Discriminatory analysis based molecular docking study for in silico identification of epigallocatechin-3-gallate (EGCG) derivatives as B-Raf^{V600E} inhibitors†

Huazhou Ying, a Jiangfeng Xie, a Xingguo Liu, c Tingting Yao, a Xiaowu Dong b,* a and Chunqi Hu a,b

Virtual screening and biological testing were utilized to identify novel B-Raf^{V600E} inhibitors. The employed LigandFit program was evaluated by examining the accuracy of the binding conformation prediction and binding affinity estimation (scoring function) via discriminative analysis training. Ten novel compound hits from the database screening were selected and subjected to in vitro biological tests. The natural product EGCG (A8) was discovered to have promising B-Raf^{V600E} inhibitory, and then we evaluated six structurally similar compounds (B1, B2, B3, C1, C2, and C3) for their B-Raf^{V600E} inhibitory activities in order to establish a structure–activity relationship. One of these compounds, B2, demonstrated a promising B-Raf^{V600E} inhibitory activity with an IC₅₀ value of 9.1 μM, providing a theoretical basis for the development of novel agents as B-Raf^{V600E} inhibitors.

Introduction

B-Raf is one of the most commonly mutating proto-oncogenes with high clinical impact and plays a significant role in the development of various cancer species. Moreover, B-Raf represents an attractive target for small molecules in anticancer drug development. B-Raf belongs to the serine/threonine Raf kinase family, which consists of three isoforms, A-Raf, B-Raf and Raf-1 (also coined C-Raf). It is also an integral member in the Ras–Raf–MEK–ERK signal transduction pathway (also known as the MAPK signaling cascade) participating in cell proliferation and cell survival. In the MAPK pathway, active Ras may induce conformational changes in B-Raf upon binding, followed by the phosphorylation status triggering kinase activity. Then, phosphorylation of B-Raf activates the MEK protein, which in turn activates the phosphorylation of several other substrates. B-Raf kinase is found to frequently mutate in melanomas, with the Val600 → Glu600 (V600E) transition representing the most common mutation, accounting for over 90% in all of the B-Raf oncogenic mutants reported to date. Furthermore, this mutation type causes B-Raf to signal independently from upstream regulation. In an effort to selectively target the B-Raf^{V600E} mutant, numerous clinically used small molecule inhibitors have been published, including LGX-818, PLX-8394, CEP-2496 (https://clinicaltrials.gov/), and BGB-283 [ref. 10] (Fig. 1).

The clinical approval of Vemurafenib11 and Dabrafenib12 for the treatment of metastatic melanoma bearing the B-Raf^{V600E} mutation has further confirmed that B-Raf^{V600E} may be an attractive target for anticancer therapy. Moreover, B-Raf^{V600E} inhibitors have been found to represent crucial drugs used in patients suffering from inoperable metastatic melanomas. However, some cutaneous adverse side effects often occur during clinical use, e.g. dermatologic reactions. In particular, the rate of cutaneous adverse effects associated with the use of Vemurafenib was reported to be 92% to 95% in patients

† Electronic supplementary information (ESI) available: The original data of MD simulation (Tables S1 and S2). See DOI: 10.1039/c7ra04788f
participating in B-Raf inhibitor melanoma (BRIM) studies. To combat this major disadvantage, we designed an efficient virtual screening method to obtain novel structures of B-Raf inhibitors with potentially reduced side effects compared to Vemurafenib.

As one method used in modern structural-based drug design, molecular docking is used routinely to predict the binding strength of ligands in the target binding pocket. Recently, Li et al. reported the identification of organohalogen drugs as B-RafV600E inhibitors by using molecular docking and bioassay. Recently, there is a growing interest in applying various methods to improve the accuracy of molecular docking, ultimately elevating the discriminatory ability of molecular docking to be more efficient. Here, the discriminatory ability to each scoring function by recognizing overlapping regions is generally affected by the receiver operating characteristic (ROC) curve. As such, the efficacy of molecular docking may be significantly enhanced through discriminatory analysis for the selection of docking scores. We have applied this method to discover and develop a novel series of JAK2 inhibitors by using the distributions and ROC curve of LigandFit docking scores to distinguish JAK2 inhibitors from inactive compounds. As part of our continued interest in rational drug design, in the present study, the LigandFit program was evaluated by examining the accuracy of the binding conformation prediction and binding affinity estimation (i.e. scoring function). Virtual screening involves the performance of docking simulations for a large number of inactive diverse compounds and known B-RafV600E inhibitors in order to obtain a rational score function according to the distribution of two series of compound scores. The optimal function was applied to screen the SPECS library and a natural product library as well as the activity of the compounds obtained by screening was evaluated. Fortunately, we found that epigallocatechin-3-gallate (EGCG, A8), a compound as the main chemical constituents of green tea, exhibits moderate inhibitory activity against B-RafV600E. Indeed, other natural products featuring large-scale structural diversity have been the major sources of bioactive agents in the fields of medicine and life sciences. Six other compounds featuring similar structural characteristics as EGCG (Fig. 5B) were biologically evaluated for their B-RafV600E inhibitory activities. Most of these compounds exhibited moderate inhibitory activity of B-RafV600E with IC_{50} values ranging between 9.1 and 43.6 μM. Among these compounds, compound B2 demonstrated a promising inhibition of B-RafV600E with an IC_{50} value of 9.1 μM. Interestingly, the screened natural products feature vastly different molecular scaffolds and are structurally distinct from current B-RafV600E inhibitors. This further demonstrates the feasible identification of novel hit compounds using our discriminatory ability training-based molecular docking study. Finally, molecular dynamics calculations were implemented to explore the potential binding mode of the promising compounds with the B-RafV600E protein, providing important information for further structural modification considerations and structure–activity relationship studies.

Methods

Docking-based virtual screening by LigandFit

Operating environment: Discovery Studio 2.5. Preparation of the target binding pocket: (a) the crystal structure of B-RafV600E (pdb: 4EHG) was employed as the template for molecular docking with the LigandFit protocol; (b) all crystallographic water molecules were removed; (c) all hydrogen atoms were added by CHARMM force field of “Prepare Protein module”; (d) the “Find Sites as Volume of Selected Ligand” tool was applied to define the docking site of the resulting protein structure. Preparation of the ligand library: (a) variable numbers of Monte Carlo simulations were implemented for different conformer generations of ligands; (b) all atoms were added by CHARMM force field; (c) all conformers were treated with the “ligand minimization” of the receptor–ligand interaction module.

LigandFit protocol: (a) the ligand–receptor interaction energies were calculated based on the Piecewise Linear Potential 1 (PLP1) force field; (b) a short rigid body minimization was performed and the top one pose for each ligand was saved; (c) scoring was performed with a set of scoring functions employed in LigandFit module, including LigScore1_Dreiding, LigScore2_Dreiding, -PLP1, -PLP2, Jain, -PMF, Dock_Score.

In vitro biological assays for B-RafV600E (or B-Raf) inhibitory activity

The inhibitory activities of the compounds against the B-RafV600E enzyme (Invitrogen, USA) were evaluated according to a published procedure. Specifically, 2.5 μL of a reaction solution in assay buffer (50 mM HEPES, pH 7.5, 10 mM MgCl2, 1 mM EGTA, 0.01% BRJ-35) was placed in a 384-well plate containing 0.6 nM B-RafV600E (or 2 nM B-Raf (Invitrogen, USA) for B-Raf assay), 0.2 μM Fluorescein-MAP2K1, and 1.5 μM ATP (or 0.1 μM ATP for B-Raf assay). The reaction was then incubated at room temperature for 1 h and subsequently quenched with 10 μL of a detection solution containing 2 nM antibody (Cat. no. PV3574, Invitrogen, USA) and 10 mM EDTA. After incubation for 30 min, the reaction mixture was analyzed on a Caliper LabChip 3000 (Caliper LifeSciences, Hopkinton, MA, USA) by electro-photoreception of the fluorescent substrate and the phosphorylated product. All compounds were first evaluated at a concentration of 10 μM and then further tested at the concentration prepared from 3-fold serial dilutions exhibiting a preliminary activity of more than 40%. The IC_{50} values were determined by the analytical software Prism 5.0 (GraphPad Software Pte Ltd.).

Molecular dynamics (MD) simulations

The docked structures of an inhibitor bound into B-RafV600E were used as the initial structures for MD calculations. A CHARMM force field was applied to the complex and the resulting system was then immersed into an “Explicit Periodic Boundary” water box with a sodium cation used for complex neutralization. Afterwards, the solvated system was subjected to double-fold minimization (10 000 cycles of steepest descent minimization and 100 000 cycles of conjugate gradient
minimization). The system was gradually heated from 50 K to 300 K over a period of 100 ps and subsequently equilibrated for 500 ps. Starting from the last frame of the equilibration, a production simulation was performed for 500 ps using the NPT ensemble under a constant temperature of 300 K and pressure of 1 atm. Other parameters of MD simulation were maintained at the default Discovery Studio configuration.

Results and discussion

Construction of drug-like library

Retrieved from the chemical SPECS database (http://www.specs.net), a total of 217 190 compounds were prepared by the Prepare Ligands protocol of Accelrys Discovery Studio 2.5 program including the following steps: (1) the two-dimensional (2D) structure was converted into a three-dimensional (3D) structure, (2) charges were calculated, (3) H-atoms were added, and (4) the “Lipinski’s rule of five” filter was applied to the resulting compounds to simplify the compound library. After the above preparation, a drug-like library containing 153 691 compounds was obtained. Another 623 compounds from the natural product library (NPD from Zinc database), were prepared as well.

Evaluation of scoring functions in LigandFit

Accurate scoring functions play a key role in identifying active B-RafV600E inhibitors through scoring. The scoring functions are expected to be able to efficiently distinguish between the active compounds and the inactive compounds. In order to accomplish this goal, 731 B-RafV600E inhibitors were acquired from the Binding DB database, and 10 000 non-active compounds were randomly retrieved from the Maybridge database using the Find Diverse Molecules module of DS 2.5. The two groups of compounds gained a maximum diversity through 2D similarity studies based on the FCFP_6 fingerprints. Subsequently, all molecules were treated with a Prepare Ligands protocol, docked into the active site of B-RafV600E and evaluated by the different scoring functions of the LigandFit module, respectively. The docking scores (LigScore1_Dreiding, LigScore2_Dreiding, -PLP1, -PLP2, Jain, -PMF, Dock_Score) and distributions of the two series (Fig. 2) were analyzed. From the data analysis of Dock_Score in Fig. 2, the differentiating capacity of Dock_Score was found to be the most promising, with the minimal overlapping area of the active B-RafV600E inhibitor and the inactive compound.

We also examined the discriminatory ability of these scores by receiver operating characteristic (ROC) curve studies. Consistent with previous results, Dock_Score exhibited a preferred corresponding AUC (area under the ROC curve) value of 0.982 when compared to the other scores. From the analysis of distributions and ROC curves, it can be concluded that the Dock_Score of the LigandFit module represents an objective scoring function for the present study and most of the scores of Dock_Scores were found between 115 and 135.

Virtual library screening using molecular docking

All of the structures in the two drug databases were docked in the active binding pocket by DS 2.5. In general, Dock_Score...
identifies B-Raf\textsuperscript{V600E} inhibitors more accurately than other scoring functions and was therefore used in virtual screening to obtain the optimal compound. Finally, further biological testing of the docking and structural diversity of the ten target compounds (Fig. 3) was carried out.

**In vitro B-Raf\textsuperscript{V600E} inhibitory activities assay**

The ten target compounds were purchased from SPECS and evaluated for their B-Raf\textsuperscript{V600E} inhibitory activities. Interestingly, compound A8 showed significant B-Raf\textsuperscript{V600E} inhibitory activity with an IC\textsubscript{50} value of 11.1 μM and an inhibition rate of 44.1% at a concentration of 10 μM (Fig. 4). Compound A8, also known as epigallocatechin-3-gallate (EGCG), represents a compound extracted from green tea and has been shown to exhibit biophysical effects as described in a number of reports found in the literature.\textsuperscript{30–32} However, some properties have never been studied before and a variety of other factors need to be considered when using EGCG. In the field of antitumor activity, EGCG has been found to reduce the proliferation of cancer cells and induce apoptosis,\textsuperscript{33,34} more specifically, mutant BRAF melanoma cell lines were found to be sensitive to EGCG.\textsuperscript{35,36} However, the metabolic reason for this anticancer activity has never been described. Here, we further explored the inhibitory activity of EGCG and its derivatives on B-Raf\textsuperscript{V600E}. Furthermore, preliminary structure–activity relationships were evaluated through a docking study.

EGCG was used as a starting point for further studies. The binding pattern of EGCG within the proposed B-Raf\textsuperscript{V600E} active site is shown in Fig. 5A. In order to improve the structural analysis of EGCG, the molecular structure was separated into three individual sections, A fragment, B fragment and linker. Of these, we hypothesize that the A fragment is important for the B-Raf\textsuperscript{V600E} activity. This notion may be due to the fact that among the previous 10 compounds, only the A8 structure was found to feature the A fragment, presumably resulting in inhibitory activity. Thus, in the following study, the A fragment as the primary functional group was retained while the other two fragments were optimized. We believe that the number of hydroxyl groups may be adjusted to increase the combinatory ability of the inhibitor. On the other hand, we also considered modifying the extension connectors to better match the
pockets. According to the docking evaluation, the structural design of the derivatives was carried out focusing on two parts: the hydrolysis of EGCG esters and the substitution of different substrates. Finally, we evaluated the activities of six compound species, B(1–3) and C(1–3), which were synthesized as reported previously (Fig. 5B). The B-RafV600E inhibitory activities of these compounds were evaluated and GW5074 was used as positive control. The results obtained represent the mean of three experiments and are expressed as IC_{50} values (Table 1). Compound B2 demonstrated an improved inhibitory activity against B-RafV600E, with an IC_{50} value of 9.1 μM. Other compounds exhibited somewhat lower B-RafV600E inhibitory activities. The data showed that two hydroxyl groups are indeed necessary for inhibitory activity and the studied inhibitors may therefore provide a reasonable basis for the future development of advanced agents to treat metastatic melanoma. Besides, we tested the wild-type Braf inhibitory activity of representative compounds EGCG and C2, both of which show weaker activity against wild-type Braf (IC_{50} values of EGCG and C2 are 57.86 μM and > 200 μM, respectively). Thus, the compound C2 exhibited even more selective to B-RafV600E than that of EGCG, with selective index more than 15.7 folds.

### Table 1 Inhibitory activity for synthesized compounds against B-RafV600E

<table>
<thead>
<tr>
<th>Compounds</th>
<th>IC_{50} (μM)</th>
<th>Compounds</th>
<th>IC_{50} (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGCG (A8)</td>
<td>11.1</td>
<td>GW5074</td>
<td>0.0015</td>
</tr>
<tr>
<td>B1</td>
<td>43.6</td>
<td>C1</td>
<td>22.1</td>
</tr>
<tr>
<td>B2</td>
<td>9.1</td>
<td>C2</td>
<td>12.7</td>
</tr>
<tr>
<td>B3</td>
<td>25.1</td>
<td>C3</td>
<td>30.1</td>
</tr>
</tbody>
</table>

#### MD simulations and mechanistic analysis of interactions

In order to investigate the interaction mode between the most promising compound, B2, and the B-RafV600E protein, molecular dynamics (MD) simulations and an interaction decomposition analysis were carried out to explore the quantitative energy contributions per residue to the binding affinity of the most promising inhibitor species B2. The RMSD value of B-Raf was found to reach equilibrium and remained an average value after a simulation time of 100 ps. The RMSD value of the protein backbone was calculated from a 200 to 500 ps trajectory and the data points were found to fluctuate 0.574 ± 0.06 nm. After 500 ps production simulation, the distance between the inhibitor B2 and the key amino acids Ser465, Thr529 Asn580 and Asp594 in B-Raf tends to converge, indicating that the system reached equilibrium conditions (Fig. 6c and Table S1, ESI†). Based on the stable conformation obtained from MD simulation, the interactions involved in the protein/ligand complex were analyzed. As shown in Fig. 5B, the scaffold was found to
establish H-bonds in inhibitor B2 with NH and the carbonyl group in the B-Raf hinge residue ASN580 establishing an L-shaped geometry. Furthermore, the hydroxy group on the scaffold formed H-bonds with Ser465. In accordance with the interaction energy decomposition, the contribution of ASN580 and Ser465 in B2 was determined to be −10.3 kcal mol$^{-1}$ and −6.4 kcal mol$^{-1}$, respectively (Table S2, ESI†). The trihydroxylphenyl substituent in inhibitor B2 was found to be located in the hydrophobic pocket adjacent to the Thr529 group. Besides, the m-dihydroxyphenyl group in B2 also forms H-bonds with Asp594. Therefore, according to the binding pattern and chemical structure of B2, we speculate that the aromatic moiety, linked by the catechins scaffold containing hydrogen-donors and receptors, may represent an important pharmacophore target of the studied B-Raf$^{V600E}$ inhibitor.

### Conclusion

In this work, we established a reliable molecular docking scoring function evaluation method and used this method for the screening of a compound library to obtain a potent small molecule inhibitor, A8 (also known as EGCG), with potent B-Raf$^{V600E}$ inhibitory activity (IC$_{50}$ value of 11.1 μM). On this basis, we screened six structurally similar compounds (B1, B2, B3, C1, C2, and C3) for their B-Raf$^{V600E}$ inhibitory activities based on a rational structure design. One of these compounds, B2, exhibited a promising B-Raf$^{V600E}$ inhibitory activity with an IC$_{50}$ value of 9.1 μM. Further studies are currently carried out in our laboratory to evaluate the pharmacological potency of this compound and a follow-up report will be published shortly.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

This study was financially supported by National Natural Science Foundation of China (81673294 and 81502926) and Zhejiang Province Public Welfare Projects (2016C33067) and China Postdoctoral Science Foundation (2017M612017).

### Notes and references