Journal of Molecular Graphics and Modelling 68 (2016) 128–139

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Identification of phenoxyacetamide derivatives as novel DOT1L inhibitors via docking screening and molecular dynamics simulation

Minghao Luo[a], Hui Wang[a], Yi Zou[a], Shengping Zhang[b], Jianhu Xiao[a], Guangde Jiang[c], Yihua Zhang[a], Yisheng Lai[a,*]

[a] State Key Laboratory of Natural Medicines, Jiangsu Key Laboratory of Drug Discovery for Metabolic Diseases, Center of Drug Discovery, China Pharmaceutical University, Nanjing 210009, China
[b] School of Chemical Sciences, The University of Auckland, 23 Symonds St., Auckland 1142, New Zealand
[c] Department of Medicinal Chemistry, College of Pharmacy, University of Florida, 1345 Center Drive, Gainesville, FL 32610, USA

ARTICLE INFO

Article history:
Received 16 March 2016
Received in revised form 16 June 2016
Accepted 17 June 2016
Available online 18 June 2016

Keywords:
DOT1L inhibitor
Virtual screening
Docking
Molecular dynamic simulation
Binding free energy

ABSTRACT

Dot1-like protein (DOT1L) is a histone methyltransferase that has become a novel and promising target for acute leukemias bearing mixed lineage leukemia (MLL) gene rearrangements. In this study, a hierarchical docking-based virtual screening combined with molecular dynamic (MD) simulation was performed to identify DOT1L inhibitors with novel scaffolds. Consequently, 8 top-ranked hits were eventually identified and were further subjected to MD simulation. It was indicated that all hits could reach equilibrium with DOT1L in the MD simulation and further binding free energy calculations suggested that phenoxyacetamide-derived hits such as **L01, L03, L04** and **L05** exhibited remarkably higher binding affinity compared to other hits. Among them, **L03** showed both the lowest glide score (−12.281) and the most favorable binding free energy (−303.9 +/− 16.5 kJ/mol), thereby making it a promising lead for further optimization.

1. Introduction

MLL-rearranged (mixed lineage leukemia gene rearranged) leukemia, which is particularly common in infant acute leukemia and secondary acute myeloid leukemia, is caused by the mixed lineage leukemia (MLL) gene rearrangements at positions such as 11q23 [1,2]. Normally, the wild-type MLL gene expresses the MLL protein containing a histone methyltransferase domain (SET domain) and a position recognition domain (AT hooks domain) [3,4]. The SET domain catalyzes the methylation of lysine 4 of histone H3 (H3K4) at specific genes, while the AT hook domain leads the MLL protein to target genes by binding to the minor groove of AT-rich DNA regions [4,5]. After rearrangements, the SET domain is replaced by a fusion-partner domain that contains sequences from proteins AF4, AF9, AF10, and ENL which have high affinity with Dot1-like protein (DOT1L). As a result, the newly encoded MLL-fusion proteins would acquire the ability to recruit DOT1L to MLL target genes with the function of the AT hooks domain and the fusion-partner domain [5–13].

DOT1L is a methyltransferase that catalyzes the methylation of lysine 79 of histone H3 (H3K79) with S-adenosyl-L-methionine (SAM) as its cofactor and the methyl donor [14–16]. Normally, DOT1L plays a critical role in gene embryonic development and maintaining normal hematopoiesis as well as normal functions of heart and kidney [17]. However, when the MLL gene rearranged, the MLL-fusion protein recruitment of excessive DOT1L around MLL target genes will result in hypermethylation at H3K79, which further activates the overexpression of several leukemogenic genes including HOXA9 and MEIS1 [18–21]. These transcriptional changes then lead to phenotypic changes and eventually induce leukemogenesis [22,23]. These studies elucidate that aberrant DOT1L activity is critical in the pathogenesis of MLL-rearranged leukemia [24]. Hence, the inhibition of DOT1L enzymatic activity has become a novel therapeutic strategy for MLL patients bearing MLL gene rearrangements. Meanwhile, this strategy also provides a promising solution for the side effects and early relapse of the conventional clinical therapies including chemotherapy and allogeneic hematopoietic stem cell transplantation [25].

The studies related to DOT1L inhibitors have reached apex in the past five years, among which the research of Epizyme Inc. [26] is the most impressive one. They have developed a small

* Corresponding author at: State Key Laboratory of Natural Medicines, Jiangsu Key Laboratory of Drug Discovery for Metabolic Diseases, Center of Drug Discovery, China Pharmaceutical University, 24 Tongjiaxiang, Nanjing 210009, China.
E-mail address: yslai@cpu.edu.cn (Y. Lai).

http://dx.doi.org/10.1016/j.jmgm.2016.06.011
1093-3263/

molecule inhibitor EPZ-5676, which is currently the first and only inhibitor that has entered clinical trials for the treatment of MLL-rearranged leukemia [27]. EPZ-5676, together with hundreds of other inhibitors developed by Epizyme, was identified by the classic ligand-based design and the following painstaking work of optimization [28–30]. As a result, all inhibitors derived from the S-adenosyl-L-homocysteine (SAH) contain the adenosine-like moiety which can be easily targeted and decomposed by endogenous enzymes. Therefore, these adenosine-containing inhibitors generally exhibit poor pharmacokinetics, such as a short plasma half-life in animal studies [31]. The metabolic instability further prevents the application of the conventional dosing method including intraperitoneal injection and oral administration. Hence, in preclinical and clinical trials, rats and leukemia patients were administrated with EPZ-5676 by continuous Intravenous (IV) infusion for a long dosing time (about 10–28 days) at a relative high dosage to present anti-leukemic function [31–33]. Considering the monotonous structural type, the poor pharmacokinetics and the defective dosing method of adenosine-containing inhibitors, there is an urgent need for discovering novel DOT1L inhibitors.

Virtual screening (VS) technologies, such as pharmacophore modeling and molecular docking simulation, have played an effective and economic role in the development of therapeutically important small molecules over decades [34–36]. In order to discover promising scaffolds that will be useful for developing novel DOT1L inhibitors, we proposed a structure-based VS strategy (Fig. 1) using molecular docking combined with molecular dynamic (MD) simulation based on a DOT1L co-crystal 4EQZ. Firstly, Specs, a small molecular database [37] was filtered by Lipinski's rule of five and the ADMET properties filters. Compounds passing through the drug-like queries entered a cascade docking screening. To determine the key residues that contribute the most to the binding, we analyzed the binding mode of the inhibitor EPZ-5676 in complex with DOT1L and found that it bound with DOT1L by forming hydrogen bonds with Asp161, Glu186, Asp222 and Phe223. Then we individually mutated these residues into alanines and calculated the binding free energy of EPZ-5676 in complex with the wild-type DOT1L and its mutants. The differences of binding free energies between the wild type and the mutants indicated that Asp161 contributed the most to the binding compared to other residues. Those compounds that had a hydrogen bond with the key residue Asp161 were then extracted. By clustering, the top-ranked hits were put into MD simulation and binding free energy calculation. The obtained results showed that phenoxyacetamide derivatives, such as **L01**, **L03**, **L04** and **L05** (especially **L03**), might have good inhibitory activity towards DOT1L.

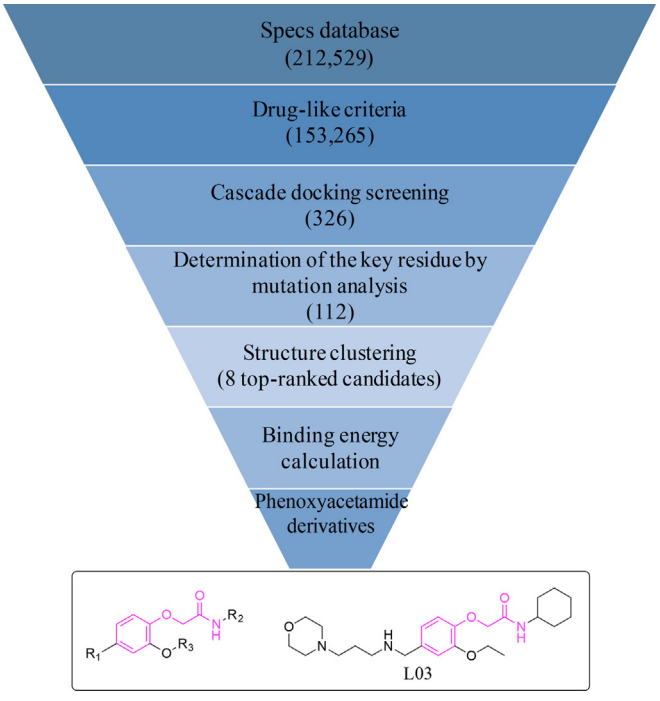

Fig. 1. Flowchart of the VS strategy adopted in this study.

2. Materials and methods

2.1. Docking software selection

In order to perform a valid docking screening, the docking algorithm that best suited our own DOT1L system should be identified firstly. Some most commonly used docking algorithms including Glide (HTVS), Glide (SP), Glide (XP) [38], Gold [39], Surflex [40], CDocker [41] were evaluated by conducting the redocking analysis. Currently, there are 19 co-crystals of DOT1L in complex with different ligands available in the RCSB Protein Data Bank (PDB) [42]. We chose nine co-crystals (PDB codes: 4HRA, 4EKG, 4EKI, 4EK9, 4EQZ, 4ER0, 4ER3, 4ER6, 4ER7) in complex with inhibitors containing adenosine-like moiety for redocking analysis. Inhibitors within the nine co-crystals were extracted and redocked into the corresponding protein structures using the six different docking algorithms, respectively. Then the root-mean-square deviation (RMSD) values between the generated poses and the corresponding original pose in the crystal were calculated.

2.2. Decoy set testing

After selecting the docking algorithm, we tried to identify the suitable co-crystal that will be used as the docking receptor. A decoy set comprising of 68 active DOT1L inhibitors with Ki value ranging from 0.5 nm to 50 μm and 1500 inactive compounds selected from the ZINC [43] database was built to investigate the nine co-crystals. The following protocol was applied in inactive molecules selection using packages from Discovery Studio 2.5 [44]. The *Find Similar Molecules by Fingerprints* protocol was used to calculate the Tanimoto similarity [45] between Zinc database molecules and the 68 active compounds using the extended-connectivity fingerprints (FCFP_6) [46]. Compounds with the scaled similarity above 0.2 were discarded. The remaining molecules were subjected to the *Find Similar Molecules by Number property* protocol to identify compounds that had similar physical properties with the active molecules. The number of similar molecules was set to 20,000. This library was further subjected to the *Find Diverse Molecules* protocol to obtain 1500 inactive molecules. This decoy set was docked into the nine co-crystals using Glide (SP). Then, the glide scores of the 68 top-ranked compounds obtained from the nine docking solutions were analyzed, and a receiver operating characteristics (ROC) curve [47] was plotted to identify the co-crystal that had the best performance to distinguish the active molecules and inactive molecules from the decoys. The curve was created by plotting the sensitivity (the true positive rate) against the 1-specificity (the false positive rate). The discrimination power of certain docking algorithm with the corresponding co-crystal was determined by the area under ROC curve (AUC). In one extreme situation, an ideal co-crystal which ranks all the actives over any decoys via certain docking algorithm would have the largest AUC, while in the another extreme, the ROC curve of a randomly ranking system would be a diagonal in which the AUC value is 0.5. Thus, co-crystal with the largest AUC value (above 0.5) was chosen as the docking receptor. After selecting the suitable co-crystal, the discrimination power of Glide (HTVS), Glide (SP) and