Ligand Based Virtual Screening

Introduction

Melanoma is a highly aggressive cancer type with a strong tendency to metastasize. When diagnosed at an early stage, removal of local tumours is mostly curative, however, late diagnosis of invasive or metastatic melanoma often carries a poor prognosis [1]. Melanomas originate from the malignant transformation of melanocytes, neural-crest derived cells specialized in the production of the pigment melanin, which contributes to skin and hair pigmentation and confers ultraviolet light protection [2]. Cutaneous melanoma, originated from the melanocytes at the basal layer of the epidermis, is the most common form of the disease.

The malignant transformation of melanocytes is the result of a complex interaction between both exogenous and genetic factors [1]. Ultraviolet (UV) exposure is the main risk factor of cutaneous melanoma. UV radiation, stimulates cyclic AMP production through a melanocortin-dependant mechanism, leading to increased proliferation and melanogenesis [3]. This proliferative process is regulated by BRAF, a serine-threonine kinase that triggers the MAP kinase pathway. The importance of this pathway in melanocyte growth might explain its extraordinarily high mutation frequency. Importantly, up to 70% of human melanomas harbour mutations in BRAF [3].

The most common BRAF mutation, which accounts for more than 90% malignancies involving this gene, is a substitution of glutamic acid for valine at position 600 (V600E) [3]. Mutated $BRAF_{V600E}$ shows approximately a 500-fold increased activity, and it induces constitutive signalling through the RAS–MEK–ERK pathway (MAP kinase signalling pathway) and nuclear factor kappa-B (NF-kB) activation [2, 4]. Through complex downstream mechanisms, BRAF activity promotes cell survival, proliferation, angiogenesis, invasion and metastasis [5].

Mutated BRAF has been observed to play an important role in malignant melanocyte transformation and is already present and in up to 80% of the benign skin lesions called naevi [2]. Moreover, mutated BRAF activity stimulates many of the hallmarks of cancer. This suggests that BRAF mutation might be a founder event in the onset of the disease and sustain its progression [2].

In view of this, we propose that BRAF inhibition would be an optimal therapeutical approach for $BRAF_{V600E}$-mutated melanomas. Through the inhibition of the BRAF serine/threonine-protein kinase we would be able to stop the signalling cascade that leads to cell proliferation and survival at an upstream position. Moreover, since BRAF mutation appears at early stages of the disease, it could be an appropriate potential treatment to avoid melanoma progression.

Methods

1. Ligand Based Virtual Screening

First, we performed a rapid search in ChEMBL database to identify out target protein, the single protein Serine/threonine-protein kinase B-RAF (CHEMBL5145). Once identified, we used the Python client for accessing ChEMBL RESTful API service (https://chembl.gitbook.io/chembl-interface-documentation/web-services) and searched approved drugs for our target. Among the 5 approved drugs obtained, Vemurafenib or PLX4032 (CHEMBL1229517) was the first approved

$BRAF_{V600E}$ inhibitor (First in Class) and the one with the largest number of Lipinski rules fulfilled. Therefore, it was selected as query molecule for similarity screening.

Next, we performed a combined 2D+3D similarity screening for our query molecule in PubChem database using its PUG-REST interface (https://pubchem.ncbi.nlm.nih.gov/rest/pug/); 2D- and 3D-similarity thresholds were fixed at 85% and 95% respectively. We analysed the resulting molecules, calculated their MACCS fingerprints using Python's library RDKit, and filtered them with a 95% Dice similarity coefficient cut-off.

2. Quantitative structure-activity relationship (QSAR)

As a further step in the ligand-based screening process, we developed a Quantitative Structure-Activity Relationship (QSAR) model, aiming to predict the physicochemical properties of our similarity-filtered molecules. To do this, we retrieved a set of 1000 molecules with biological activity towards our target protein (CHEMBL5145), considering only assays evaluating half maximal inhibitory concentration (IC50). With this data, we trained a Random Forest Regression model correlating molecule fingerprints with their experimental IC50 values, and used it to predict the IC50 of our similarity-filtered molecules. Only drug candidates with a predicted IC50 value lower than 180nM were saved for further analysis.

3. ADMET Properties

Before analysing binding affinity, we want to ensure that shortlisted drug candidates fulfil the Lipinski's Rule of Five criteria (RO5) [8]. These rules assess the druglikeness of chemical compounds based on specific chemical and physical properties important for pharmacokinetics -Absorption, Distribution, Metabolism, Elimination and Toxicity (ADMET)- [9]. Lipinski's RO5 evaluates the number of hydrogen (H) bond donors (>5), the number of H bond acceptors (>10), molecular mass (<500 Da) and octanol-water partition coefficient (log P, <5). Molecules fulfilling these criteria were considered for further structure-based analysis.

4. Protein-Ligand Docking

As a last step in our ligand-based virtual screening procedure, we performed protein-ligand docking to predict the complex structure formed by our target protein ($BRAF_{V600E}$) and the shortlisted drug candidates, and asses their binding affinity. We fetched the 3-dimensional structure of mutant $BRAF_{V600E}$ in complex with PLX4032 (Vemurafenib) from Protein Data Bank (PDB ID: 3OG7) (https://www.rcsb.org/structure/3OG7). $BRAF_{V600E}$ kinase assembles forming homodimers and both subunits are equally sensitive to PLX4032 inhibition [10]. In our fetched 3OG7 PDB experimental structure, PLX4032 is anchored to only one monomer, therefore, this subunit was used as a structure for docking.

Before docking, protein structure was protonated and treated to remove the ligand and unnecessary molecules (water, ions, crystallographic solvents) using LePhar's LePro software (http://www.lephar.com/software.htm). Similarly, ligands were processed using RDKit and Openbabel's Python implementation (Pybel) to add hydrogens, generate 3D coordinates and perform local optimization though 500 Merck molecular force field (MMFF94) steps. Coordinates of the binding pocket were manually calculated as the average coordinates of every

atom of the PLX4032 ligand obtained from the 3OG7 PDB file, and a standard pocket size of 20Å x 20Å x 20Å was fixed.

Once the target protein and the ligands were properly processed, we performed docking using Smina (https://sourceforge.net/projects/smina/), a fork of AutoDock Vina for scoring and high-performance energy minimization. Smina was run for every ligand with the previously obtained coordinates and dimensions, a maximum number of 3 nodes, a maximum exhaustiveness of 8, and a randomly defined seed. As a result of the analysis, we obtained the locally optimized binding affinity of each ligand and their 3-dimenisonal coordinates in a sdf file. Docking results of the best performing molecules were visualized and double-checked using Chimera software and its AutoDock Vina implementation.

Results

Five serine/threonine-protein kinase B-RAF (CHEMBL5145) inhibitors approved by the FDA were detected. Vemurafenib (CHEMBL1229517) was the first in approval date and one of the best performing regarding ADMET properties, with only one RO5 violation (2 H bond donors). Among the 6759 molecules initially fetched from the similarity screening, only 682 overcome Dice similarity coefficient threshold using MACCS fingerprints.

The QSAR-predicted IC50 values for the similarity screened molecules ranged from 94.8nM to 23.25µM, and only 45 molecules showed a predicted IC50 value greater than the 180nM cut-off. Experimental IC50 values obtained for Vemurafenib vary largely in different assays and cell lines, ranging from suspiciously low (4nM) to higher (290nM) values in the nanomolar scale, as observed in ChEMBL database. Therefore, we can conclude that similarity-filtered molecules show at least similar or improved inhibitory activity.

Among the 45 shortlisted molecules, 33 of them fulfilled Lipinski's RO5 criteria. The most common "limiting" rules among these compounds, proximate to RO5 violation, were molecular weight and log P value. This filtering results suggest that our drug candidates have, at least theoretically, better pharmacokinetics with respect to Vemurafenib.

The docking results can be observed in Table 1. All drug candidates showed a minimized binding affinity lower than -9, and among them, 2 compounds had affinity values below -11. We can consider that molecules 8 (CID: 54586539), 5 (CID: 91896025) and 11 (CID: 91896025), with the lowest binding affinities, are optimal potential druglike candidates to be tested experimentally.

	minimizedAffinity	ID
8	-11.14773	mol_54586539_8
5	-11.05601	mol_91896025_5
11	-10.95070	mol_44587235_11
2	-10.77955	mol_11653652_2
0	-10.73782	mol_11531192_0

Table1. Docking results sorted by binding affinity. Only best 5 candidates are shown

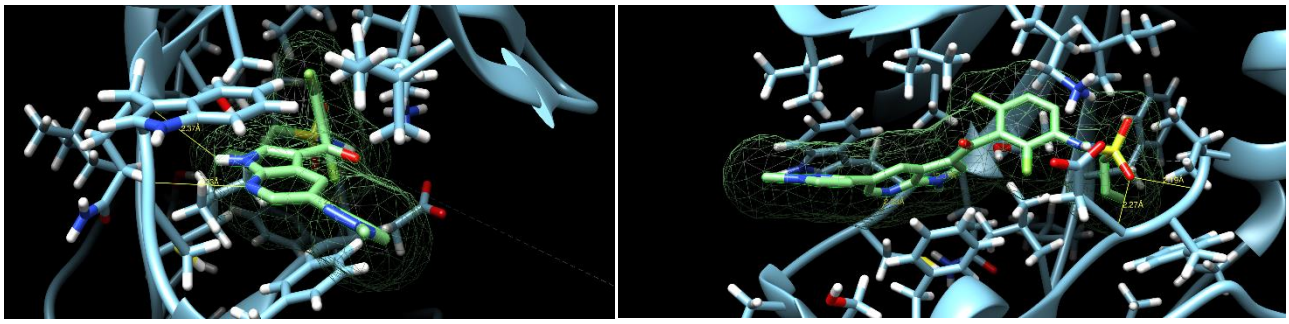

Figure 1. Docking results visualized on Chimera software. Four hydrogen bonds can be observed between the ligand and the target structure.

Discussion

In this work, we performed a ligand based virtual screening aiming to identify a novel potential drug candidate for $BRAF_{V600E}$-mutant melanomas. Using a pipeline of different computational tools, we have been able to identify a small subset of compounds with high druglikeness potential, which should be considered for further experimental testing.

Importantly, these compounds show similar or higher inhibitory capacity, and improved pharmacokinetic properties as compared to the current approved $BRAF_{V600E}$ inhibitor Vemurafenib (PLX4032). Therefore, novel drug candidates proposed in this work could have great therapeutic potential in $BRAF_{V600E}$-harbouring melanomas.

Supplementary Code

Complete code of the analysis can be found in the following Google Drive folder:

https://drive.google.com/drive/folders/10zZuoWy-DU6JB49XIHUUJ7iE4KcoULz3?usp=sharing

References

1. Schadendorf, D., van Akkooi, et al. (2018). Melanoma. The Lancet, 392(10151), 971–984.
2. Dhomen N. and Marais, R. (2007). New insight into BRAF mutations in cancer. Current Opinion in Genetics & Development, 17:31–39.
3. Davies H, Bignell GR, et al. (2002). Mutations of the BRAF gene in human cancer. Nature, 417:949-954.
4. Satyamoorthy, K., Li, G., et al. (2003). Constitutive mitogen-activated protein kinase activation in melanoma is mediated by both BRAF mutations and autocrine growth factor stimulation. Cancer Res, 63:756-759.
5. Chavda J., Bhatt H. (2020). Systemic review on B-RafV600E mutation as potential therapeutic target for the treatment of cancer. European Journal of Medicinal Chemistry, 206, 112675.
6. Ullah, R., Yin, Q., et al. (2021). RAF-MEK-ERK pathway in cancer evolution and treatment. Seminars in Cancer Biology.
7. Hansen, J. D., Grina, J., et al. (2008). Potent and selective pyrazole-based inhibitors of B-Raf kinase. Bioorganic & Medicinal Chemistry Letters, 18(16), 4692–4695.

8. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development. Advanced Drug Delivery Reviews, 46(1-3), 3–26.
9. Ferreira LLG, Andricopulo AD. ADMET modeling approaches in drug discovery. Drug Discov Today. 2019 May;24(5):1157-1165.
10. Bollag, G., Hirth, P., Tsai, J. *et al.* Clinical efficacy of a RAF inhibitor needs broad target blockade in *BRAF*-mutant melanoma. *Nature* **467**, 596–599 (2010).