Journal of Molecular Structure 1098 (2015) 199–205

Contents lists available at ScienceDirect

Journal of Molecular Structure

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

QSAR studies on triazole derivatives as sglt inhibitors via CoMFA and CoMSIA

Hui Zhi [a], Junxia Zheng [b], Yiqun Chang [c], Qingguo Li [a], Guochao Liao [d], Qi Wang [e, *], Pinghua Sun [c, *]

[a] School of Chinese Materia Medica, Guangzhou University of Chinese Medicine, Guangzhou 510006, PR China
[b] Faculty of Chemical Engineering and Light Industry, Guangdong University of Technology, Guangzhou 510006, PR China
[c] Guangdong Province Key Laboratory of Pharmacodynamic Constituents of TCM and New Drugs Research, College of Pharmacy, Jinan University, Guangzhou 510632, PR China
[d] Department of Chemistry, Wayne State University, Detroit, MI 48202, USA
[e] Institute of Clinical Pharmacology, Guangzhou University of Chinese Medicine, Guangzhou 510006, PR China

ARTICLE INFO

Article history:
Received 4 December 2014
Received in revised form
1 June 2015
Accepted 1 June 2015
Available online 6 June 2015

Keywords:
QSAR
CoMFA
CoMSIA
SGLT inhibitory triazoles

ABSTRACT

Forty-six sodium-dependent glucose cotransporters-2 (SGLT-2) inhibitors with hypoglycemic activity were selected to develop three-dimensional quantitative structure-activity relationship (3D-QSAR) using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) models. A training set of 39 compounds were used to build up the models, which were then evaluated by a series of internal and external cross-validation techniques. A test set of 7 compounds was used for the external validation. The CoMFA model predicted a q^2 value of 0.792 and an r^2 value of 0.985. The best CoMSIA model predicted a q^2 value of 0.633 and an r^2 value of 0.895 based on a combination of steric, electrostatic, hydrophobic and hydrogen-bond acceptor effects. The predictive correlation coefficients (r^2_{pred}) of CoMFA and CoMSIA models were 0.872 and 0.839, respectively. The analysis of the contour maps from each model provided insight into the structural requirements for the development of more active sglt inhibitors, and on the basis of the models 8 new sglt inhibitors were designed and predicted.

1. Introduction

Sodium-dependent glucose cotransporters-2 (SGLT-2) has been recognized as a new and attractive target for diabetes patients because of its vital role in the reduction of blood glucose level [1]. Administration of SGLT2 inhibitors reduces both preprandial and postprandial blood glucose, and may decrease glucotoxicity. SGLT2 inhibitors selectively target renal glucose transporters and do not directly influence islet, thereby leaving the insulin secretion function totally intact without affecting the counter-regulatory hormones. Thus, there is a low risk of hypoglycaemia even if they are used in combination with oral anti-diabetic agents such as sulfonylurea and metformin [2]. In recent years, several SGLT-2 inhibitors have been launched or in clinical trials as drug candidates for the treatment of type II diabetes, including: Canagliflozin, Dapagliflozin, Tofogliflozin, Ipragliflozin, Luseogliflozin, Empagliflozin, Ertugliflozin, Sergliflozin etabonate, and Remogliflozin etabonate (Fig. 1) [3]. Interestingly, the majority of currently reported SGLT-2 inhibitors were various glucosides derived from natural product phlorizin. Therefore, the non-glycosides triazole derivatives with SGLT-2 inhibitory activities may offer unforeseen advantages, and be prescribed especially when the secondary failure occurred, which is very common in traditional anti-diabetic therapy regimens.

3D-QSAR techniques such as CoMFA [4] and CoMSIA [5] are routinely used in modern drug design, especially when the molecular target is either unknown or has not been structurally resolved. QSAR studies are performed by correlating physicochemical descriptors from a set of related compounds to their known biological activity or molecular property values [6,7]. These computational techniques have been proved particularly helpful in the design of novel, more potent drugs while some other computational approaches were applied to provide complementary information and to discover novel potent inhibitors. Homology

* Corresponding authors.
E-mail addresses: wqitcm@foxmail.com (Q. Wang), biochemdoctor@sina.com (P. Sun).

http://dx.doi.org/10.1016/j.molstruc.2015.06.004
0022-2860/

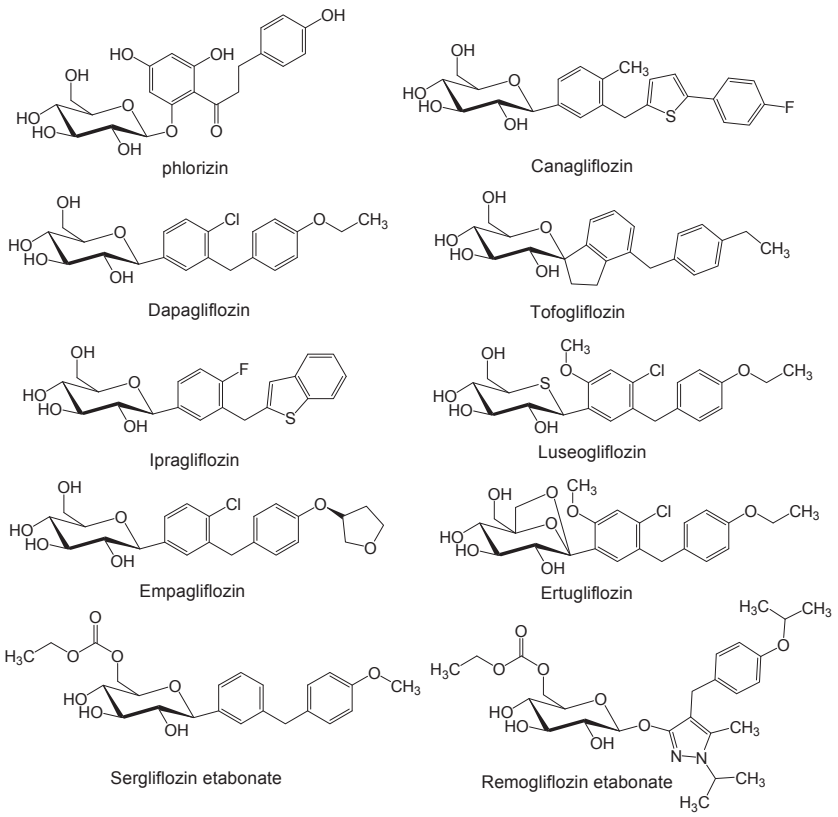

Fig. 1. SGLT-2 inhibitors as launched drugs or drug candidates in clinical trials.

modeling, molecular docking were administered on 54 C-aryl glycosides analogues, providing fundamental structural insights into the active site and key interactions involved in the binding of inhibitors to hSGLT2 [8]. 3D pharmacophore models and atom-based 3D-QSAR models have been developed using the PHASE module of Schrodinger with a series of N-β-D-xylosylindole derivatives to determine the structural requisite of these SGLT2 inhibitors [9]. Optimization of Gaussian Kernel Function in Support Vector Machine Aided QSAR studies were reported, based on 45 C-aryl glycosides SGLT2 inhibitors [10]. 2D-QSAR and 3D-QSAR models by k-nearest neighbor (kNN) method for molecular field analysis (MFA) were built based on 33 thiophenyl C-aryl glycoside SGLT2 inhibitors, revealed that a specific group or type of descriptor was not sufficient to capture the true factors responsible for the activity in 2D-QSAR [11]. CoMFA and CoMSIA studies were also made on 180 C-aryl glycosides SGLT2 inhibitors with satisfactory statistical results [12]. All these works are applied on glycosides molecules while our work focuses on non-glycosides triazole molecules, and the applications of QSAR methodology to some triazole SGLT inhibitors have not been reported. we expect that our work could provide complementary and useful information to discover novel potent SGLT2 inhibitors. The present 3D CoMFA and CoMSIA studies of forty-six triazole SGLT inhibitors will not only derive the essential structure requirements for the hypoglycemic activity but also provide useful rationale for further design of new drug candidates for type II diabetes.

2. Materials and method

2.1. Data sets

All triazole SGLT-2 inhibitors and their biological activities were collected from literature [13,14]. (Table 1) The IC_{50} values were converted into the corresponding PIC_{50} ($-\log IC_{50}$), whose activity range from 5.6 to 8.5 log units provided a broad and homogenous data for 3D-QSAR study. The "Every Nth Value" method in the SYBYL 8.1 package (Tripos Inc., St. Louis, USA) was used to creating the test set and sorted by "descending" the pIC_{50} [$-\log IC_{50}$] value, from which the value of N was set to be 7. Therefore 7 compounds were selected as test subset (15.2%), and the remaining 39 compounds as training subset (84.8%).

2.2. Molecular modeling and alignment

The 3D structures of the compounds were sketched in SYBYL 8.1 package. MMFF94 (Merck Molecular Force Field) force field and its charge were assigned to fully minimize each structure of 46 triazole derivatives, with the gradient convergence criterion set to 0.01 kcal/mol. The conformations obtained were subjected to simulated annealing to identify the global minimum energy conformations, in which the system was heated at 2000 K for 1 ps and then cooled at 200 K for 1 ps. 100 such cycles were run for each molecule, and the lowest energy conformation was determined and subsequently minimized with the same criteria as mentioned above [15]. A good alignment is the most essential for the quality and the predictive ability of CoMFA and CoMSIA models. Thus, we applied molecular alignment using Align Database – Rigid body alignment of molecules in a Mol2 database with a user-defined common core of Sybyl tools in present study [16]. The most active compound **42** was selected as the template molecule and the other molecules were aligned onto it by common substructure with bold lines as shown in Fig. 2. It can be seen that all the studied compounds have similar minimum conformations.