A MACHINE LEARNING-BASED QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDY FOR THE PSYCHOMIMETIC ACTIVITY OF PHENETHYLAMINE DERIVATES

Jamshid Kayumov[a], Abdushukur G'ofurov[a], Durbek Usmanov[b], Bakhtiyor Rasulev[a,c]

[a] National University of Uzbekistan, named after Mirzo Ulugbek, Tashkent, Uzbekistan
[b]Institute of the Chemistry of Plant Substances, Academy of Sciences, Tashkent, Uzbekistan
[c] Department of Coatings and Polymeric Materials, North Dakota State University, Fargo ND, USA

Abstract

A novel model is developed for the prediction of the psychotomimetic activity of the substituted phenethylamines as psychedelic drugs. The structure-activity analysis was performed using a quantitative structure-activity relationship (QSAR) method where molecular structures of the desired phenethylamine derivatives used and their activities. In this study 118 different substituted phenethylamines are used with reported psychotomimetic activity values. The QSAR analysis was carried out by application of combined approach of genetic algorithm for variables selection and multiple linear regression analysis. In order to find a stable conformation and generate additional descriptors for QSAR study a quantum-chemical analysis was performed by applying a semi-empirical method. As a result, a number of models were developed, where a ten-variable model showed the best predictive performance with r^2 = 0.7428 and q^2_{LOO} = 0.6738. The robustness and predictability of the best model was validated using a leave-one-out technique, external set and *y*-scrambling methods. The predictive ability of the model was confirmed with the external set, showing the r^2_{ext} = 0.7365. The developed model can be used in the prediction of the psychotomimetic activity of new and untested organic compounds.

Materials and Methods.

A specific QSAR modeling techniques used in this study [1]. The dataset of the compounds for the present research work was collected from several published experimental data [2-3] with psychotomimetic activity (PA). All original activity data were converted into molar 1/log(PA) response variables.

Results and Discussion.

The whole set of 118 compounds was divided into the training set consisted of 98 compounds and a test set (predicting set) of 20 compounds. GA-MLRA technique has identified several models.

The following equation represent the developed model towards the PA:

Log(PA)= 0.4937 (±0.3728)**Jhetv** -0.0315(±0.0157) **MPC07**-0.8057 (±0.4383) **MATS8v** -0.2671(±0.2819) **GATS5e**+11.0819 (±4.0085) **BELm1**+1.0233 (±0.4618) **BEHe8** 3.1745(±1.3014) **E2m**-20.5801(±13.8738) **R6u+**

+10.8007 (±5.1876) **R4p+** +0.1922 (±0.1824) **H-051**-19.1914(±7.7205)

REFERENCES

1. Patnode K., Demchuk Z., Johnson S., Voronov A., Rasulev B. Combined Computational Protein-ligand Docking and Experimental Study of Bioplastic Films from Soybean Protein, Zein and Natural Modifiers, ACS Sustain. Chem. Eng., 2021, 9, 10740-10748

2. Ashby J., Tennant R.W. Mutat. Res. 1991, 257, 229-306

3. Rosenkranz H.S., Klopman G. Toxicol. Ind. Health. 1988, 4, 533-540