suite 2013 [18, 19] software. Here, all ligands were minimized in Ligprep2.5 wizard [20], by implementing OPLS 2005 [21] force field. During the minimization, the module Epik 2.2 [22] was utilized to fix the ionization state of the ligand at pH 7.0 ± 2.0 [23]. From this analysis, up to 32 possible stereoisomers of each compound were generated; we selected only the conformer having the lowest energy for further processing.

Protein preparation and grid generation

Protein structures for virtual screening were prepared using the protein preparation module of Schrödinger Suite 2013 [19] software, in which the crystal structure is initially assigned proper hydrogen, charges and bond orders. As the main target of the virtual screening was VEGFR-2, its tyrosine kinase domain was retrieved from the RCSB protein data bank (PDB id: 3B8R) [24]. At neutral pH, all hydrogen bonds in the

<table>
<thead>
<tr>
<th>PME</th>
<th>Particle mesh Ewald</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rg</td>
<td>Radius of gyration</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean square deviation</td>
</tr>
<tr>
<td>RMSF</td>
<td>Root mean square fluctuation</td>
</tr>
<tr>
<td>SAR</td>
<td>Structure activity relationship</td>
</tr>
<tr>
<td>SASA</td>
<td>Solvent accessible surface area</td>
</tr>
<tr>
<td>SGB</td>
<td>Surface generalized Born</td>
</tr>
<tr>
<td>SP</td>
<td>Standard precision</td>
</tr>
<tr>
<td>TIP3P</td>
<td>The transferable intermolecular potential3 points</td>
</tr>
<tr>
<td>VEGFR-2</td>
<td>Vascular endothelial growth</td>
</tr>
<tr>
<td>2</td>
<td>factor receptor 2</td>
</tr>
<tr>
<td>XP</td>
<td>Extra precision</td>
</tr>
<tr>
<td>YASARA</td>
<td>Yet another scientific artificial reality application</td>
</tr>
</tbody>
</table>

Due to their high efficacy and lower toxicity, compounds from natural sources have recently gained more attention as anti-angiogenesis agents [15]. Nevertheless, they still suffer from poor bioavailability [16]. Consequently, in vivo metabolites of some herbal compounds were found to show better pharmacokinetic, potency and safety profiles in comparison to their parent compounds [17]. For this reason, to find potent VEGFR-2 inhibitors with optimum pharmacokinetic profiles, this study applied a virtual in silico screening approach using a database of in vivo metabolites of herbal ingredients. The virtual screening procedure was integrated with a three-step filtering approach, including Glide XP docking, ADME/T and MM-PBSA analysis, to sort out the top hits from the database (Fig. S1, see Supplementary file 1). After that, each docked protein–ligand complex was subjected to 100-ns MD simulations along with binding energy calculations through MM-PBSA analysis, in order to elucidate the VEGFR-2 binding mechanisms of identified compounds. The study identified two virtual hits, ZINC14762520 and ZINC36470466, that could be potent VEGFR-2 inhibitors, and also suggested a basis for further rational design of their various analogs to find more effective inhibitors in the future.
structure were optimized, deleting unnecessary water. Afterwards, the minimization process was run with the OPLS 2005 force field, considering structural changes of not more than 0.30 Å of RMSD. The active site of the protein was fixed for docking simulation by generating a grid box at the reference ligand binding the protein. Grid generation parameters were kept at default, with a box size of 18 Å × 18 Å × 18 Å, and the OPLS 2005 force field utilized for post minimization. The charge cutoff and van der Waals scaling factor were set to 0.25 and 1.00, respectively [25].

**Virtual screening**

We carried out extra precision (XP) flexible docking using the Glide module of Schrödinger-Maestro v9.4 [26, 27], which is more sophisticated than SP/HTVS in scoring function [27]. Here, all ligands were treated flexibly, considering the partial charge and van der Waals factor of 0.15 and 0.80, respectively. Minimization was performed to the docked complex after docking using the OPLS 2005 force field. The best-docked pose with lowest Glide score value was recorded for each ligand [25].

**ADME/T analysis**

In order to predict the pharmacokinetic profile of the identified hits, QikProp (Schrodinger Suite 2013) [28] was utilized, where the absorption, distribution and metabolism of each ligand were calculated. Pharmacokinetics and the drug-like properties of the ligands, such as the rate of oral absorption, QlogP, solvent accessible surface area (SASA), and Lipinski's rule of five [29], were predicted, and maximum percentage of oral absorption rate with a cutoff of 50% was considered for selecting the top hits for further processing.

**Prime MM-GBSA**

Binding affinity calculations were carried out using the Prime MM-GBSA module of the Schrodinger suite [30], where greater negative affinity denotes higher stability. The docked Pose viewer file from Glide XP docking was introduced for calculation, using the sampling minimization protocol used. Generalized Born Surface Accessible (GBSA) as a continuum model, and OPLS 2005 force field as molecular mechanics (MM), keeping the protein flexible [31–35]. For correcting empirical functions of π-stacking and H-bond interactions, a dielectric solvent model such as VSGB 2.0 [36] was used.

**MD simulations**

MD simulations were performed using YASARA Dynamics software [37], and the AMBER14 force field [38]. A total of six different systems were used to run MD simulation. The systems included top four virtual hits, (ZINC13370550-VEGFR-2, ZINC02585769-VEGFR-2, ZINC14762520-VEGFR-2 and ZINC36470466-VEGFR-2 complexes), sunitinib-VEGFR-2, and apo-VEGFR-2. In each protein–ligand complex, the parameters of each ligand were assigned through AutoSMILES [39] algorithms, where unknown organic molecules are parameterized fully automatically by the calculation of semi-empirical AM1 Mulliken point charges [39] with the COSMO solvation model, assigning of AM1BCC [40] atom and bond types, and also assigning GAff [41] (General AMBER Force Field) atom types and the remaining force field parameters. Before simulation, the hydrogen bonding network of the protein–ligand complex was optimized and solvated by a TIP3P [42] water model in a simulation cell. Periodic boundary conditions were maintained with a solvent density of 0.997 gL⁻¹. Titratable amino acids in the protein complex were subjected to pH calculation during solvation. The initial energy minimization process of each simulation system, consisting of 62,521 ± 10 atoms was performed by a simulated annealing method, using the steepest gradient approach (5000 cycles). Each simulation was run with a multiple time step algorithm [43], using a time-step interval of 2.50 fs under physiological conditions (298 K, pH 7.4, 0.9% NaCl) [44]. All bond lengths were constrained using the linear constraint solver (LINCS) [45] algorithm, and SETTLE [46] was used for water molecules. Long-range electrostatic interactions were described by the PME [47] methods, and, finally, 100 ns MD simulation was accomplished at Berendsen thermostat [48] and constant pressure. For further analysis, the trajectories were saved every 250 ps, and subsequent analysis was performed by default script of YASARA [49] macro and VMD [50, 51] software.

All snapshots were further subjected to MM-PBSA (MM-Poisson–Boltzmann surface area) binding free energy calculation by YASARA software using the following formula [52].

\[
\text{Binding Energy} = E_{\text{potRecept}} + E_{\text{solvRecept}} + E_{\text{potLigand}} + E_{\text{solvLigand}} - E_{\text{potComplex}} - E_{\text{solvComplex}}
\]

Here, YASARA built-in macros was used to calculate MM-PBSA binding energy, using AMBER 14 as a force field, where more positive energies indicate better binding [53].

**Results and discussion**

**Virtual screening**

Overexpression of VEGF and its receptor VEGFR-2 is the most common phenomenon in angiogenesis and plays a significant role in cell survival and promoting tumorigenesis.
In current drug discovery processes, virtual screening is the approach most widely used to find novel hits for a target by reducing time, cost and the massive compound library. Apparently, XP docking assists in filtering out false positives from the compound database by using a strict scoring function; hence, it was considered in this study [57]. However, prior to docking simulations, we first validated docking accuracy through redocking of a known reference ligand using XP docking methodology. Here, the reference co-crystal ligand was first separated and re-docked using the Glide XP algorithm. Conformation of the complex with the highest negative docking score was selected from each run. The RMSD of the docked conformation was then calculated, where a lower RMSD value denotes a high level of the accuracy of docking protocol. Interestingly, Glide docking with XP setting produced the docked complex with lowest RMSD of 0.61 Å, which denoted that the position of docked ligand was almost similar to that of the crystal, as shown in Fig. S2 (see supplementary file 1). Therefore, the candidate ligands including reference inhibitors, sunitinib and sorafenib were docked into VEGFR-2, following the same docking protocol.

Initially, the top 5% compounds were selected from Glide XP docking based on high docking score. As a result, 35 compounds were considered, with docking score range from −12.00 to −10.00 kcal mol\(^{-1}\), in which the highest docking score was found to be −11.80 kcal mol\(^{-1}\) for ZINC95099443 and the lowest docking score was −10.04 kcal mol\(^{-1}\) for ZINC95099251. On the contrary, the reference inhibitors, sunitinib and sorafenib resulted in docking scores of −6.60 kcal mol\(^{-1}\), and −7.65 kcal mol\(^{-1}\), respectively (Table 1). However, bioavailability and toxicity are the primary concerns for drug development; therefore, we subjected these compounds for ADMET profiling, as discussed below.

### Bioavailability of ligands

Pharmacologically relevant attributes and physiochemical descriptors of the identified hypothetical hits were calculated by QikProp software. Each descriptor’s range satisfies 95% of known drugs [58–60]. The pharmacokinetic properties of 35 ligands were assessed (see supplementary file 2), where cell permeability, the number of rotatable bonds, number of metabolites, solubility, logP, and oral absorption rate were predicted, from which the top six candidates were selected (Fig. S3, see supplementary file 1).

As described in Table 2, all calculated attributes were within the satisfactory range, with properties including, QPlogPo/w (octanol/water partition coefficient), QPlogHERG (log IC\(_{50}\) value for blockage of K\(^+\) channels) [61, 62], SASA [59], QPlogS (aqueous solubility) [59], molecular weight, and H bond acceptor/donor [61]. Although the standard inhibitors obtained highest oral absorption, they were also predicted to fall outside the recommended range for QPlogHERG value. The HERG K\(^+\) channel, which is encoded by the human ether-a-go-go related gene (HERG), is one of the main reasons for the cardiac toxicity produced by a wide range of therapeutic drugs [63]. This channel promotes electrical activity of the heart and also modulates various functions of the nervous system [64]; thus, blockers of this channel cause potential toxicity, which is a major consideration in drug discovery processes [65]. All these findings indicating that the ligands identified showing ideal behavior of drug molecules, and that they are well tolerated and less toxic than standard inhibitors.

### MM-GBSA binding affinity estimation

The top six identified ligands were re-ranked by their binding affinity calculated through MM-GBSA analysis. ZINC14762520 obtained highest binding affinity (Δ\(G\)\(_{\text{bind}}\) = −78.09 kcal mol\(^{-1}\)) with highest contribution from covalent and lipophilic binding energies. In contrast, ZINC95099475 showed the lowest binding affinity of −58.29 kcal mol\(^{-1}\) (Table S1, see supplementary file 1). In comparison to the identified ligands, the reference inhibitors sunitinib and sorafenib showed binding affinities of −45.26 kcal mol\(^{-1}\) and −57.44 kcal mol\(^{-1}\) respectively. On the basis of MM-GBSA result, we thus selected the top four compounds and subjected them to detailed binding analysis.

<table>
<thead>
<tr>
<th>Compound name</th>
<th>Docking score (kcal mol(^{-1}))</th>
<th>Glide energy (kcal mol(^{-1}))</th>
<th>Glide ligand efficiency</th>
<th>Strain penalty</th>
<th>Glide Emodel (kcal mol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunitinib</td>
<td>−6.60</td>
<td>−51.69</td>
<td>−0.22</td>
<td>0</td>
<td>−72.91</td>
</tr>
<tr>
<td>Sorafenib</td>
<td>−7.65</td>
<td>−53.60</td>
<td>−0.23</td>
<td>0</td>
<td>−88.69</td>
</tr>
<tr>
<td>ZINC13370550</td>
<td>−10.68</td>
<td>−55.73</td>
<td>−0.35</td>
<td>0</td>
<td>−88.02</td>
</tr>
<tr>
<td>ZINC02585769</td>
<td>−10.64</td>
<td>−44.22</td>
<td>−0.50</td>
<td>0</td>
<td>−70.24</td>
</tr>
<tr>
<td>ZINC95099475</td>
<td>−10.19</td>
<td>−42.89</td>
<td>−0.48</td>
<td>0</td>
<td>−65.61</td>
</tr>
<tr>
<td>ZINC36470466</td>
<td>−11.06</td>
<td>−48.05</td>
<td>−0.48</td>
<td>0</td>
<td>−66.98</td>
</tr>
<tr>
<td>ZINC38957372</td>
<td>−11.43</td>
<td>−43.63</td>
<td>−0.54</td>
<td>0</td>
<td>−69.10</td>
</tr>
<tr>
<td>ZINC14762520</td>
<td>−10.64</td>
<td>−34.85</td>
<td>−0.39</td>
<td>0</td>
<td>−70.95</td>
</tr>
</tbody>
</table>
Table 2 Pharmacokinetic and toxicological properties of the top six candidates and standard inhibitors, calculated from QikProp

<table>
<thead>
<tr>
<th>Name</th>
<th>MW (g/mol)</th>
<th>HB donor</th>
<th>HB acceptor</th>
<th>SASA (Å²)</th>
<th>log HERG</th>
<th>Qlog S</th>
<th>log Po/w</th>
<th>% Human oral absorption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunitinib</td>
<td>398.47</td>
<td>2.00</td>
<td>6.00</td>
<td>751.53</td>
<td>-6.68</td>
<td>-5.32</td>
<td>3.84</td>
<td>89.98</td>
</tr>
<tr>
<td>Sorafenib</td>
<td>464.83</td>
<td>3.00</td>
<td>6.00</td>
<td>748.03</td>
<td>-5.64</td>
<td>-6.97</td>
<td>3.94</td>
<td>96.88</td>
</tr>
<tr>
<td>ZINC13370550</td>
<td>416.38</td>
<td>4</td>
<td>10.55</td>
<td>660.94</td>
<td>-5.006</td>
<td>-3.66</td>
<td>0.59</td>
<td>59.78</td>
</tr>
<tr>
<td>ZINC02585769</td>
<td>286.24</td>
<td>4</td>
<td>5.5</td>
<td>496.06</td>
<td>-4.097</td>
<td>-2.53</td>
<td>0.44</td>
<td>60.22</td>
</tr>
<tr>
<td>ZINC09509475</td>
<td>288.25</td>
<td>3</td>
<td>4.75</td>
<td>494.56</td>
<td>-4.676</td>
<td>-2.90</td>
<td>0.95</td>
<td>65.55</td>
</tr>
<tr>
<td>ZINC36470466</td>
<td>318.36</td>
<td>5</td>
<td>5.65</td>
<td>578.41</td>
<td>-5.003</td>
<td>-2.44</td>
<td>1.37</td>
<td>65.62</td>
</tr>
<tr>
<td>ZINC38957372</td>
<td>290.27</td>
<td>4</td>
<td>5.5</td>
<td>523.10</td>
<td>-5.047</td>
<td>-2.59</td>
<td>0.90</td>
<td>66.99</td>
</tr>
<tr>
<td>ZINC14762520</td>
<td>372.41</td>
<td>3</td>
<td>4.75</td>
<td>661.36</td>
<td>-5.000</td>
<td>-4.34</td>
<td>3.2</td>
<td>84.28</td>
</tr>
</tbody>
</table>

a Molecular weight (acceptable range: <500)
b Hydrogen bond donor (acceptable range: ≤5)
c Hydrogen bond acceptor (acceptable range: ≤10)
d Total solvent accessible surface area in using a probe with a 1.4 radius (acceptable range: 300–1000)
e Predicted IC50 value for blockage of HERG K+ channels (concern: below -5)
f Predicted aqueous solubility, S in mol dm-3 (acceptable range: -6.5 to 0.5)
g Predicted octanol/water partition coefficient (acceptable range: -2 to 6.5)
h Predicted human oral absorption on 0 to 100% scale (<25% is poor and >80% is high)

Ligand binding analysis

The docked complexes of the top four virtual hits from Glide XP docking were further considered for protein–ligand interaction profiling. As shown in Table 3, and Fig. 1a, ZINC02585769 showed four hydrogen bonds with C919, D1046, E885 and G1047 residues, while C919 formed double hydrophobic interactions with the ligand, maintaining corresponding binding distances of 2.23 Å and 2.41 Å, respectively. Although the ligand formed major hydrophobic interactions with K868, V899, V848, L1035 and L840 residues and were thus involved in hydrophobic interactions via π-alkyl bonding.

In the case of the ZINC13770550 (Fig. 1b) formed several hydrogen bonds with the active site residues, where most of the interactions were mediated by E868, C919 and N923 residues. C919 was also involved in π-donor hydrogen bonding with the ligand, while L840, L1035, V848, A866 and C1045 were responsible for hydrophobic interactions via π-alkyl bonding.

According to the detailed analysis of molecular interactions, except ZINC14762520, all compounds formed hydrophobic interactions with L840 and L848 residues of the glycine rich loop (G-loop) that makes a hydrophobic interaction with ATP [66, 67], while ZINC14762520 showed interactions with several residues from the hinge region and DFG motif, including C919, D1046 and F1047. All ligands also made contacts with the residues of the linker region, including C1045 and D1046, which confirm their VEGFR-2 inhibition, as these residues are directly involved in polar interactions with ATP, assist the shifting of γ-phosphate through coordinating Mg2+ ions [68]. Furthermore, non-covalent interactions of the residues of the hinge region (T917 to N923) with all four ligands revealed that these compounds may act as ATP competitive inhibitors [69, 70].

MD simulations

To describe the binding mechanisms, flexibilities and structural behaviors of the top virtual hits, we conducted MD simulations for 100 ns for each complex, including the reference inhibitor, sunitinib and also the ligand free protein. Different energies such as coloumb, van der Waals and total energies in all complexes were calculated from the simulation trajectories and are plotted in Fig. 2. As shown in Fig. 2, the highest RMSD values were observed for ZINC36470466, while ZINC02585769 obtained the lowest RMSDs in the 100 ns simulation. As can be seen from the plot, ZINC02585769 showed a more rigid conformation than the other proteins, and also achieved equilibrium at 10 ns and remained stable afterward.
Table 3: Detailed molecular interactions analysis for best hits obtained from Glide XP docking

<table>
<thead>
<tr>
<th>Compound</th>
<th>Conventional hydrogen bonds</th>
<th>π-Donor hydrogen bond</th>
<th>Hydrophobic</th>
<th>π-Alkyl</th>
<th>Amide-π-stacked</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZINC02585769</td>
<td>Cα19 (2.23 Å), Cα19 (2.41 Å), Cα19 (1.88 Å), Cα19 (2.15 Å)</td>
<td>Kα868 (5.15 Å), Kα868 (4.18 Å), Kβ20 (4.88 Å), Kβ20 (4.73 Å), Kβ20 (3.76 Å), L1015 (4.46 Å), L1015 (4.46 Å), L1015 (5.02 Å), L1015 (5.31 Å), L1015 (4.94 Å), L1015 (4.11 Å)</td>
<td>L1035 (4.36 Å), L1035 (5.12 Å), L1035 (5.33 Å), L1035 (4.92 Å), L1035 (4.16 Å), L1035 (4.11 Å), L1035 (4.52 Å), L1035 (4.65 Å), L1035 (5.29 Å), L1035 (4.60 Å)</td>
<td>C919 (1.98 Å), C919 (2.09 Å), C919 (2.12 Å), C919 (1.98 Å), C919 (1.96 Å)</td>
<td>C919 (1.98 Å), C919 (2.09 Å), C919 (2.12 Å), C919 (1.98 Å), C919 (1.96 Å)</td>
</tr>
<tr>
<td>ZINC13370550</td>
<td>E1868 (2.83 Å), E1868 (2.90 Å), E1868 (2.05 Å), E1868 (2.26 Å), E1868 (2.52 Å), E1868 (2.23 Å), E1868 (2.62 Å), E1868 (1.83 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å), E1868 (1.81 Å)</td>
<td>Kα868 (4.60 Å), Kα868 (5.35 Å), Kα868 (4.36 Å), Kα868 (5.29 Å), Kα868 (4.60 Å), Kα868 (5.12 Å), Kα868 (5.33 Å), Kα868 (4.92 Å), Kα868 (4.16 Å), Kα868 (4.11 Å), Kα868 (4.52 Å), Kα868 (4.65 Å), Kα868 (5.29 Å), Kα868 (4.60 Å)</td>
<td>L1035 (4.36 Å), L1035 (5.12 Å), L1035 (5.33 Å), L1035 (4.92 Å), L1035 (4.16 Å), L1035 (4.11 Å), L1035 (4.52 Å), L1035 (4.65 Å), L1035 (5.29 Å), L1035 (4.60 Å)</td>
<td>C919 (1.98 Å), C919 (2.09 Å), C919 (2.12 Å), C919 (1.98 Å), C919 (1.96 Å)</td>
<td>C919 (1.98 Å), C919 (2.09 Å), C919 (2.12 Å), C919 (1.98 Å), C919 (1.96 Å)</td>
</tr>
</tbody>
</table>

In contrast, ZINC36470466 showed a dramatic increase in flexibility, with RMSD values rising gradually from 1.5 Å to 2.75 Å over time. ZINC14762520 was in equilibrium up to 70 ns, but showed fluctuations thereafter. On the other hand, the apo form also reached equilibrium after 10 ns and maintained RMSD up to 1.9 Å for 50 ns, after which the RMSD was seen to increase gradually. Similarly, sunitinib displayed stable RMSD up to 50 ns; however, after that it decreased to 1.5 Å and remained steady thereafter. Overall, the results of RMSD analysis clearly depicted that the conformation of the protein was changed upon ligand binding, as the apo structure represented the lowest RMSD fluctuations over the course of the simulation. In order to investigate local fluctuations in protein structure induced by the ligands, the root mean square fluctuation (RMSF) of each protein backbone was calculated. As shown in Fig. 3a, five flexible regions were involved in the ATP binding site, including G-loop L836 to L849, hyd-1 A866 to K868, hinge T916 to N923, catalytic subunit H1026 to L1035 and DFG motif C1045 to G1048. Among these compounds, ZINC36470466 produced the highest fluctuations in G-loop regions, while ZINC14762520 showed low RMSF compared to the apo form. Again in all cases, amino acids ranging from 114 to 128 fluctuated more in apo form; however, fluctuations were seen to reduce after ligand binding. In the case of the standard inhibitor, sunitinib increased flexibility in the hinge region, up to 1.9 Å, while no significant changes in RMSF were observed for ZINC14762520. It is also noteworthy that binding of all ligands except ZINC14762520 increased local fluctuations of the catalytic subunit region in comparison to the apo protein. This was due to their strong interactions with residues and rotation of the particular pharmacophore involved in intermolecular interactions. In the DFG motif, ZINC13370550 produced the highest RMSF, which denotes increased flexibility in this region. Overall, RMSF analysis suggested that all compounds formed favorable interactions with ligand binding regions. Although dynamic changes occurred upon binding to these ligands, protein stability was increased, which is similar to the reference ligand.

To further clarify the structural deviation derived by the ligands, flexibilities of the protein in terms of SASA, radius of gyration and number of hydrogen bond were calculated. Figure 4 depicts the total SASA of each protein, i.e., the molecular surface area that is assessable to solvent molecules. Here, decreased values of SASA denote the shrunken nature [71]. The results of SASA analysis suggest that ZINC02585769 increased the SASA of proteins with several large fluctuations in 100 ns simulation. Conversely, ZINC13370550 showed a pattern similar to that of the standard inhibitor sunitinib, indicates that ZINC13370550 was dynamic over the simulation. Comparatively, ZINC36470466 showed lower SASA values over time, allowing us to conclude that the ligand caused tight packing of the protein. Interestingly, ZINC14762520 maintained a steady SASA profile over time; however, some large fluctuations were seen in 100 ns of simulation. On the other hand, the radius of gyration of the proteins, shown in Fig. 5a, revealed that ZINC14762520 produced a higher radius of gyration at the beginning, which was then maintained over the entire simulation.
simulation time. In comparison to the other ligands, ZINC36470466 showed lower values; these results therefore suggest flexibility of ZINC14762520 as well as tight packing of ZINC36470466, as higher radius of gyration value denotes loose packing of the protein structure, which clearly supports the results calculated from SASA analysis. In addition, sunitinib maintained a radius of gyration profile similar to that of the apo state after 17 ns; no significant changes were seen.
thereafter. Figure 5b represents the total intra-residue hydrogen bonds formed during the simulations, in which maximum highest picks were observed for apo form and ZINC14762520. On average, there was insignificant variation of hydrogen bonding among the proteins, ranging from 50 to 100 (Table S2, see supplementary file 1). Nevertheless, it is concluded that binding of ligands, including standard inhibitors, reduced intra-residue hydrogen bonds in the protein, therefore increasing flexibility. The degree of flexibility increased by binding of ZINC13370550 and thus the overall dimension of the structure was enhanced, while it was reduced by the binding of ZINC36470466. The total number of H-bonds between the ligand and residues of the ATP binding site of protein for each complex are represented in Fig. 6, where maximum hydrogen bonding was observed for ZINC36470466 (Fig. 6e) during the simulation, while ZINC13370550 (Fig. 6c) showed minimum interaction (Table 2). For a detailed understanding, H-bond occupancy was calculated and represented in Fig. 6f. Here Fig. 6f shows the percentage of H-bond occupancy, where each residue either acted as a donor or as an acceptor.

As shown in Fig. 6f, ZINC14762520 formed several H-bonds with the residues of DFG motif and hinge region, including C919, E885, F1047 and D1046 residues during the simulation, with total occupancy of more than 15%. The major interaction with C919 was mediated by the hydroxyl group present in the 4-hydroxyphenyl part of ZINC14762520, where the oxygen atom acted as acceptor. ZINC36470466 formed H-bonds with D1046 and E885 residues as H-bond donor, occupying more than 50% of the simulation, while less than 20% of H-bonds were occupied as acceptor with C919 residue. Here, hydroxyl groups present at the 3-hydroxybenzyl part of the compound act as donors in the interactions with D1046 residues of DFG motif. The main contribution of ZINC02585769 in the hydrogen bonding interaction was with residue E885, with 121.04% H-bond occupancy. Here hydroxyl groups present at the 3,4-dihydroxyphenyl part are major contributors in H-bond formation, acting as donors. In contrast, ZINC13370550 showed maximum H-bond occupancy with T926 and C919 residues as both H-bond donor and acceptor, with interaction maintained by
the d-glucopyranose ring and ketone group present in the flavones part, respectively. Sunitinib interacted with residue E^917^ more than 35% as a donor with an additional H-bond with L^840^ having a percentage occupancy of around 10%. As an acceptor, sunitinib maintained H-bonds with C^919^ residues about 34% of total simulation time.

Fig. 3  a Structural representation of VEGFR-2 kinase highlighting the different motifs involved in ATP binding. b Root mean square fluctuation (RMSF) plots of protein–ligand complexes: Blue apo, orange sunitinib, red ZINC02585769, olive green ZINC13370550, black ZINC14762520, pink ZINC36470466. c Comparative RMSF plot between protein-ligand complexes and apo structure. Residues involved: light green G-loop, orange hyd-1, blue hinge, brick red catalytic subunit, cyan DFG motif.
In computer-aided drug design approaches, Poisson-Boltzmann surface area (PBSA) is one of the most attractive solvation systems that are applied widely to calculate binding energy of protein–ligand complexes [51, 72]. Hence, we subjected each protein-ligand complex to a MM-PBSA binding energy calculation to see how structural changes influence ligand binding. With this aim, we calculated the binding energy of every snapshot produced from MD simulations. The results are plotted in Fig. 7a, where more positive energies indicated better binding, according to the theory of nuclear physics [73]. According to the laws of physics, the energy required to disassemble a whole into separate parts is usually positive [74]. On average, ZINC14762520 obtained maximum energy of binding of 166.12 ± 6.5 kJ mol\(^{-1}\), while ZINC02585769 showed the least energy of 19.45 ± 1.4 kJ mol\(^{-1}\). The reference inhibitor exhibited a binding energy of 115.43 ± 1.2 kJ mol\(^{-1}\). By contrast, the ZINC13370550 and
ZINC36470466 complexes showed positive binding energy values of 49.52 ± 5.2 kJ mol$^{-1}$ and 60.92 ± 4.7 kJ mol$^{-1}$, respectively.

Although ZINC02585769 produced higher binding energy at initial stages of simulation, the binding energy was seen to decrease afterward. ZINC36470466 and ZINC14762520 showed similar patterns of binding energy profiles; however, ZINC36470466 obtained lower binding energy after 50 ns of simulation. As ligand flexibility is involved in the formation and breakdown of new bonds in the active site, we analyzed ligand RMSD to gain more detailed understanding. As shown in Fig. 7b, the conformation of the sunitinib fluctuated up to 2.4 Å, and several large decreases in fluctuation were observed one by one over time. Despite several fluctuations, sunitinib maintained binding energy above 40 kJ mol$^{-1}$ in total time of MD simulation. Likewise, over the course of simulation, ZINC14762520 supported stable conformations with RMSD up to 0.7 Å, although several oscillations were observed, which therefore caused variations in binding energy. With regard to ZINC13370550, the ligand showed conformational shifting at 19 ns, which thus produced higher binding energy at that time; however, several large shifting occurred afterward which indicates greater fluctuation in binding energy profile. As a result, these conformational shifts caused weaker binding of the ligand with the active site residues of the protein, therefore resulting in low MM-PBSA binding free energy, as can be observed in Fig. 7a. On the other hand, ZINC02585769 showed a rigid conformation over time, thus showing that flexibility of the ligand plays an important role in the formation of a stable protein-ligand complex.
Overall, MD trajectory analysis infers that ZINC14762520 and ZINC36470466 are the most potent candidates among the herbal secondary in vivo metabolites to inhibit VEGFR-2 kinase. Both these compounds achieved highest average binding free energies in MM-PBSA analysis among the top virtual hits, and the binding energy of ZINC14762520 was greater than the reference inhibitor sunitinib. They also maintained maximum non-covalent interactions in 100 ns with residues of the three most important ligand binding regions in VEGFR-2 kinase, including G-loop, DFG motif and hinge regions [75]. In all kinase domains, G-loop and hinge regions serve as important components for ligand binding, where the G-loop interacts with the purine ring of adenine, and the hinge region acts as the ribose binding pocket [76]. Interestingly, both compounds ZINC14762520 and ZINC36470466 interacted with the residues involved in the hinge region and the DGF motif. Comparatively, ZINC36470466 sustained maximum interactions with hinge regions by its benzyl rings, although ZINC14762520 contains a hydroxyl pharmacophore in the 4-hydroxyphenyl portion that stabilized hinge binding by hydrogen bonding more than 35% in total simulation (Fig. 3f). Furthermore, docking analysis showed that ZINC36470466 formed hydrophobic interactions through its aromatic ring with L840 and V848 residues of the G-loop region. Hence, we calculated the atomic distances between the interacting ring of the ligands and the beta carbon of L840 and V848 residues, respectively. As shown in Fig. S4 (see supplementary file 1), ZINC36470466 maintained distances less than 5.5 Å in 100 ns simulation with these hydrophobic residues, despite exhibiting some fluctuation. Furthermore, L840 maintained maximum intermolecular distance with the ligand from 4 to 5.5 Å in the total simulation, which indicates that L840 of the G-loop contributed stronger hydrophobic interaction than V848. Taken together, the above analysis suggests strong stability and VEGFR-2 kinase inhibition with both ZINC14762520 and ZINC36470466.

Conclusion

In this in silico experiment, we tried to identify potent inhibitors of VEGFR-2 from a database of in vivo metabolites of herbal ingredients, using Glide XP docking and ADME/T analysis to identify the six best candidates. Among them, the top four virtual hits were specified by MM-GBSA calculations. And finally, ZINC14762520 and ZINC36470466 have been selected as the best virtual hits by MM-PBSA and MD simulations. Molecular docking analysis revealed the detailed binding mode of the selected virtual hits with VEGFR-2. According to MD simulation analysis, these compounds energetically formed stable complexes with VEGFR-2 kinase owing to different conformational changes. In addition, all identified compounds were predicted to be orally active and satisfied a drug likeness profile. Compared to standard inhibitors, they are also predicted to be less susceptible to cardiac toxicity. The results of this study therefore offer a suitable starting point for further development as well as detailed in vitro and in vivo analyses.

Acknowledgments We thank Dr. Elmar Krieger, YASARA Biosciences GmbH, for providing an academic version of YASARA dynamics software. The authors acknowledge Prof. Gert Vriend (WHAT IF Foundation / CMBI, Netherlands) for his critical suggestions in Protein Dynamics Simulations. The authors are grateful to the Bangladesh Council of Scientific and Industrial Research for funding under the R&D project (SL. No. 42, 2016-17) to build a computational platform in Bangladesh.

Springer
Compliance with ethical standards

Competing financial interests The authors declare no competing financial interests.

References

4. Plate KH, Breiger G, Weich HA, Risau W (1992) Vascular endothelial growth factor is a potential tumour angiogenesis factor in hu-

AABC 10:11

Advances and Applications in Bioinformatics and Chemistry:


Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
AUTHOR PLEASE ANSWER QUERY.

No Query.