IN SILICO IDENTIFICATION OF APOBEC3B SMALL MOLECULE INHIBITORS FROM DTP-NCI LIBRARIES

MARYAN MOHAMUD MOHAMED1, NOR ATIQAH JUSRIL1, MOHD ILHAM ADENAN1,2, N. G. KWOK WEN*3

1Faculty of Applied Sciences, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia, 2Universiti Teknologi MARA, Cawangan Pahang, 26400, Bandar Tun Abdul Razak Jengka, Pahang, Malaysia, 3School of Pharmacy, Management and Science University, Seksyen 13, 40100, Shah Alam, Selangor, Malaysia

Email: kwng@msu.edu.my

Received: 04 Jan 2021, Revised and Accepted: 20 Mar 2021

ABSTRACT

Objective: APOBEC3B (A3B) enzyme causes C-to-T or C-to-G somatic alteration in the cancer genome, leading to the evolution of a broad spectrum of human cancers. The present study aims to identify A3B small molecule inhibitors using a top-down approach via pharmacoinformatic virtual screening.

Methods: Virtual screening of 2951 drug-alkie molecules with diversified structures from the National Cancer Institute Development Therapeutics Program (DTP-NCI) compounds library was performed using GOLD and AutoDock Vina docking programs against the 3D structure of A3B (PDB ID: 5TD5).

Results: Amongst the docked compounds, Nordracorubin, NSC641233 and Rakoxifene hydrochloride showed the most potent binding affinities towards A3B on both AutoDock/Vina and GOLD. Several significant similarities were observed between A3B and the three hits, including hydrogen bonds and pi-pi stacking. The three compounds also exhibited interaction with the centralized zinc cofactor and amino acid residues that directly contribute the deaminase activity of A3B enzyme.

Conclusion: We hypothesize that the findings from this study could significantly shorten the quest for novel molecules against the A3B after verification with subsequent in vitro and in vivo studies in the near future.

Keywords: APOBEC3B, Molecular docking, DTP-NCI, AutoDock-Vina, Gold

INTRODUCTION

The APOBEC ("apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like") in human is part of a broader superfamily of zinc-dependent cytidine deaminase enzymes that consists of 11 family members, including APOBEC1, activation-induced deaminase (AID), APOBEC3A, APOBEC3B, APOBEC3C, and APOBEC4. These proteins remove the 4-NH2 group from cytosines/cytidines on single-stranded (ssDNA) or RNA and convert them into uracils (C-to-U conversion). This deamination process is part of the innate immune responses to hypermutate pathogenic viral genomes, such as retroviruses, DNA viruses and endogenous retroelements [1].

APOBEC3B (A3B) is a sub-unit member of the APOBEC3 (A3s) family localized in the nucleus. The core structure of A3B is supported by a backbone of five β-sheet strands and six α-helices motifs with a centralized zinc ion. Despite having both C-terminal catalytic (CTD) and N-terminal pseudo-catalytic (NTD) zinc domains, only the CTD in A3B has been identified to have enzymatic activity. The amino acids around the CTD of A3B are approximately 91% similar to that of A3A [2-3]. Upon binding to the ssDNA, the substrate target cytosine will be inserted deep into the A3B active binding cavity containing a zinc ion and amino acid residues of Cys104, Cys106, and His253 [2].

Despite having intrinsic ability to ablate viral infection, dysregulation of A3s activities, particularly A3A and A3B, were linked with various human cancers progression through DNA hypermutation and promiscuous RNA editing [3]. Mutational signatures attributed to the action of A3B, characterized by C-to-T and C-to-G mutations predominantly in the 5'-TCG motifs, were detected in a large subset of human cervical, lung, bladder, and breast cancers [4-6]. Overexpression of A3B in breast cancer results in an adverse effect for estrogen-receptor-positive (ER+) breast cancer patients, as well as other undesirable outcomes in cancer development, metastasis, therapy resistance, and cell death [7-9]. Collectively, these studies demonstrate the role of A3B in cancer mutagenesis, and such this enzyme may present as a promisingly therapeutic target and marker for the development of a new class of cancer treatments.

Over the last decades, computer-aided virtual screening applications have played an imperative role in discovering and developing bioactive molecules [10]. Docking algorithms that markedly enhance the precision and efficiency in hit identification and lead optimization have been designed and released. AutoDock, Vina, Gold, Libdock, and CDOCKER are the most commonly used [11]. Molecular dynamics simulation of A3B phosphorylation revealed that phosphorylation of Thr214 restrained the enzyme deaminase activity by blocking ssDNA to bind within the A3B enzyme active site [12]. Nevertheless, there has been little discussion on the receptor-based virtual screening of A3B using different docking algorithms against commercial or publicly available compound libraries. Hence, the search for new active anticancer drugs based on the inhibition of A3B enzyme is still required. In this present study, virtual screening of 3000 drug-alkie chemical scaffolds obtained from the Developmental Therapeutics Program, National Cancer Institute (DTP-NCI) was carried out to identify the potential inhibitory compounds against A3B deaminase enzyme.

MATERIALS AND METHODS

Molecular docking (MD) was performed against small compounds obtained from the DTP compound databases into the A3B active site pocket. The schematic diagram of the study is illustrated in fig. 1. Protein identification

The X-ray crystallography of human A3B bound with ssDNA was downloaded from the Protein Data Bank (PDB) (ID: 5TD5) with 1.72 Å resolution [2]. Subsequently, to authenticate the identity of A3B, the protein sequence was compared with all other A3s amino acid sequences using the Basic Local Alignment Search Tool (BLAST). The A3B sequence from PDB was 100% identical to the similar gene clusters that encode the APOBEC3B_HUMAN DNA dC->dU-editing...
from BLAST. Water and ligands bonded to the A3B structure were then removed to clear the binding pocket for later docking computations.

![Image](https://PLACEHOLDER)

**Fig. 1:** Schematic diagram of *in silico* screening on NCI diversity compounds for A3B enzyme inhibitors

**Ligand preparation**

The small molecule chemical structures were downloaded from the National Cancer Institute Development Therapeutics Program's (DTP-NCI) database. The DTP-NCI data set we used in the current study includes Approved Oncology Drugs Set with 166 compounds; Diversity Set VI with 1584 compounds; Mechanistic Set V with 811 compounds, and; Natural Products Set V with 390 compounds. The complete details of the compounds can be obtained from the DTP-NCI website (https://dtp.cancer.gov/). All the chemical structures were optimized using Merck Molecular Force Field 94 (MMFF94) calculations within CambridgeSoft ChemBio3D software (PerkinElmer Inc., USA).

**Molecular docking studies**

**Vina and AutoDock 4.2**

*In silico* screening was performed using Vina and AutoDock 4.2 docking program [13-17]. The ligands were allowed to have flexible rotational bonds, while the protein was kept rigid. Before the virtual screening, an in-depth analysis of the A3B structure was conducted to understand its binding pocket. Using AutoGrid, a grid map with the following settings was generated to cover the binding pocket: a) grid box with the dimensions of 50 x 50 x 50 and grid spacing of 0.375 Å, and b) grid center designated at x, y, and z dimensions of 7.14, 30.17-14.07, respectively. The grid map files of the atoms present in the ligands were also calculated. In the docking protocol, the Lamarckian genetic algorithm was employed as the conformational search, and the population size was set to 150, while the other settings were left as defaults. The docking protocol was saved as a DPF file. The setting in the DPF file was then used by the AutoDock command tool to generate a DLG file that contained data on the ligand-protein binding positions [10].

**Genetic optimization for ligand docking (GOLD)**

GOLD Suite (Hermes 1.7.0) was employed for the docking analysis [18]. Protein preparation was done using the GOLD Suite Hermes visualizer. First, hydrogen atoms were added to the protein, whereas any excessive water molecules in the active site pocket were removed. The free rotation of the sidechain was restricted. Default GOLD parameters were employed to perform all the docking simulations within the pre-defined binding site. A hundred percent search efficiency was applied with a minimum of 10,000 and a maximum of 125,000 operations for each ligand molecule. The docking was performed with a full range of ligand conformational flexibility. Different scoring functions, including ChemScore Fitness, Gold Score, and Piecewise Linear Potential (PLP), were used to rank the ligand [10].

**Structural analysis and visualization**

AutoDock tools produced multiple predictions for each protein-ligand bound structure and ranked the docked conformations according to ascending order of energy and distance. The best protein-ligand poses were selected based on the lowest energy representative of each ranked cluster, and the results from Autodock Vina were exported as pdbqt output. The protein-ligand interaction was then visualized and further analyzed using Discovery Studio 2.5 (Accelrys Inc., USA).

**RESULTS AND DISCUSSION**

We used two automated docking programs, AutoDock/Vina, and GOLD to systematically investigate the potential A3B inhibitors from 2951 selected DTP small molecules. The loop region within the A3B X-ray crystallography structure (PDB: STD5) (fig 2) was used as the starting point and search target. This loop region within the structure may be the hotspots for APOBEC mutagenesis, as revealed by co-crystal structures prepared by other researchers and validated through biochemical analyses [19].

![Image](https://PLACEHOLDER)

**Fig. 2:** Crystal structure of A3B (PDB: STD5) with 1.72 Å resolution

The screening results were ranked based on the compounds binding energy, and hits with the most substantial binding energies consensus in both AutoDock/Vina and GOLD algorithms were shortlisted. Aaurintricarboxylic acid (ATA), a non-specific compound inhibitor for A3s, was used as a reference ligand. Based on the simulations performed, three compounds, noradcarobin, NSG541233, and 3aloxifen (RAL), were strongly associated with the active protein site, more than the reference ligand (fig 3, table 1).

Visual inspections of the interactions between A3B and the three compounds were shown in fig. 4. The docking results revealed that all the three compounds were well-positioned in the A3B active site domain, mainly with the centralized zinc cofactor and amino acid residues of Trp281, Tyr313, Tyr315, His212, Trp287, Asn240, Arg211, Thr214, and His253 which directly contribute the deaminase activity of A3B enzyme. Shi et al. [2] reported that the ssDNA bound to A3B active site pocket containing a zinc ion coordinated by His253, Cys284, and Cys289. Our docking results also reviewed a similar establishment of molecular interactions network, mainly hydrogen bonds and π-bonds with A3B active residues, especially His253. The aromatic rings of interacting amino acid residues and the small compounds resulted in strong π-bonds, which increases the stability of the ligand-enzyme complexity [20].
Fig. 3: The chemical structures of identified active hits against A3B enzyme. From upper right to left: (i) Nordracorubin; (ii) NSC641223; (iii) Raloxifene hydrochloride (RAL), and; (iv) Aurintricarboxylic acid (ATA)

Table 1: GOLD scores and AutoDock/Vina binding energy of three potential A3B inhibitory compounds

| Compounds                     | NSC/NCI diversity Set | Chemical formula | Binding amino acid residues | GOLD (Score/fitness) | AutoDock/Vina (Kcal/mol) |
|-------------------------------|------------------------|------------------|----------------------------|----------------------|--------------------------|
| Nordracorubin                 | 376,248                | C_{31}H_{22}O_{5} | His212, Trp281, Tyr313, Tyr315 | 37.1536/37.2649      | 67.9874/6.8              |
| NSC641233                     | 641,233                | C_{32}H_{18}ClNO_{7} | Ala254, Asn240, His212, His253, Thr214, Trp287, Tyr313 | 61.6349/36.3795      | 85.8062/7.35             |
| Raloxifene hydrochloride      | 747,974                | C_{28}H_{27}NO_{4}S | His212, His253, Trp287, Tyr313, Tyr315 | 30.0083/37.5152      | 65.2962/7.28             |
| Aurintricarboxylic Acid (ATA) | 4056                   | C_{22}H_{14}O_{9} | Asn240, His212, His253, Tyr313, Tyr315 | 47.512/32.934        | 65.321/4.0              |

Nordracorubin formed a conventional hydrogen bond with Trp281 at O2 and π-donor hydrogen bond with Tyr313 (Fig. 4i). Hydrophobic π-σ interaction occurred between aromatic rings of the compound with Tyr315 and His212. Besides that, π-σ stacked and π-alkyl with these residues were also identified. The H-bond and π-σ interaction significantly enhanced the enzyme-substrate complex’s stability [21].

Meanwhile, NSC641233 developed a combination of hydrophobic π-π interactions, π-cation, and sidechain hydrogen bond interactions with A3B (Fig. 4ii). The compound formed conventional H-bonds with Trp287, Tyr313, and Asn240 residues and a hydrophobic π-π stacked with His253. A T-shaped π-stacking with His212 was also visualized. Simultaneously, π-alkyl interaction occurred between the alkyl group of His253 and Ala254 and the compound's aromatic rings. The two aromatic rings of the compound also made double T-shaped π-stacking with Tyr313 along with Zn 401. These interactions with zinc ion that occurs inside the active pocket might promote stable substrate binding with A3B.

On the other hand, the aromatic rings of compound RAL made a hydrophobic π-π stacking interaction with Tyr313, π-σ with His253, and T-shaped π-stacking with Trp287 and His212 (Fig. 4iii). H-bond interactions of the NH with Tyr313, His212, and His253 were also observed. Consequently, π-sulfur interaction was seen between the aromatic ring attaching to Tyr313 and the sulfur atom's lone pair of electrons. Besides, π-alkyl interaction was observed between the alkyl group of A3B and the aromatic ring of the compound bonding with Trp287.

ATA formed a hydrophobic π-π stacking interaction with Tyr315. The aromatic ring established T-shaped π-stacking with Tyr313. A conventional hydrogen bonding interaction with side chain residues of Asn240 and His253 at oxygen atoms in the ortho position was observed. These binding modes indicated the stability of the compounds to interact well within the receptor (Fig. 4iv).
The 2D and 3D ligand-A3B interaction analysis using discovery studio 2.5. From top to bottom: i. Nordracorubin; ii. NSC641233; iii. RAL, and; iv. Aurantricarboxylic acid (ATA). The interactions of hydrogen bonds, π-π stacking, and zinc ion between the ligands and amino acid residues are shown as green, magenta, and gold dashed lines, respectively.
The anticancer activities of these compounds have been mentioned elsewhere. Nordacorubin can be found in the fruit part of *Daenonrops draco*, commonly known as dragon-blood palm. The name "Dragon’s blood" refers to a reddish resinous product of the plant. The plant extract has been widely used as a traditional medicine in many cultures as a hemostatic, anti-diabetic, anti-cancer, anti-inflammatory, and wound healing agent [22]. Diverse range of secondary metabolites have been identified and isolated from the reddish resin, including flavones, flavanes, homoisoflavonanes, chalcones, dihydrochalcones, stilbenes, and flavonoid oligomers [23-25]. Nordacorubin is derived from the dimerization of monomeric flavonoids within the plant [26]. The compound has mild antimicrobial and antiplatelet properties. This compound’s derivatives, especially Dracorhodin perchlorate (DP), a synthetic analogue of Dracorhodin, demonstrated tremendous antiproliferation activity against human prostate cancer, melanoma, and leukemia cancer cells [27, 28]. The apoptotic activities of DP on several human cancer cells, especially HeLa and MCF-7 cells, are also reported [29, 30]. Despite the wide usage of *Daenonrops draco*, little research has been done on the biological effects of other active compounds from Dragon’s blood.

A synthetic NSC641233 compound was found in the DTP-NCI Mechanistic Set V. The set consists of 811 compounds derived from 37,836 open compounds that have been tested in the NCI human tumor 60 cell line screen [10]. Our team performed a systematic search from several databases only to discover limited literature was reported on this compound. This compound exhibited a strong binding energy value of -9.5kcal/mol, resulting in a stable complexation with the A3B enzyme. Interestingly, the compound bound to the A3B-competitive active site specifically with Zn ion, which is essential for deactivating the enzymatic activities of the A3 deaminase family.

Compound RAL, a selective benzothiophene estrogen receptor modulator, has been approved by the Food and Drug Administration (FDA) to treat and prevent osteoporosis in postmenopausal women. It was also used to reinforce the risk of breast cancer in those at high risk. RAL acts like estrogen to prevent bone loss and improve lipid profiles by decreases total and LDL cholesterol. It can denote the proliferation and progression of estrogen-dependent breast and uterine cancer [32]. RAL could significantly sensitize estrogen receptor multidrug-resistant (MDR) breast tumors to paclitaxel both in vitro and in vivo. A combination of RAL and paclitaxel could significantly enhance paclitaxel-induced cell apoptosis, G2-M arrest, as well as inhibition of cell proliferation in MDR tumors [32]. These observations indicated that RAL might be a vital chemosensitizer drug for breast cancer therapy. Furthermore, a study reported the combination of RAL with trichostatin A (TSA) inhibits E2-induced proliferation of MCF-7 breast cancer cells in a dose-responsive and time-dependent manner [33]. Several RAL derivatives have also exhibited prominent cytotoxic activities against cancer cell lines by inhibiting nucleotide pyrophosphatase/phosphodiesterase-1 and -3 enzymes [34].

**CONCLUSION**

This study discovered three prospective inhibitory compounds against A3B enzyme activity using Autodock/Vina, and GOLD. Compound Nordacorubin, RAL and NSC641233 showed the best binding energy among the 2900 docked compounds, using ATA binding energy as the reference cut-off point. These compounds binding to His253 amino acid residue and obligatory zinc ion within the A3B active pocket could potentially perturb the enzyme deamination activity on cytosine. Shi and coworkers [2] have demonstrated that the target cytosine on ssDNA was deeply inserted into the Zn-coordinating active site pocket in the co-crystal structures A3A-ssDNA and A3Bctd-ssDNA. Hence, we believe that A3B deaminase activity on ssDNA cytosine could be prevented due to the compounds’ occupation in the active catalytic site. Validation of compounds inhibitory activities using a specifically designed in vitro Fluorescence Resonance Energy Transfer (FRET) assay is currently on-going in our laboratory.

**ABBREVIATION**

A3; APOBEC3; A3B; APOBEC3B; AID; Activation-induced deaminase; APOBEC. Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; ATA; Aurintricarboxylic acid; BLAST: Basic Local Alignment Search Tool; CTD: Catalytic-terminal domain; DTP: Dracorhodin perchlorate; DTP: Developmental Therapeutics Program; FDA: Food and Drug Administration; FRET: Fluorescence Resonance Energy Transfer; HBV: Hepatitis B virus; HIV: Human Immunodeficiency virus; HTLV: Human T-lymphotropic virus; MD: Molecular docking; MDR: Multidrug-resistant; NCi: National Cancer Institute; NSC: National Service Center number; NTD: Amine-terminus, NHz-terminus or N-terminal end; PDB: Protein Data Bank; PLP: Piecwise Linear Potential; PyVs: Polyomavirus; ssDNA: single-stranded DNA; Tsa: Trichostatin A.

**ACKNOWLEDGMENT**

This work was supported by the BESTARI Research Grant (600-IRMI/PD/RD/5/3 BESTARI (072/2018)) from the Institute of Research Management and Innovation (IRMI), Universiti Teknologi MARA. We would like to appreciate and thank the Atta-ur-Rahman Institute for Natural Product Discovery and Institute of Science (IOS) for providing us with the computational resources, especially the GOLD docking program.

**FUNDING**

Nil

**AUTHORS CONTRIBUTIONS**

All authors have contributed equally.

**CONFLICT OF INTERESTS**

The authors reported no potential conflict of interest.

**REFERENCES**

1. Rfsland EW, Harris RS. The APOBEC3 family of retroelement restriction factors eric. Curr Top Microbiol Immunol 2013;371:1-27.

2. Shi K, Carpenter MA, Banerjee S, Shaban NM, Kurahashi K, Salamango DJ, et al. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. Nat Struct Mol Biol 2017;24:131-9.

3. Salter JD, Bennett RP, Smith HC. The APOBEC protein family: united by structure, divergent in function. Trends Biochem Sci 2016;41:578-94.

4. Alexandrò IB, Nik Zainal S, Wedge D. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.

5. Burns MB, Temiz NA, Harris RS. Evidence for APOBEC3B mutagenesis in multiple human cancers. Nat Genet 2013;45:977-83.

6. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, et al. An APOBEC3B cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 2013;45:970-6.

7. Siewuerts AM, Willis S, Burns MB, Look MP, Meijer Van Gelder ME, Schlicker A, et al. Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. Horm Cancer 2014;5:405-13.

8. Harris RS. Molecular mechanism and clinical impact of APOBEC3B-catalyzed mutagenesis in breast cancer. Cancer Res 2015;75:17:8.

9. Law EK, Siewuerts AM, Lapara K, Leonard B, Starrett GJ, Molan ME, Schlicker A, Wedge D, Stojanov P, et al. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. Sci Adv 2016;2:a6101737.

10. Choowongkomon K, Sawatdichaikul O, Songtawee N, Limtrakul J. Receptor-based virtual screening of EGFR kinase inhibitors from the NCI diversity database. Molecules 2010;15:4041-54.

11. Zhou X, Yu S, Su J, Sun L. Computational study on new natural compound inhibitors of pyruvate dehydrogenase kinases. Int J Mol Sci 2016;17:340.

12. Matsutomo T, Shirakawa K, Yokoyama M. Protein kinase a inhibits tumor mutator APOBEC3B through phosphorylation. Sci Rep 2019;9:8307.

13. Trott O, Olson AJ. AutoDock vina: improving the speed and accuracy of docking. J Comput Chem 2010;31:455-61.
