A combination of pharmacophore modeling, molecular docking and virtual screening for NPC1L1 receptor inhibitors from Chinese herbs

Xiaoqian Huo, Ludi Jiang, Xi Chen, Yusu He, Yongqiang Yang, Yanling Zhang*
School of Chinese Pharmacy
Beijing University of Chinese Medicine
Beijing, 100102, China
aixue_126@126.com

***Abstract*—NPC1L1, a protein localized in jejunal enterocytes, is critical for cholesterol absorption. As the receptor inhibitors are effective solutions for hyperlipidaemia, NPC1L1 receptor is becoming a hot spot in drug targets. In this study, pharmacophore modeling and molecular docking were combined to discover potential NPC1L1 inhibitors from traditional Chinese medicine. The best pharmacophore model, Hypo1, which was generated by 9 known inhibitors, comprised of two Hydrogen bond acceptor lipid and two Hydrophobic aromatic regions. And the active compounds hit rate (A%), identification index (N), and comprehensive evaluation index (CAI) are 100%, 3.852, and 3.852 respectively. Hypo1 was used to screen TCMD (version 2009) to identify potential inhibitors, which resulted in a hit list of 38 compounds with Lipinski's rule of five. In addition, docking was used to refine pharmacophore-based screening results by using ezetimibe as a reference. Then, 11 compounds with higher docking score than ezetimibe had been reserved. This paper provides a reliable utility for discovering natural NPC1L1 receptor inhibitors from traditional Chinese herbs.**

Keywords—NPC1L1 receptor inhibitors; pharmacophore; docking; virtual screening; Traditional Chinese Medicine

I. INTRODUCTION

Hyperlipidemia, a medical condition, is characterized by excessive amounts of fatty substances such as cholesterol and triglycerides in the blood of an individual. Along with its complication, hyperlipidemia has become a health-threatening issue. Nowadays, the marketed lipid-lowing drugs have been developed by acting on various targets to exert therapeutic action. For instance, ezetimibe, as a representative inhibitor of NPC1L1, can decrease cholesterol absorption in the small intestine [1], while statins lower cholesterol levels by inhibiting HMG-CoA reductase to control the synthesis of endogenous cholesterol. Ballantyne CM indicated that treatment that combines NPC1L1 inhibitors and HMG-CoA reductase inhibitors appears to be superior to treatment with either drug alone because of the complementary functional mechanism [2].

The NPC1L1 protein, mainly found on the small intestine and liver [3], is a vital factor for cholesterol absorption which made it become a hot drug target for the treatment of Hyperlipidemia. Therefore, the more selective and secure inhibitors of NPC1L1 are urgently needed to be designed and discovered for the treatment.

The 3D-pharmacophore model and molecular docking, serving as two basic virtual screening techniques in drug discovery, can mitigate the high-cost and time-consuming problems which caused by traditional drug design methods. Ligand-based pharmacophore model is an abstract description of the essential features and corresponding 3D locations of the chemical structures, which can explain how structurally diverse ligands exert similar biological activities. Through the use of pharmacophore for virtual screening, molecules which may act on the same receptor can be screened from a database. Molecular docking is a method which predicts the binding orientation of candidate compounds to their protein targets in order to in turn evaluate the affinity and activity of the ligands. It is always utilized for screening the initial results further by providing more structural information and the pattern of receptor recognizing ligands.

The task of this study is to build the pharmacophore models for NPC1L1 receptor inhibitors. HipHop, one of the classic algorithms of ligand-based pharmocophore, was selected to construct qualitative models. By using several validation methods of pharmacophore, the best pharmacophore was selected as a template to screen potential NPC1L1 inhibitors from the Traditional Chinese Medicine Database (TCMD, version 2009) [4]. In addition, with the purpose of making up the missing information of receptor and reducing the false positive rate of the screening result, the structure of NPC1L1 protein (PDB ID: 3QNT) was used to establish the docking model.

II. MATERIALS AND METHODS

A. HipHop pharmacophore model generation

The ligand-based pharmacophore was built by HipHop (Common Feature Pharmacophore Generation) within Accelrys Discovery Studio 4.0(DS).In this study, "NPC1L1 receptor inhibitors" was used as a search term to select NPC1L1 receptor inhibitors from the Binding Database (http://www.bindingdb.org/). Then, 87 compounds were obtained as a universal set. In order to ensure the compounds of the training set had relatively good representative of activity,

This work was supported by the National Natural Science Foundation of China (No. 81173522). *Correspondence author: Y. Zhang (collean_zhang@163.com)

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

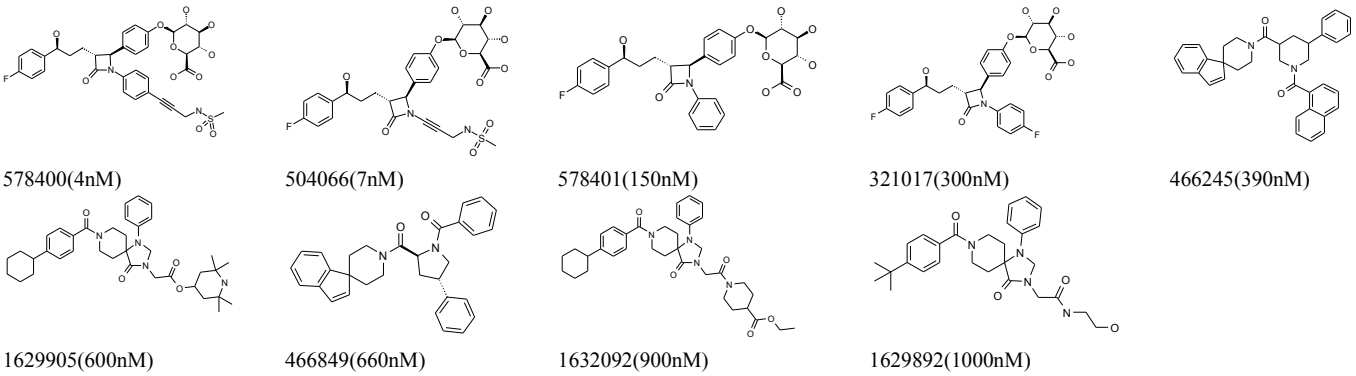


Fig. 1. Structures of all 9 training set compounds employed in pharmacophore generation. IC50 values were listed in parentheses.

compounds. Satisfied these two requirements, the training set can have a representative of all the active compounds. After that, the remaining 78 active compounds were regarded as test set. The information of the biological activities and structures of the training set is showed in Figure1.

All the compounds were constructed and minimized by DS. Diverse conformations were created by the BEST mode within the relative energy threshold of 20kcal/mol. And maximum number of conformations was set to 255. In addition, an initial analysis of the training set demonstrated that five features, such as Hydrogen bond acceptor(A), Hydrogen bond acceptor lipid(ALi), Hydrogen bond donor(D), Hydrophobic region(H), Hydrophobic aromatic region(Haro), Ring aromatic(R), well mapped all the training set ligands.

In order to produce a Hiphop pharmocophore model, Principal values of compounds in the training set were set to 2 to indicate the high activity, and set MaxOmitFeat value to 0 to show that there is no feature should be ignored. According to the initial analysis of training set, A (0-5), D (0-5), H (0-5), N (0-5), and R (0-5) features were selected during the pharmacophore construction. Besides, Maximum Excluded Volumes (Ev) value was set to 5.And all other parameters were automatically set to the default value.

B. Validation of the pharamacophore model

In this study, in order to identify features which were commonly existed in highly active NPC1L1 inhibitors, a test set was used to evaluate all the pharmacophore models. And the evaluation indicators were presented as follow: Rank score, A%, N, and CAI (Comprehensive Appraisal Index) [5]. Then, by considering all these factors, the best pharmacophore model was selected as TCMD query.

C. Database search

The optimal pharmacophore model was utilized for screening potential NPC1L1 inhibitors from TCMD by using the BEST search method. 233033 compounds from 6735 natural plants and animals were contained in TCMD. The compounds which map all the pharmaphore features were hit. And then they were filtered by Lipinski's rule (≥3), in order to be further analyzed by docking studies.

D. Molecular docking studies

3QNT, a structure of NPC1L1 receptor, was used as the reference model, common problems of which were cleaned up by DS. Define and Edit Binding Site tools were applied to identify and edit binding sites of receptor. This study located binding sites automatically according to Receptor Cavities because there is no inhibitor binding to the crystal structure. Meanwhile, all the ligands used for docking studies were minimized by CHARMm forcefield and were changed the ionization state in PH value of 6.5.

The training set was docked into each binding pocket by using three algorithms, LibDock, LigandFit and CDOCKER, respectively, to determine the preferable docking module. The binding site which can accept maximum active compounds in training set may be the potential binding pocket. And the corresponding algorithms would be the optimal one which is suitable for this docking system. After that, the selected binding site and docking algorithm would be applied to evaluate the hits.

III. RESULTS AND ANALYSIS

A. Pharmacophore model generation

The top 10 pharmacophore models with high Rank scores (79.83-95.83) were shown in Table1. From Table1, the main features contained Hydrogen bond acceptor(A), Hydrogen bond acceptor lipid(ALi), Hydrogen bond donor(D), Hydrophobic region(H), Hydrophobic aromatic region(Haro), Ring aromatic(R). These ten pharmacophores contain all the kinds of feature. One pharmacophore may include two or three tapes of features. To select the best hypothesis model, these models were compared by using the evaluation indicators. The ligand-based pharmacophore show the features that training molecules have in common.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

B. *Validation of the pharamacophore model*

78 active compounds and 234 inactive compounds constituted the test set for validate the pharmacophore model. The result was shown in Table2. The Rank score, A%, N and CAI were used to validate the pharmacophores. The Rankscore was the methods in DS to evaluate the pharmacophore. A% was used to assess the ability of distinguish the active compounds. N shows the capacity of telling active compounds from inactive compounds. In this study we proposed the CAI to compressively evaluate the resolution of pharmacophore. We used the value 2, 1.5 as a criterion value to evaluate the N, CAI. The pharmacophore excluded the hypo4, 5 get a high N value. All the pharmacophore pass the CAI cut-off scores. The pharmacophores get the highest rankscore, A%, N and the CAI, it shows that they are quality pharmcophores in differentiating the active compounds.

These pharmacophores all have a decent randscore, N and CAI. Thinking of the four methods is In order to impersonally appraise all 3 factors to select the best pharmacophore for our virtual search, ranking values were assigned to the observations of the 4 factors (Table3). And Σ Ranking was used to calculate a summation of these ordinal numbers from the 4 factors. In present study, for the models was sorted by the ordinal numbers, a lower Σ Ranking value indicated a more favorable model we obtained. With this strategy, Hypo1 was chosen as the best pharmacophore model (Fig. 2) The Hypo1 contains two Hydrophobic aromatic region features and two Hydrogen bond acceptor lipid features.

C. *Database search*

By using the best pharmacophore model, a hit list of 1643 compounds was obtained. Then, the hits were filtered by Lipinski's rule (≥3), and 38 compounds were reserved. Among them, Salvianolic acid C, derived from Salvianolic, had higher estimated activity. The Salvianolic is the major ingredient of Danshen Jiangzhi Yin. Meanwhile, Danshen Jiangzhi Yin could reduce total cholesterol (TC) in serum [6]. The Salvianolic acid C got 3.8364 fitvalue which explained that the compounds can be well fit of the pharmacophore. The benzene rings were fit into the Haro features. And two hydroxies were fit into the Ali feature. The mapping of the feature was showed in the fig 3.

TABLE1 THE CALCULATING RESULTS OF PHARMACOPHORE MODELS

Hypo	HipHop Features	Rank
1	HaroHaroAliAli	95.83
2	RRAliAli	92.84
3	HaroHaroAA	91.60
4	RRAA	89.64
5	RRAA	83.30
6	HHAliAli	83.30
7	HHAliAli	80.10
8	HHAA	80.09
9	HHAA	80.06
10	HHAA	79.83

TABLE2 THE VALIDATION RESULTS OF PHARMACOPHORE MODELS

Hypo	D[a]	A[b]	Ha[c]	Ht[d]	A%[e]	N[f]	CAI[g]
1	312	78	78	81	100.00%	3.852	3.852
2	312	78	77	125	98.72%	2.464	2.432
3	312	78	75	137	96.15%	2.190	2.106
4	312	78	77	123	98.72%	2.504	2.472
5	312	78	75	135	96.15%	2.222	2.137
6	312	78	76	163	97.44%	0.466	1.817
7	312	78	77	161	98.72%	0.478	1.889
8	312	78	75	153	96.15%	1.961	1.885
9	312	78	77	126	98.72%	2.444	2.413
10	312	78	75	153	96.15%	1.961	1.885

[a]D is the number of all compounds the test database contains. [b]A is the number of active compounds in the test database. [c]Ha is the number of active hits though the pharmacophores search. [d]Ht is the number of all compounds hits though pharmacophores search. [e]A% identify the ability to identify active compounds from the test database. [f]N represents represents the ability to active compounds from inactive compounds. [g]CAI is Comprehensive Appraisal Index

TABLE 3 THE RESULTS OF RANKING

Hypo	Ranking of Rankscore[a]	Ranking of N[a]	Ranking of CAI[a]	∑Ranking[b]
1	1	1	1	3
2	2	3	3	8
3	3	6	6	15
4	4	2	2	8
5	5	5	5	15
6	6	9	9	24
7	7	8	7	22
8	8	7	8	23
9	9	4	4	17
10	10	8	8	26

[a]The 3 validation methods need to be comprehensively appraised. [b]The summation of ranking values based on the 3 factors ranking results.

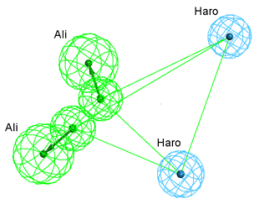


Fig.2 The best pharmacophore model

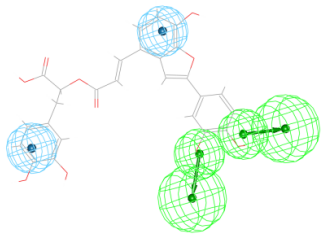


Fig3.The best pharmacophore model (Hypo1) mapped with Salvianolic acid C

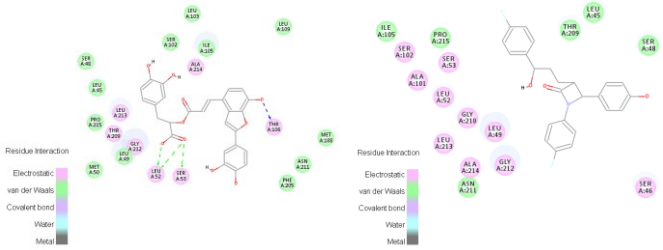
D. *Molecular docking*

In this study, 6 potential binding sites were automatically located. Then, three aforementioned docking algorithms were respectively used to calculate the training set in different binding sites.

The training set was used to docking into every site to find the active site in highest likelihood. By using LibDock, 7 active ligands docked into site 1 and other sites didn't accept any pose. All 6 sites didn't accept ligands by using

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

LigandFit. With the method of CDOCKER, just 1 ligand was accepted by site 1. Therefore, LibDock was better at analyzing the NPC1L1-ligand interactions and site 1 might be the active pocket of NPC1L1. According to the selection of LibDock algorithm, Libdocking score was used to estimate the binding affinity of a ligand. The binding pocket size was created with the sphere radius of 12.6Å. Then, 38 drug-like compounds were docked into site 1, resulting in a



(a)Salvianolic acid C (b)Ezetimibe

Fig.4. The interactions between the ligands ([a]Salvianolic acid C and [b]Ezetimibe) and the receptor. Pink circles represent the hydrogen-bond, charge or polar interactions, and green circles represent the other residues involved in van der Waals interactions.

hit list of 34 compounds. Only the site1 has compounds docked into, so we apply the site1 in the TCM docking.

Ezetimibe was used as a reference to appraise the hits. 11 compounds got higher score than ezetimibe (Libdocking score = 134.845). Among the top 11 compounds, Salvianolic acid C, Platyphyllonol-5-O- β -D-xylopyranoside and Globoidnan A achieved outstanding ranking in both pharmacophore and molecular docking, could be lead candidates for design of novel NPC1L1 inhibitors. The docking schematic diagram of Salvianolic acid C and ezetimibe were shown in Figure4. The higher score means the Salvianolic acid C can be better fit than the ezetimibe. So it can be in better biological activity.

Meanwhile, the Salvianolic acid C was also docking into the protein. The compounds which can fit the protein structure have more advantages in the application. From Figure4, the interaction residues of the Salvianolic acid C, such as LEU52, SER53, and THR106 were similar to ezetimibe. It can be speculated that these amino acid residues may play a key role in the interaction between NPC1L1 and inhibitors.

IV. CONCLUSION

Based on pharmacophore modeling and molecular docking, potential inhibitors of NPC1L1 were discovered from TCMD. More specifically, the qualitative pharmacophore models were generated from 9 NPC1L1 inhibitors by using HipHop. The best pharmacophore model Hypo1 was validated with the test set and used for the TCMD virtual screening. With the help of Lipinski's rules, the hits were filtered and further evaluated by molecular docking studies. Molecule docking verified the compounds from TCM that map the structure of the protein. The active compounds we got may be potential inhibition drug of NPC1L1. In summary, we expect this study could be applied in the research of potential natural NPC1L1 inhibitors.

ACKNOWLEDGMENT

At the point of finishing this paper, the authors would like to express sincere thanks to the National Natural Science Foundation of China (No. 81173522) in Beijing University of Chinese Medicine.

REFERENCES

[1] Altnann SW, Davis HR Jr, Zhu LJ, et Niemann-Pick C1 Like 1 (NPC1L1) is critical for intestinal cholesterol absorption [J]. Science,2004,303:1201-1204.

[2] Ballantyne CM, Abate N, Yuan Z, et Dose-comparison study of the combination of ezetimibe and simvastatin (Vytorin) versus atorvastatin in patients with hypercholesterolemia: the Vytorin Versus Atorvastatin (VYVA) study [J]. Am Heart, 2005 149:464-473.

[3] Davis HR Jr, Zhu LJ, Hoos LM, et, NPC1L1 is the intestinal phytosterol and cholesterol transporter and akey omdulator of whole body cholesterol homemstasis [J] BiolChem, 2004,279:33586-33592.

[4] Traditional Chinese Medicine Database (TCMD), Accelrys Inc., Beijing, Chineseedition,2009.Available:http://www.neotrident.com/product/detail.aspx.id=18.

[5] Yang Z, Zhang Y, Wang X, et al. Pharmacophore Model Generation of P2Y12 Inhibitor[C]. Biomedical Engineering and Biotechnology (iCBEB), 2012 International Conference on. IEEE, 2012: 396-399..

[6] Zhi W, Liqiang H, Shuang W, Fei G. Research on hypolipidemic effect and mechanism of reducing blood fat of compound Danshen Jiangzhi Yin[J]. Chinese Archives of Traditional Chinese Medicine, 2013, 31(12): 2659-2661.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE