Data in Brief 15 (2017) 281–299

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Comprehensive data on a 2D-QSAR model for Heme Oxygenase isoform 1 inhibitors

Emanuele Amata [a,*], Agostino Marrazzo [a], Maria Dichiara [a], Maria N. Modica [a], Loredana Salerno [a], Orazio Prezzavento [a], Giovanni Nastasi [b], Antonio Rescifina [a], Giuseppe Romeo [a], Valeria Pittalà [a,*]

[a] Department of Drug Sciences, University of Catania, Viale A. Doria 6, 95125 Catania, Italy
[b] Department of Mathematics and Computer Sciences, University of Catania, Viale A. Doria 6, 95125 Catania, Italy

ARTICLE INFO

Article history:
Received 18 July 2017
Received in revised form
7 September 2017
Accepted 19 September 2017
Available online 21 September 2017

Keywords:
Heme Oxygenase
2D-QSAR
pIC50 prediction
FDA
CORAL

ABSTRACT

The data have been obtained from the Heme Oxygenase Database (HemeOxDB) and refined according to the 2D-QSAR requirements. These data provide information about a set of more than 380 Heme Oxygenase-1 (HO-1) inhibitors. The development of the 2D-QSAR model has been undertaken with the use of CORAL software using SMILES, molecular graphs and hybrid descriptors (SMILES and graph together). The 2D-QSAR model regressions for HO-1 half maximal inhibitory concentration (IC_{50}) expressed as pIC_{50} ($pIC_{50} = -LogIC_{50}$) are here included. The 2D-QSAR model was also employed to predict the HO-1 pIC_{50}values of the FDA approved drugs that are herewith reported.

Specifications Table

Subject area	Computational Chemistry
More specific subject area	Quantitative Structure-Activity Relationship (QSAR) modeling

* Corresponding authors.
E-mail addresses: eamata@unict.it (E. Amata), vpittala@unict.it (V. Pittalà).

http://dx.doi.org/10.1016/j.dib.2017.09.036

Type of data	Table, figure
How data was acquired	Statistical modeling and online databases
Data format	Raw and analyzed
Experimental factors	The whole dataset consists of 382 HO-1 inhibitors which were randomly split and divided into training, invisible training, calibration, and validation sets.
Experimental features	The 2D-QSAR models have been developed using CORAL software. Chemical structure descriptors and pIC_{50} were used as variables.
Data source location	Department of Drug Sciences, Department of Mathematics and Computer Sciences, University of Catania, Italy
Data accessibility	With this article

Value of the data

- HO-1 is a crucial enzyme involved in the catabolism of heme and overexpressed in a number of tumors with poor clinical outcome.
- 2D-QSAR modeling data was generated to provide a method useful in finding or repurposing novel HO-1 inhibitors.
- The model has also been used to predict the HO-1 pIC_{50} for the FDA-approved drugs.

1. Data

HO-1 is a crucial enzyme involved in the regioselective catabolism of heme. Strongly induced upon stressful condition, HO-1 is recognized to fulfil crucial roles in cytoprotection and in the maintenance of endogenous homeostasis, playing a role in metabolic, cardiovascular, and pulmonary diseases [1–3]. Nevertheless, under adverse circumstances it has been demonstrated that aberrant levels of HO-1 may sustain cancerous diseases. Therefore, its inhibition is of interest in all such pathological conditions [4–7]. QSAR models as well as other methods are regression, classification or statistical methods used in the chemical and biological sciences, helping in predicting variables or in understanding patterns [8–11]. Data here reported provide information about a set of HO-1 inhibitors, recovered from the Heme Oxygenase Database (HemeOxDB) together with their pIC_{50} ($-\log IC_{50}$) [12]. These latter have been used in building up the first hybrid 2D-QSAR model embracing the all set of known HO-1 inhibitors. The model has also been used to predict the HO-1 pIC_{50} for the Food and Drug Administration approved drugs. These latter predicted HO-1 pIC_{50} data are also here reported.

2. Experimental design, materials and methods

2.1. Dataset preparation

The dataset consists of 382 HO-1 inhibitors which were randomly split three times and then divided into training (131 compounds), invisible training (131 compounds), calibration (60 compounds) sets for model development and a validation set (60 compounds) for invisible model validation. The three splits and four sets have been randomly generated, and their pIC_{50} minimum, maximum and middle are reported in Table 1.