Data in Brief 9 (2016) 35–42

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

3D-QSAR modelling dataset of bioflavonoids for predicting the potential modulatory effect on P-glycoprotein activity

Pathomwat Wongrattanakamon [a], Vannajan Sanghiran Lee [b], Piyarat Nimmanpipug [c], Supat Jiranusornkul [a,*]

[a] Laboratory for Molecular Design and Simulation (LMDS), Department of Pharmaceutical Sciences, Faculty of Pharmacy, Chiang Mai University, Chiang Mai 50200, Thailand
[b] Department of Chemistry, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia
[c] Computational Simulation and Modelling Laboratory (CSML), Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form
27 July 2016
Accepted 1 August 2016
Available online 4 August 2016

Keywords:
Molecular modelling
QSAR
Multiple linear regression
P-glycoprotein
Flavonoids
Herb-drug interaction

ABSTRACT

The data is obtained from exploring the modulatory activities of bioflavonoids on P-glycoprotein function by ligand-based approaches. Multivariate Linear-QSAR models for predicting the induced/inhibitory activities of the flavonoids were created. Molecular descriptors were initially used as independent variables and a dependent variable was expressed as pFAR. The variables were then used in MLR analysis by stepwise regression calculation to build the linear QSAR data. The entire dataset consisted of 23 bioflavonoids was used as a training set. Regarding the obtained MLR QSAR model, R of 0.963, $R^2=0.927$, $R^2_{adj}=0.900$, SEE=0.197, $F=33.849$ and $q^2=0.927$ were achieved. The true predictabilities of QSAR model were justified by evaluation with the external dataset (Table 4). The pFARs of representative flavonoids were predicted by MLR QSAR modelling. The data showed that internal and external validations may generate the same conclusion.

* Corresponding author.
E-mail address: supat.jira@cmu.ac.th (S. Jiranusornkul).

http://dx.doi.org/10.1016/j.dib.2016.08.004

Specifications Table

Subject area	*Computational Chemistry*
More specific subject area	*Quantitative Structure-Activity Relationship (QSAR) modelling*
Type of data	*Equation, tables, graphs*
How data was acquired	*In silico analysis and statistical modelling*
Data format	*Analysed*
Experimental factors	*Multivariate Linear-QSAR models for predicting the induced/inhibitory activities of the flavonoids were created. Molecular descriptors were initially used as independent variables and a dependent variable was expressed as pFAR; −log (fluorescence activity ratio).*
Experimental features	*The molecular descriptors and pFAR values were used in multiple linear regression (MLR) analysis by stepwise regression calculation to generate the model. The entire dataset consisted of 23 bioflavonoids was used as a training set.*
Data source location	*Laboratory for Molecular Design and Simulation (LMDS), Department of Pharmaceutical Sciences, Faculty of Pharmacy, Chiang Mai University, Chiang Mai, Thailand*
Data accessibility	*The data is with this article.*

Value of the data

- P-gp is an important clinically mediated target of herbal compounds including flavonoids in herb-drug interactions that physicians must be aware for a safe prescription.
- 3D-QSAR modelling data was constructed for predicting P-gp inhibitory activity as pFAR values of flavonoids that may allow a primary screening for healthcare providers and benefit for patients who take more than one medication.
- The model could be utilised to screen the potential herb-drug interaction risks of flavonoids and be an alternative strategy to scrutinise flavonoids which are used to recover the pharmacological outcomes of anticancers agents which are P-gp's substrates.

1. Data

The data shown here regarding a QSAR equation construction that is used to predict the induction/inhibition of P-glycoprotein modulators.

2. Experimental design, materials and methods

2.1. Dataset for analysis

The 23 flavonoids and their induced/inhibitory activities were obtained from two publications [1,2]. The bioassay (fluorescence activity ratio; FAR at 40 μg/ml which represents P-gp induction or inhibition) values of the 23 flavonoids cover the range from 0.5 to 46.4. From the preliminary investigation using bioassay (FAR) as a dependent variable, the obtained correlation was low and increased higher in models with excessive descriptors. The FAR values were transformed becoming the corresponding pFAR ($-\log$ FAR) values, which is in the range of -1.67 to 0.3. The use of pFAR is to represent a negative value ($-$) as a P-gp inhibitory activity and a positive value ($+$) as a P-gp induced activity. Flavonoids with FAR values >1 but <10 (pFAR <0 but >-1) were regarded to be active