ЖУРНАЛ СТРУКТУРНОЙ ХИМИИ

2017. Том 58, № 2 *Февраль – март* *С. 361 – 365*

КРАТКИЕ СООБЩЕНИЯ

UDC 541.6:539.2

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDY ON THE BINDING AFFINITY OF SOME AMINOTHIAZOLE DERIVATIVES WITH A DOPAMINE RECEPTOR IN BRAIN

M. Mahani[1], S. Sheikhghomi[2], H. Sheikhghomi[3], J. Fasihi[3]

[1]*Department of Chemistry and Nanochemistry, Faculty of Sciences and Modern Technologies, Graduate University of Advanced Technology, Kerman, Iran*
E-mail: mohmahani@gmail.com
[2]*Department of Medicine, Iran University of Medical Sciences (IUMS), Tehran, Iran*
[3]*Nuclear Science and Technology Institute, Tehran, Iran*

Received August, 3, 2015

Quantitative structure-activity relationship (QSAR) models can be applied as a powerful tool for predicting the response in biological and chemical systems. D_2 receptor subtypes as dopamine receptors assist dopaminergic neurotransmission in brain. In this work, binding affinities of synthesized agonists with D_2-like receptors in binding assays using rat brain were related to the structural properties of these agonists. The structural descriptors of these compounds are calculated. A stepwise variable selection is applied for PLS modeling. The walk and path counts, 2D autocorrelations, 3D atom pairs, RDF, 3D-MoRSE, WHIM, GETAWAY blocks are among the selected descriptors. The PLS model is built with 5 latent variables. The predictive ability of the model is evaluated on a set of 5 ligands, which are not used in modeling steps and the acceptable results are obtained.

DOI: 10.15372/JSC20170215

Keywords: QSAR, partial least squares, soft modeling, structural descriptors, aminothiazole derivatives, dopamine receptor.

Dopamine plays essential roles in most neural pathways of the central nervous system (CNS). First of them is the nigrostriatal pathway comprising the neurons of the substantianigra (A9) and projecting to neurons of the neostriatum, and the second is the mesocorticolimbic pathway composed of neurons of the ventral tegmental area (A10), connecting with those of the limbic cortex and other limbic structures [1]. The dopamine molecule has been associated with many physiological functions, such as fine movement coordination, cognition, emotion, and memory, by the pituitary reward system [2, 3]. Alterations in the dopaminergic function are involved in the pathogenesis of Parkinson′s disease [4], psychomotor diseases, and schizophrenia. Therefore, the dopamine receptor families have been targeted in the development of treatment medications for these disorders [2]. Five subtypes of the dopamine receptor have been identified (D_1—D_5). Dopamine receptors may be classified into two subfamilies (the D_1-like receptors (D_1, D_5) and the D_2-like receptors (D_2, D_3, D_4)) based on the DNA sequence and functional properties [3, 5—8]. The D_2 dopamine receptor is located on postsynaptic dopaminergic neurons that are centrally involved in reward-mediating mesocorticolimbic pathways. Signaling through D_2 receptors governs the physiologic functions related to locomotion, hormone production, and drug abuse. D_2 receptors are also known targets of antipsychotic drugs used to treat psychomotor diseases such as schizophrenia, a debilitating mental illness that affects 0.5—1.5 % of the worldwide population. Although the biophysical and pharmacological properties of D_2 receptors have been studied, many problems remain unresolved due to the lack of three-dimensional structures and

experimental limitations [2]. Older dopamine agonists used for Parkinson′s disease such as bromocriptine and cabergoline are poorly selective for one dopamine receptor over another, and, although most of these agents do act as D_2 agonists, they affect the other subtypes as well. Several selective D_2 ligands are now available and this number is likely to increase as further research progresses. Hence, considerable effort has been devoted to the discovery and development of potent and selective D_2 ligands. The prediction of the binding affinity of agonists at D_2 as a dopamine receptor has been considered as an attractive subject for researchers.

Chemometric modeling methods can relate the molecular structure with the binding affinity of agonists, antagonists, and the other compounds due to the facilitation of the synthesis of the expected structures. Wilcox et al. have predicted agonist affinities at recombinant D1 vs. D2 dopamine receptors by the comparative molecular field analysis (CoMFA) method [9]. CoMFA is a 3D QSAR technique based on data from the known active molecules. The aim of CoMFA is to derive a correlation between the biological activity of a set of molecules and their 3D shape, electrostatic and hydrogen bonding characteristics. Pettersson synthesized new dopaminergic ligands, a set of 4-phenylpiperidines and 4-phenylpiperazines, and studied the pharmacological, characterizing and quantitative structural-retention relationship (QSAR) modeling of them in both *in vivo* and *in vitro* assays [10]. Avram et al. predicted inhibition constants *K*i for 71 antipsychotics, already approved for clinical treatment, as well as new representative chemical structures that exhibit the antipsychotic activity were evaluated using 3D-QSAR—CoMSIA models [11]. A series of new compounds with the goal to identify the potent and selective dopamine receptors were synthesized by Chen et al. [12].

In the present work, the binding affinity of a series of agonists with D_2 as a dopamine receptor in rat brain was modeled by partial least squares (PLS) as a powerful modeling tool. The resulted model was examined on a test set to evaluate the prediction of the binding affinity of agonists.

Material and methods. Software and methods. PLS algorithms were written in MATLAB (R 2010, version 7.10). HyperChem for Windows ver.7.0 (Hypercube Inc., Alberta, Canada) and Dragon (Milano Chemometrics and QSAR Research Group, Milano, Italy) were used for optimizing the molecular structures and calculating the molecular descriptors, respectively. All programs were run on a PC (Pentium IV, 2.5 GHz) with the Microsoft Windows 7 operating system.

Dataset and descriptors. Binding affinities of 41 new synthesized agonists (see http://jsc.niic.nsc.ru) at the D2-like receptors were taken from the literature [12]. The dataset was randomly divided into two groups: a training set (thirty six agonists) and a prediction set (five agonists). HyperChem was used for drawing the molecular structures. Before calculating the molecular descriptors, the optimization of the 3D molecular structures was performed by a semi-empirical AM1 method. A total of 1022 molecular descriptors (constitutional, 3D-MoRSE …) of different kinds were calculated by Dragon for each chemical compound.

Chemometric method. PLS. In the PLS regression the Y-variables are regressed on the selected latent variables that are a linear combination of original variables. The PLS method takes into account the information about Y-variables during the decomposition of the independent variable data matrix. A non-linear iterative partial least squares (NIPALS) algorithm was used in the present work. With this algorithm, the data matrix of the descriptors is compressed to a smaller data matrix, called ″scores″, in a new coordinate system. The new coordinate axes are called latent variables (LVs), or factors, and are linear combinations of the original variables. LV represents a systematic variation found in the dataset. The regression coefficients from each original variable to each LV are called loadings. The number of LVs to be used within the PLS algorithm is an important parameter to achieve better performance in the prediction. This allows modeling of the system with the optimum amount of information and the avoidance of overfitting or under-fitting. The leave-one-out cross validation procedures consists of the systematic removing of one of the training compounds in turn, and the use of only remaining ones for constructing the model. In our study, the predicted values were compared with the actual ones for a calibration set, and the predicted error sum of squares (PRESS) was calculated. PRESS was computed in the same manner, and each time a new factor was added to the PLS model.

$$\text{PRESS} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2.$$

Here y_i is the actual value and $\hat{y}_i$ is the predicted value for the i compound, and n is the number of compounds.

Results and discussion. A receptor is a molecule or a polymeric structure in a cell, which acts as the biological target, recognizing and binding a compound. Agonist is a chemical that binds to a receptor and activates the receptor to produce a biological response. A set of structural descriptors were generated for the 41 new synthesized agonists. This pool of descriptors was further treated by mathematical procedures.

Variable selection. The calculated descriptors contain the physicochemical properties of a series of D_2-like receptors. The descriptors that have zero or constant values for about 40 percent of the compounds were removed from the dataset. Also the correlation coefficients of the descriptors with each other were calculated and the descriptors with the correlation coefficients of 0.8 and lower correlation with the binding affinities were removed from the dataset. A stepwise variable selection was applied for extracting meaningful descriptors. In the stepwise selection, the forward selection and backward elimination procedures were made with alpha to enter and alpha to remove values equal to 0.15. Alpha is the maximum acceptable level of risk for rejecting a true null hypothesis (type I error) and is expressed as a probability ranging between 0 and 1 and alpha is frequently referred to as the level of significance. In the forward selection and backward elimination procedures p-values are compared to alpha to determine the significance. If the p-value is less than or equal to the alpha level, the null hypothesis is rejected in favor of the alternative hypothesis, if else, the null hypothesis is not rejected. At the 0.15 alpha level the chance of finding an effect that does not really exist is only 15 %. A decrease in the alpha value also means a decreased chance of detecting an effect if one truly exists. Finally, thirteen descriptors were selected for PLS modeling by the stepwise selection (Table 1).

PLS model. The PLS model was trained by13 selected descriptors (Table 1). All of the 13 descriptors were pretreated by mean centering and scaling (with each variable variance) and then modeled by PLS. The leave-one-out (LOO) cross validation was used for the validation and selection of the best number of LVs during the training step. In LOO-CV, PRESS was plotted against the number of LVs. As shown in Fig. 1, the minimum PRESS was obtained with 5 LVs. The regression coefficients

Table 1

Thirteen selected descriptors for optimized modeling

Descriptor	Descriptor block (Dimension)	Meaning
Ss	Constitutional descriptors (0D)	Sum of Kier-Hall electrotopological states
SRW05	Molecular walk counts (2D)	Self-returning walk count of order 05
GATS4v	2D autocorrelations (2D)	Geary autocorrelation — lag4 / weighted by atomic van der Waals volumes
G (N…O)	Geometrical descriptors (3D)	Sum of geometrical distances between N…O
RDF085u	RDF descriptors (3D)	Radical Distribution Function — 8.5 / unweighted
RDF125u	RDF descriptors (3D)	Radical Distribution Function — 12.5 / unweighted
RDF090m	RDF descriptors (3D)	Radical Distribution Function — 9.0 / weighted by atomic masses
Mor25v	3D-MoRSE descriptors(3D)	3D-MoRSE— signal 25 / weighted by atomic van der Waals volumes
L2u	WHIM descriptors (3D)	1st component size directional WHIM index / unweighted
G2v	WHIM descriptors (3D)	2st component symmetry directional WHIM index / weighted by atomic van der Waals volumes
G3e	WHIM descriptors (3D)	2st component symmetry directional WHIM index / weighted by atomic Sanderson electronegativities
R6m	GERAWAY descriptors (3D)	R autocorrelation of lag 6 / weighted by atomic masses
R5p+	GERAWAY descriptors (3D)	R maximal autocorrelation of lag 5 / weighted by atomic polarizabilities

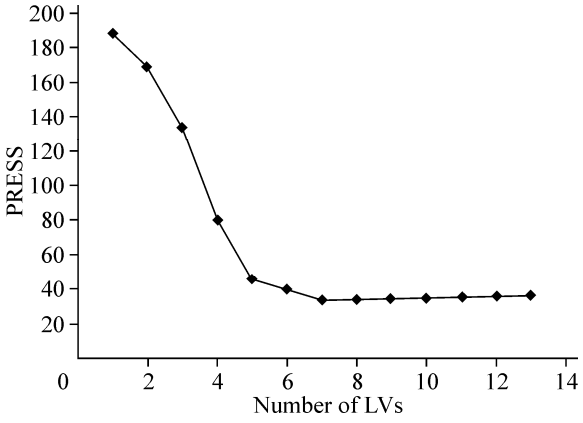

Fig. 1. PRESS vs. the number of LVs for the PLS model

for this model are listed in Table 2. The descriptors in the model were used to interpret the agonist action at the D2 receptor. The model was applied on new agonists to predict the binding affinity and the results were gathered in Table 3. The results show that it is possible to make a prediction of the agonist affinity and interpret the mechanism of the action.

Model descriptors. In this study, the response of a series of new synthesized dopamine agonists at the D_2 dopamine receptor was examined using a chemometric approach. The aim of this work was to try to understand the mechanistic basis of the agonist action. Kier and Hall have developed the concept of the electrotopological state index for atoms in a molecule [13, 14]. Examples of uses of these descriptors in QSAR in the areas of NMR chemical shifts, the inhibition of monoamine oxidase and receptor binding affinities have been published [15]. Carbon, nitrogen, oxygen, sulfur, and the halogens are the main building blocks of drug-like molecules. Kier and Hall used 35 E-state-atom-types of these main elements to calculate E-states. The advantage of E-states over the counts of the same atom types is that the steric and electronic effects of the surrounding atoms accounted in E-states. Among the selected descriptors ″Ss″ is the sum of Kier-Hall electrotopological states. Another descriptor in the model was the self-returning walk count of order 05. Self-returning walk counts are a particular case of walk counts that contain the atomic and molecular descriptors obtained from an H-depleted molecular graph, based on graph walks starting and ending at the same vertex [16]. Molecular self-returning walk counts are the spectral moments of the adjacency matrix, which were also expressed as linear combinations of counts of certain fragments contained in the molecular graph. Another class of descriptors is derived from the 3D structure of the molecules. This class of descriptors is generally calculated based on the computationally optimized molecular geometry. Other descriptors in our model were in the geometrical class of descriptors.

Conclusions. The data presented here show that it is possible to obtain measures of the agonist binding and efficacy at D_2 dopamine receptors using a chemometrics method. Differences in agonist efficacy patterns may be apparent for different receptor isoforms.

T a b l e 2

Regression coefficients for the PLS *model*

Descriptor	Regression Coefficient	Descriptor	Regression Coefficient
Ss	–0.24	Mor25v	–0.64
SRW05	0.16	L2u	0.28
GATS4v	–0.32	G2v	0.19
G (N…O)	0.25	G3e	0.16
RDF085u	0.14	R6m	0.06
RDF125u	–0.16	R5p+	–0.10
RDF090	0.39		

T a b l e 3

Prediction results of the PLS *model on the test set*

Name	Actual value	PLS	
		Prediction	Error (%)
11	185	194.84	5.3172
22	328	355.74	8.4586
23	470	521.65	10.989
25	491	522	6.3144
33	127	118.33	–6.8238

REFERENCES

1. *Björklund A., Hökfelt T.* Handbook of chemical neuroanatomy. – Amsterdam: Elsevier, 1983.
2. *Wang Y.T., Su Z.Y., Hsieh C.H. et al.* // J. Chem. Inf. Model. – 2009. – **49**. – P. 2369 – 2375.
3. *Missale C., Nash S.R., Robinson S.W. et al.* // Physiological Rev. – 1998. – **78**. – P. 189 – 225.
4. *Lotharius J., Brundin P.* // Nat. Rev. Neurosci. – 2002. – **3**. – P. 932 – 942.
5. *Gardner B., Strange P.G.* // Br. J. Pharmacol. – 1998. – **124**. – P. 978 – 984.
6. *Sibley D.R., Monsma F.J.* // Trends Pharmacol. Sci. – 1992. – **15**. – P. 264 – 270.
7. *Civelli O., Bunzow J.R., Grandy G.K.* // Annu. Rev. Pharmacol. Toxicol. – 1993. – **32**. – P. 281 – 307.
8. *Neve K., Neve R.A.* The Dopamine Receptors. – New Jersey, Totowa: Humana Press, 1997.
9. *Wilcox R.E., Tseng T., Brusniak M.Y. et al.* // J. Med. Chem. – 1998. – **41**. – P. 4385 – 4399.
10. *Pettersson F.* Synthesis, pharmacological, characterization, QSAR modelling of 4-phenylpiperidines, 4-phenylpiperazines effects on the dopaminergic neurotransmission in vivo, Department of Chemistry, Molecular Biology. – University of Gothenburg, Göteborg, Sweden, 2012.
11. *Avram S., Duda-Seiman D., Borcan F. et al.* // J. Serb. Chem. Soc. – 2011. – **76**. – P. 263 – 281.
12. *Chen J., Levant B., Wang S.* // Bioorg. Med. Chem. Lett. – 2012. – **22**. – P. 5612 – 5617.
13. *Kier L.B., Hall L.H.* Molecular structure description: the electrotopological state. – San Diego: Academic Press, 1999.
14. *Hall L.H., Mohney B., Kier L.B.* // J. Chem. Inf. Comp. Sci. – 1991. – **31**. – P. 76 – 82.
15. *Hall L.H., Mohney B., Kier L.B.* // Quant. Struct.-Act. Rel. – 1991. – **10**. – P. 43 – 51.
16. *Harary F.* Graph theory, Addison—Wesley Pub. Co., Reading, Mass., 1969.