Journal of the Taiwan Institute of Chemical Engineers xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of the Taiwan Institute of Chemical Engineers

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

The discovery of potential tubulin inhibitors: A combination of pharmacophore modeling, virtual screening, and molecular docking studies

Miaomiao Niu, Ke Wang, Congying Zhang, Yaru Dong, Guissi Fida, Xue Dong, Jiyu Chen, Yueqing Gu *

Department of Biomedical Engineering, School of Life Science and Technology, State Key Laboratory of Natural Medicines, China Pharmaceutical University, 24 Tongjia Lane, Nanjing 210009, China

ARTICLE INFO

Article history:
Received 2 May 2014
Received in revised form 26 June 2014
Accepted 23 July 2014
Available online xxx

Keywords:
Tubulin
Pharmacophore modeling
Virtual screening
Molecular docking

ABSTRACT

The microtubules play a significant role in the growth and functions of cells, and hence, inhibition of microtubule assembly has emerged as one of the most promising strategies for the treatment of cancer. To clarify the essential structure-activity relationship of tubulin inhibitors as well as identify novel leads, three-dimensional pharmacophore models were established on the basis of known tubulin inhibitors. The model (Hypo1) that had the highest correlation coefficient (0.9330), the largest cost difference (66.911) and the lowest RMSD value (0.70258) was chosen as the best pharmacophore model. Hypo1 was validated using test set, Fischer randomization and leave-one-out methods and a decoy set. Then, the model was used as a 3D structural query to search the Specs database. We subsequently refined the retrieved hit compounds by subjecting them to drug-like filtrations and molecular docking studies. Finally, we selected 10 compounds on the basis of the necessary interactions with the key residues to be experimentally validated in a bioassay. We showed that all compounds had some level of inhibitory effect on MCF-7 breast cancer cell line where the most promising three candidates exhibited an inhibition rate of more than 80% at a concentration of 100 μmol/L.

1. Introduction

The microtubule system of eukaryotic cells plays a very important role in regulating cell architecture, and it has a crucial role in cell division, since microtubules are a key component of the mitotic spindle [1]. Microtubules as a target for anticancer drugs are involved in many essential cellular processes, such as maintenance of cellular shape, regulation of motility, transport of material within the cell and cell signaling [1,2].

Antimitotic agents arrest the cell cycle at the G_2/M phase and result in tumor regression and apoptotic cell death [3–5]. Classic tubulin-binding agents including taxanes and vinca alkaloids, are widely used to treat human cancers by interfere with microtubule dynamics [5]. However, due to neurotoxicity, difficult synthesis and drug resistance, their clinical use has been limited [5,6]. Therefore, it is very important to develop novel tubulin-binding inhibitors with novel modes of action [5,7,8].

There has long been considerable interest in the discovery and development of novel inhibitors able to interfere with tubulin polymerization [9–11]. In recent years, a variety of small molecules have been reported as inhibitors of tubulin polymerization that bind to the colchicine-binding site of tubulin [12–15]. Though many different tubulin inhibitors had been synthesized and experimentally assessed, there is no information available regarding the discovery for structurally novel leads up to date. The development of such kind of drug is focused on how to design novel tubulin inhibitors.

In our study, a combination of pharmacophore modeling, virtual screening, and molecular docking studies served as a guide aiming to identify potential leads with antitumor activities. A high-correlation quantitative pharmacophore model was generated using the observed structure-activity relationship of known tubulin inhibitors. After validation, this pharmacophore model is used as a 3D structural search query to find new classes of

* Corresponding author. Tel.: +86 25 83271080; fax: +86 25 83271046.
E-mail addresses: 295248209@qq.com (K. Wang), djsld999@163.com (C. Zhang), zhuxituan00@126.com (Y. Dong), eons99d@126.com (G. Fida), woshioo4@126.com (X. Dong), 474962426@qq.com (J. Chen), guyueqingsubmission@hotmail.com, niumm2012@126.com (Y. Gu).

http://dx.doi.org/10.1016/j.jtice.2014.07.016
1876-1070/

compounds from the Specs database. The hit compounds were further subjected to drug-like filtrations and molecular docking studies to refine the retrieved hits. The binding free energy and molecular interactions with the active site residues were considered as important components to identify the potential inhibitory leads. Finally, potential inhibitory leads showing strong molecular interactions with the key amino acids were proposed as potential leads.

2. Materials and methods

2.1. Pharmacophore model generation

HypoGen module of Discovery Studio v2.5 (Accelrys, San Diego, USA) was used to perform all pharmacophore modeling calculations. The accurate choice of the training set is a key issue in a computerized pharmacophore generation process. The built pharmacophore hypothesis can be as good as the input data information. Based on the principle of structural diversity and wide coverage of activity range, a total of 62 compounds from the literature resources [1,2,5,9,16–18] were selected to be used as primary data set in 3D QSAR pharmacophore modeling study. To ensure the statistical relevance, a training set containing 31 diverse compounds with the experimental activity values (IC_{50}) were selected from 62 dataset compounds and used as training set and the remaining 31 compounds were used as test set compounds to be utilized in pharmacophore validation. The following criteria have been considered during the selection of data set in order to achieve a significant pharmacophore hypothesis. (1) All 62 compounds with inhibitory activity against CEM cancer cell lines have been shown to bind at the colchicine site; (2) The data set must be widely populated covering an activity range of at least 4 orders of magnitude. The inhibitory activity values of the training set compounds span over a range of five orders of magnitude, from 0.52 nM to 13,800 nM. The activity values of test set compounds have a range of four orders of magnitude, from 2.8 nM to 14,900 nM; (3) All inhibitory activity of 62 compounds used in the data set is from same wet-lab assay. These compounds were built and subsequently minimized to the closest local minimum based on the Charmm-like force field (DS). All 31 compounds in the training set were submitted to 3D QSAR pharmacophore generation of DS. Best conformer generation option, a maximum number of 250, and an energy threshold of 10 kcal/mol above the energy minimum for conformation searching were selected to generate multiple conformations. Hydrogen bond donor (HBD), Hydrogen bond acceptor (HBA), hydrophobic (HY) and ring aromatic (RA) features were used to generate 10 pharmacophore models [19,20]. In this study, the top 10 hypotheses returned by the generation process were selected for further calculation.

The quality of a pharmacophore model is mainly determined by two theoretical cost calculations that are represented in bit units. One is the 'null cost' that represents the highest cost of a pharmacophore model with no features which estimates every activity to be the average of the activity data of the training set molecules. The second cost is the 'fixed cost' also known as cost of an ideal model, which represents the simplest model that fits all the data perfectly. The total cost of any model should always be away from the null cost and close to the fixed cost to be the meaningful model. The cost difference between null and fixed cost values should be larger for a significant pharmacophore model. A value of 40–60 bits in a model implies that it has 75–90% probability of representing a true correlation in the data [19–21]. The hypotheses are also evaluated based on other cost components. The cost value for every hypothesis is the summation of the weight cost (W), the configuration cost (C) and the error cost (E). The weight cost is a value that increases in a Gaussian form as the feature weights in a model deviate from the ideal value of two. The configuration cost measures the entropy of the hypothesis space. The error cost is the value represents the root-mean-squared difference (RMSD) between estimated and experimental activity value of the training set compounds. If the input training set compounds are too multiplex, due to too flexible training set molecules, this will lead to an effusive number of hypotheses as an outcome of the subtractive phase. This configuration cost should always be less than a maximum value of 17. The correlation coefficient of the pharmacophore model should be close to 1.

2.2. Pharmacophore model evaluation

Pharmacophore model evaluation are provided as Supplementary material.

2.3. Virtual screening

To identify novel hit compounds, the validated pharmacophore model was used as 3D query to screen the Specs database (2,02,919). The two-dimensional (2D) chemical structures of 2,02,919 compounds in the chemical database were converted to 3D structures in Discovery Studio v2.5 (DS). The Ligand Pharmacophore Mapping protocol was applied in database screening [19,20]. The screened compounds were further filtered based on estimated activity and Lipinski's rule of five. A Lipinski-positive compound has (i) a molecular weight < 500; (ii) <5 hydrogen bond donor groups; (iii) <10 hydrogen bond acceptor groups and (iv) an octanol/water partition coefficient ($\log P$) value < 5 [19–21].

2.4. Molecular docking

Molecular docking are provided as Supplementary material.

2.5. Cell proliferation inhibition assay

The biological assays were performed using the MTT assay against human breast cancer cell line (MCF-7). The cell line was cultured in DMEM medium supplemented with 10% fetal bovine serum. For *in vitro* treatment, the carcinoma cells were seeded in 96 well plates and incubated in incubator at 37 °C and 5% CO_2. After 24 h, the cells were treated with the concentration of each test compound for 48 h. At the end of the drug exposure period, the cells were incubated at 37 °C for 4 h to 6 h by adding 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Sigma) (20 μL/well). Then the medium was removed, and DMSO was added to the insoluble fraction. The absorbance values at a wave-length of 490 nm were determined with the Spectramax M5 Microtiter Plate Luminometer (Molecular Devices, USA). Values were calculated using percentage of growth *versus* untreated control.

3. Results and discussion

3.1. Pharmacophore modeling

To quantitatively correlate the chemical structure of tubulin inhibitors to their biological activity, the HypoGen algorithm available in 3D QSAR Pharmacophore Generation protocol of DS was carried out. In pharmacophore model generation, the training set of 31 compounds (Fig. 1) with activity values ranging from 0.52 to 13,800 nM was used to generate 10 top-scored pharmacophore hypotheses. The results of the top 10 pharmacophore hypotheses and their statistical parameters are shown in Table 1. In this study, the first pharmacophore hypothesis (Hypo1) is the best one characterized by the lowest total cost value (123.717), the largest

Please cite this article in press as: Niu M, et al. The discovery of potential tubulin inhibitors: A combination of pharmacophore modeling, virtual screening, and molecular docking studies. J Taiwan Inst Chem Eng (2014), http://dx.doi.org/10.1016/j.jtice.2014.07.016