Journal of Molecular Structure 1049 (2013) 315–325

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Structure

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

A combined pharmacophore modeling, 3D-QSAR and molecular docking study of substituted bicyclo-[3.3.0]oct-2-enes as liver receptor homolog-1 (LRH-1) agonists

Manisha Lalit [1], Rahul P. Gangwal [1], Gaurao V. Dhoke, Mangesh V. Damre, Kanchan Khandelwal, Abhay T. Sangamwar *

Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), Sect-67, S.A.S. Nagar, Punjab 160 062, India

HIGHLIGHTS

- Ligand based pharmacophore model was developed and validated for LRH-1 agonists.
- Three different databases were screened by means of validated pharmacophore model.
- CoMFA and CoMSIA models were developed and validated.
- Docking analysis shows that His390 and Arg393 plays important role in binding of agonists.

ARTICLE INFO

Article history:
Received 25 March 2013
Received in revised form 13 June 2013
Accepted 13 June 2013
Available online 22 June 2013

Keywords:
CoMFA
CoMSIA
Liver receptor homolog-1
Molecular docking
Pharmacophore
Virtual screening

ABSTRACT

A combined pharmacophore modelling, 3D-QSAR and molecular docking approach was employed to reveal structural and chemical features essential for the development of small molecules as LRH-1 agonists. The best HypoGen pharmacophore hypothesis (Hypo1) consists of one hydrogen-bond donor (HBD), two general hydrophobic (H), one hydrophobic aromatic (HYAr) and one hydrophobic aliphatic (HYA) feature. It has exhibited high correlation coefficient of 0.927, cost difference of 85.178 bit and low RMS value of 1.411. This pharmacophore hypothesis was cross-validated using test set, decoy set and Cat-Scramble methodology. Subsequently, validated pharmacophore hypothesis was used in the screening of small chemical databases. Further, 3D-QSAR models were developed based on the alignment obtained using substructure alignment. The best CoMFA and CoMSIA model has exhibited excellent r^2_{ncv} values of 0.991 and 0.987, and r^2_{cv} values of 0.767 and 0.703, respectively. CoMFA predicted r^2_{pred} of 0.87 and CoMSIA predicted r^2_{pred} of 0.78 showed that the predicted values were in good agreement with the experimental values. Molecular docking analysis reveals that π–π interaction with His390 and hydrogen bond interaction with His390/Arg393 is essential for LRH-1 agonistic activity. The results from pharmacophore modelling, 3D-QSAR and molecular docking are complementary to each other and could serve as a powerful tool for the discovery of potent small molecules as LRH-1 agonists.

1. Introduction

Nuclear receptors (NRs) are a class of proteins found inside the cells, which act as transcriptional factors. They regulate expression of many genes and control the development, homeostasis, and metabolism of the organism. Their activity is regulated by binding of small lipophilic compounds including hormones, metabolites and few synthetic ligands. NRs get activated upon ligand binding and/or phosphorylation causing a conformational change, which results in dissociation of co-repressor complexes and recruitment of co-activator complexes. According to sequence homology, NR superfamily is classified into seven subfamilies (NR0-NR6) [1]. NRs for which no ligand has been identified are categorized as orphan NRs. Liver receptor homolog-1 (LRH-1, NR5A2) is one such orphan NR, which belongs to NR5A subfamily. In adult mammals, it is mainly confined to liver, pancreas and intestine. It is also expressed in ovary, pre-adipocyte and at lower levels in placenta. It plays an important role in early development and also in regulation of bile acid synthesis, cholesterol metabolism and steroidogenesis in the adult. It also regulates the expression of aromatase in the breast and ovaries, which also exhibit its utility in cancer therapy [2–7].

LRH-1 was found to be constitutive active when expressed in a variety of cell types [8]. However, phospholipids have been found

* Corresponding author. Tel.: +91 0172 2214682.
E-mail address: abhays@niper.ac.in (A.T. Sangamwar).
[1] These authors contributed equally to this work.

0022-2860/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.molstruc.2013.06.035

in the ligand binding pockets of LRH-1 [9–11], but their role in regulation of LRH-1 activity still remains to be established. Two phosphatidylcholines, dilauryl and diundecadeoyl have been found as LRH-1 agonists and depicted LRH-1 as a target for type-2 diabetes [12]. Medicinal chemistry approach for the development of LRH-1 agonists has been utilized by Whitby et al. [13,14]. They have used cis-bicyclo-[3.3.0]oct-2-ene skeleton and designed various compounds of series 1-anilino, 1-alkoxy- and 1-alken-2-yl-substituted bicyclo-[3.3.0]oct-2-enes. Among the 1-anilino series, compound GSK8470 showed good pEC_{50}, but its primary limitation is acid instability. The compound RJW100 was found to have good activity and stability. In the cell-based studies, it is found to be an active agonist of LRH-1 and now, it is under preclinical evaluation for toxicity study. According to our knowledge, till now no compound has reached the clinical trials, warranting the novelty of LRH-1 agonists.

The aim of this study is to identify the basic structural requirements for LRH-1 agonistic activity and thereby designing novel and potent agonists. The combination of pharmacophore modelling, 3D-QSAR and molecular docking approach has been employed to achieve this goal. The HypoGen algorithm based pharmacophore hypothesis was generated. The validated pharmacophore hypothesis was subsequently used in virtual screening process to identify novel and potent LRH-1 agonists. CoMFA and CoMSIA model were developed and validated based on the substructure alignment. Further, molecular docking analysis has rendered the complimentary information to pharmacophore and 3D-QSAR studies.

2. Materials and methods

2.1. Pharmacophore modelling

2.1.1. Data sets

Since the last half decade, a number of small molecules were reported as LRH-1 agonists. Out of these, 47 agonists assayed by same biological methods were selected and used further for pharmacophore modelling study [13,14]. The agonistic activity of these compounds was expressed in terms of EC_{50} (i.e., concentration of a compound where 50% of its maximal effect is observed) and pEC_{50} ($pEC_{50} = -\log EC_{50}$) values. The most important step in the pharmacophore modelling is the selection of a suitable training set with wide activity range of at least four orders of magnitude, responsible for determining the quality of the generated pharmacophore. The reported pEC_{50} values of dataset spanned across a small range from 5.32 to 7.92. These activity values were rescaled to the range of four log units to develop statistically reliable pharmacophore. For HypoGen pharmacophore model generation, a training set of 16 compounds (Fig. 1) was selected based on the principles of structural diversity and activity range (As per manual of Discovery Studio 2.5 (DS2.5), minimum 16 compounds are required in the training set for the development of HypoGen based Pharmacophore hypothesis). The rest of the 31 compounds (Fig. S1) from the dataset were used as a test set for pharmacophore model validation.

2.1.2. Pharmacophore model generation

The 3D QSAR Pharmacophore Generation and Ligand Pharmacophore Mapping module within DS2.5 software package [15] were used to carry out pharmacophore modelling studies. The conformations for all training set compounds were generated by Cat-Conf program within DS2.5 software package. The BEST method was employed during generation of multiple acceptable conformations. The BEST method provides complete and improved coverage of conformational space by performing a rigorous energy minimization and optimizing the conformations in both torsional and cartesian space using the poling algorithm [16]. The features such as hydrogen bond acceptor (HBD), general hydrophobic feature (H), hydrophobic aliphatic (HYA), hydrophobic aromatic (HYAr) and ring aromatic (RA) were included for the pharmacophore generation assuming common features present in the studied compounds. The uncertainty value during pharmacophore generation was set to two, as the activity range in the training set compounds barely spans the minimum requirement of four orders of magnitude as well as to correlate the training set compounds with their activity values accurately [17]. The uncertainty value of two means the biological activity of a particular agonist is assumed to be located somewhere in the range two times higher to two times lower of the true value of that agonist. Top ten pharmacophore hypotheses were generated using training set of LRH-1 agonists. The best pharmacophore hypothesis was selected based on significant statistical parameters (high correlation coefficient (r^2), lowest total cost, highest cost difference and low RMS values).

2.1.3. Validation of pharmacophore model

The validation of developed pharmacophore model was done to determine whether it is capable of differentiating between active, least active and inactive compounds [18]. To validate the best pharmacophore hypothesis three different methods were employed (test set prediction, decoy test and Fischer randomization test). A test set of 31 diverse LRH-1 agonists were used to validate the best pharmacophore hypothesis. The cost functions such as weight cost, configuration cost and error cost, calculated during the pharmacophore generation process were initially used to validate the best pharmacophore hypothesis. In test set validation method, Ligand Pharmacophore Mapping protocol with the BEST flexible search option was employed to map the test set compounds upon the best pharmacophore hypothesis. The pEC_{50} value for each test set compound was also estimated. In decoy set validation method, a small database of decoys was generated using DecoyFinder1.1. Five active LRH-1 agonists were included in the decoy database to calculate goodness of hit score (GH) and enrichment factor (E value). GH and E value are the two major parameters, playing important role in identifying capability of the generated pharmacophore hypothesis. Finally, Fischer randomization methodology was employed as the third validation procedure with a goal to check whether there is a strong correlation between the chemical structures and the biological activity in the training set. In this validation method, 19 random spreadsheets (hypotheses) were generated by randomizing the activity data of the training set compounds to achieve 95% confidence level.

2.1.4. Virtual screening

Virtual screening of chemical databases is a fast and accurate method to find potential leads suitable for further development [19]. In our study, the validated pharmacophore hypothesis (Hypo1) was used as a 3D query in database screening. Three commercially available databases (ChemDiv, Specs and NCI) of diverse chemical compounds were screened to identify novel LRH-1 agonists. From the available Fast/Flexible and Best/Flexible search option, Best/Flexible option was utilized for screening the databases. Maximum Omitted Features option was set to '−1' during screening of the chemical databases. Hit compounds were screened for their predicted biological activity values using the Hypo1 pharmacophore model. The compounds which were showing estimated pEC_{50} values greater than 7 were selected and subsequently subjected to molecular docking analysis using Glide5.5.