DOI: 10.6060/ivkkt.20196205.5933

УДК: 543.683

ПРОГНОЗИРОВАНИЕ ЭСТРОГЕННОЙ АКТИВНОСТИ БИСФЕНОЛА А И ЕГО АНАЛОГОВ С ИСПОЛЬЗОВАНИЕМ КВАНТОВО-ХИМИЧЕСКИХ ВЫЧИСЛЕНИЙ И ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Ву Ван Дат, Ле Ким Лонг, Доан Ван Фук, Нгуен Хоанг Чанг, Нгуен Ван Чанг, Нгуен Тхи Тху Ха

Ву Ван Дат *

Химический Факультет, Институт Естественных Наук, Вьетнамский Национальный Университет, Ханой, 19 Ле Тхань Тонг, Хоан Кием, Ханой, Вьетнам
Email: vvdat@most.gov.vn*

Ле Ким Лонг, Нгуен Хоанг Чанг

Педагогический Факультет, Институт Образования, Вьетнамский Национальный Университет, Ханой, 144 Суан Тху, Кау Зьау, Ханой, Вьетнам
Email: longlk@vnu.edu.vn, hoangtrang@mail.ru

Доан Ван Фук

Институт химии и материалов, 17 Хоанг Шам, Кау Зьау, Ханой, Вьетнам.
Email: doanphucspbstu@mail.ru

Нгуен Ван Чанг

Институт тропической технологии, Вьетнамская академия наук и технологий, 18 Хоанг Суок Вьет, Кау Зьау, Ханой, Вьетнам
Email: nguyenvantrangsphn@gmail.com

Нгуен Тхи Тху Ха

Ханойский национальный университет образования, 136 Суан Тху, Кау Зьау, Ханой, Вьетнам.
Email: ntt.ha@hnue.edu.vn

В данной работе представлены результаты исследования количественной взаимосвязи между структурой и активностью (QSAR) бисфенола А и его аналогов с использованием квантовых химических расчетов и метода искусственных нейронных сетей (ANN). Анализ молекулярной структуры был выполнен на основе теории функционала плотности (DFT) методом B3LYP/6-31+G(d). Квантовые химические расчеты концентрируются на оптимизировании молекулярной структуры, частоты колебаний и энергии молекулярной орбиты. Распределение электронной плотности атомов было изучено в рамках естественного орбитального анализа связи (NBO). Все расчеты были выполнены в программных пакетах Gaussian 9 и Gaussview 6. Полученные в результате расчетов структурно-квантовые параметры и ранее известные экспериментально наблюдаемые биологические данные об эстрогенной активности рабочей выборки были использованы в качестве входных данных для построения модели QSAR с использованием метода искусственных нейронных сетей. С помощью метода искусственных нейронных сетей были выявлены также структурно-квантовые параметры, наиболее эффективно влияющие на эстрогенную активность исследованных веществ. Для проверки эффективности и стабильности построенных моделей был использован метод скользящего контроля и формирования внешней тестовой выборки. Полученые QSAR модели имеют довольно хорошие статистические параметры: $R^2 = 0,99$; $Q^2_{LOO} = 0,98$; $R^2_{Predict} = 0,98$. В соответствии с этим результатом следует отметить, что предлагаемая нами модель, построенная методом искусственных нейронных сетей с использованием параметров квантовой химии, достаточно адекватна и может быть полезна для прогнозирования эстрогенной активности неизученных производных и аналогов BPA с достаточной надежностью.

Ключевые слова: количественные соотношения «структура-свойство», модель QSAR, бисфенол А, эстрогенная активность, теория функционала плотности, 6-31+G(d) базисная функция, метод искусственных нейронных сетей

PREDICTING ESTROGEN ACTIVITIES OF BISPHENOL A AND ITS ANALOGS USING QUANTUM CHEMISTRY CALCULATIONS AND ARTIFICIAL NEURAL NETWORKS

Vu Van Dat, Le Kim Long, Doan Van Phuc, Nguyen Hoang Trang, Nguyen Van Trang, Nguyen Thi Thu Ha

Vu Van Dat*

Faculty of Chemistry, University of Science, Vietnam National University, Le Thanh Tong st., 19, Hoan Kiem Dist., Hanoi, Vietnam
Email: vvdat@most.gov.vn*

Le Kim Long, Nguyen Hoang Trang

Faculty of Education, University of Education, Vietnam National University, Xuan Thuy st., 144, Cau Giay Dist., Hanoi, Vietnam
Email: longlk@vnu.edu.vn, hoangtrang@mail.ru

Doan Van Phuc

Institute of Chemistry and Material, Hoang Sam st., 17, Cay Giay Dist., Hanoi, Vietnam
Email: doanphucspbstu@mail.ru

Nguyen Van Trang

Institute for Tropical Technology, Vietnam Academy of Science and Technology, Hoang Quoc Viet st., 18, Cau Giay Dist., Hanoi, Vietnam
Email: nguyenvantrangsphn@gmail.com

Nguyen Thi Thu Ha

Hanoi National University of Education, 136 Xuan Thuy st., Cau Giay, Hanoi, Vietnam
Email: ntt.ha@hnue.edu.vn

This article presents the results of the quantitative structure – activity relationship (QSAR) study of bisphenol A (BPA) and its analogs using quantum chemistry calculations and method of artificial neural networks (ANN). Molecular structural analysis is performed using Density Functional Theory (DFT) at the B3LYP/6-31+G(d) level. The quantum calculations focus on finding the optimized molecular structures, vibrational frequencies, the molecular orbital energies with reasonable accuracy. The study of electron density distribution was carried out in the framework of the natural bond orbital (NBO) methods. The obtained parameters and known observable estrogen activities are used as input data for constructing the QSAR model, using the artificial neural network method. Based on the artificial neural network method the quantum parameters having the strongest impact on the estrogen activity of the compounds were revealed. The internal and external validation methods have been performed to test the performance and the stability of the model. The statistical parameters obtained of the QSAR model were: $R^2 = 0.99$; $Q^2_{LOO} = 0.98$; $R^2_{Predict} = 0.98$. According to the obtained results, our proposed model, constructing by method of artificial neural network using the parameters of quantum chemistry is adequate and may be useful to predict of estrogen activities for unexplored derivatives and BPA analogs with moderate reliability.

Key words: quantitative structure – activity relationship, QSAR model, bisphenol A, estrogen, Density Functional Theory, 6-31+G(d) basis function, artificial neural networks – ANN

Для цитирования:
Ву Ван Дат, Ле Ким Лонг, Доан Ван Фук, Нгуен Хоанг Чанг, Нгуен Ван Чанг, Нгуен Тхи Тху Ха Прогнозирование эстрогенной активности бисфенола А и его аналогов с использованием квантово-химических вычислений и искусственных нейронных сетей. *Изв. вузов. Химия и хим. технология*. 2019. Т. 62. Вып. 5. С. 31–37

For citation:
Vu Van Dat, Le Kim Long, Doan Van Phuc, Nguyen Hoang Trang, Nguyen Van Trang, Nguyen Thi Thu Ha Predicting estrogen activities of bisphenol A and its analogs using quantum chemistry calculations and artificial neural networks. Izv. *Vyssh. Uchebn. Zaved. Khim. Khim. Tekhnol*. 2019. V. 62. N 5. P. 31–37

INTRODUCTION

Quantitative structure – activity relationship (QSAR) is a statistical research method that establishes empirical models describing relationship between the biological activity of compounds and their chemical structures. In fact, the structure-activity relationship is often complex and difficult to express by mathematical functions [1]. The application of classical data processing methods (Multivariate Regression Analysis, Partial Least Squares, etc.) to QSAR models is often problematic. The requirement for QSAR researchers is to apply state-of-the-art data processing and statistical analysis methods through specialized software to create a highly stable and predictable model. One of the most widely used data processing methods today is artificial neural network (ANN).

In this article, we developed a QSAR model for Bisphenol A and its analogs using quantum chemistry and artificial neural networks. The basis for selecting this substance for investigation is derived from the need to use BPA and its analogs [2, 3] as well as their severity [4-9].

METHODS AND MATERIALS

A general schema of QSAR method is showed in Fig. 1.

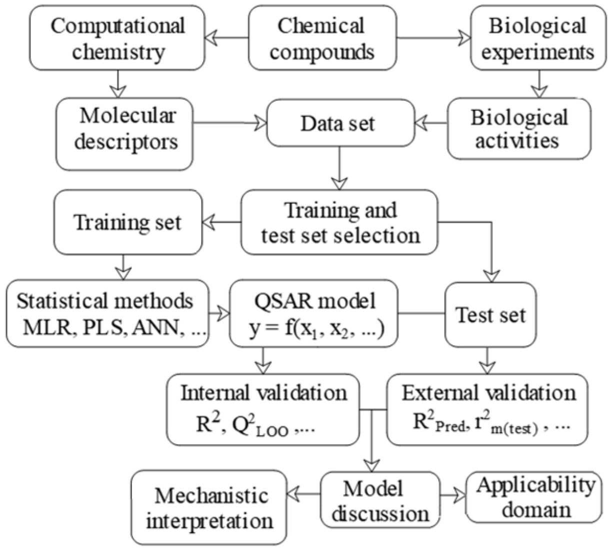


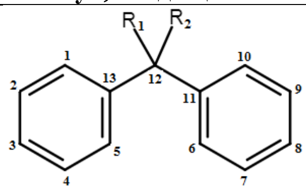
Fig. 1. Schema of QSAR method
Рис. 1. Схема построения QSAR модели

2.1. Biological data

The data set used in this study include 23 compounds synthesized and their biological activities analyzed by a team of researchers from the University of Minnesota and the University of New Orleans, USA [10]. The biological activity selected for this study is estrogen level of activity evaluated in the form of biological expression of the reporter genes embedded in the cell. The structures of the molecules in the data set are presented in Table 1.

Table 1

Structures of the molecules in the data set [10]
***Таблица 1*. Структуры изучаемых молекул, входящих в набор данных**



Compound		Substituents										
		1	2	3	4	5	6	7	8	9	R_1	R_2
1	DM DMB Bis A	H	CH_3	OH	H	H	H	CH_3	OH	H	CH_3	$CH_2CH(CH_3)_2$
2	DMB Bis A	H	H	OH	H	H	H	H	OH	H	CH_3	$CH_2CH(CH_3)_2$
3	MM4	H	H	OH	H	H	H	H	OH	H	C_2H_5	C_2H_5
4	Bis A	H	H	OH	H	H	H	H	OH	H	CH_3	CH_3
5	HF Bis A	H	H	OH	H	H	H	H	OH	H	CF_3	CF_3
6	DM HPTE	H	CH_3	OH	H	H	H	CH_3	OH	H	H	CCl_3
7	MM1	H	H	OH	H	H	H	H	OH	H	H	CH_3
8	Bis F	H	H	OH	H	H	H	H	OH	H	H	H
9	Bis B	H	H	OH	H	H	H	H	OH	H	CH_3	C_2H_5
10	DM Bis A	H	CH_3	OH	H	H	H	H	OH	CH_3	CH_3	CH_3
11	HPTE	H	H	OH	H	H	H	H	OH	H	H	CCl_3
12	1844-00-44	H	H	OH	H	H	H	H	OH	H	H	$CH(CH_3)_2$
13	MM2	H	H	OH	H	H	H	H	OH	H	H	C_2H_5
14	TM Bis A	H	CH_3	OH	CH_3	H	H	CH_3	OH	CH_3	CH_3	CH_3
15	o,p'-Bis A	H	H	H	H	OH	H	H	OH	H	CH_3	CH_3

	Compound	1	2	3	4	5	6	7	8	9	R_1	R_2
16	Mono Mxy Bis A	H	H	OH	H	H	H	H	OCH_3	H	CH_3	CH_3
17	P Bis A	H	H	OH	H	H	H	H	OH	H	CH_3	C_6H_5
18	PCP	H	H	H	H	H	H	H	OH	H	CH_3	CH_3
19	MH MM1	H	H	OH	H	H	H	H	H	H	CH_3	H
20	MH Bis F	H	H	H	H	H	H	H	OH	H	H	H
21	TC Bis A	H	Cl	OH	Cl	H	H	Cl	OH	Cl	CH_3	CH_3
22	TB Bis A	H	Br	OH	Br	H	H	Br	OH	Br	CH_3	CH_3
23	Mxy Bis A	H	H	OCH_3	H	H	H	H	OCH_3	H	CH_3	CH_3

2.2. Structural parameter calculation

The quantum parameters characterizing the molecular structure are calculated based on the Density Functional Theory (DFT) at B3LYP/6-31+G(d) level [11-13]. The calculation methods are implemented in the Gaussian 09 software [12]. The visualization of the studied molecules and the initial structures are conducted using ChemCraft software [14]. The list of structural quantum parameters calculations is presented in Table 2.

Table 2

List of quantum descriptorsof of bisphenol A analogs

***Таблица 2*. Список квантовых дескрипторов исследованных веществ**

Symbol	Definition	Unit
Length	Bond length	Å
Angel	Bond angle	Degree
Dihedral	Dihedral	Degree
E_{HOMO}	Energy of the Highest Occupied Molecular Orbital	eV
E_{LUMO}	Energy of the lowest unoccupied Molecular Orbital	eV
μ	Dipole Moment	Debye
E_{sp}	Total Energy of Molecular	eV
ΔE	$\Delta E = E_{LUMO} - E_{HOMO}$	eV
χ	$\chi = \frac{-(E_{LUMO} + E_{HOMO})}{2}$ – absolute electronegativity	eV
η	$\eta = \frac{E_{LUMO} - E_{HOMO}}{2}$ – absolute hardness	eV
ω	$\omega = \mu^2 / 2\eta$ – reactivity index	eV
C1, C2, … C13	Atomic charges on carbon atoms 1,2,.. respectively	e

2.3. Statistical analysis and variable selection

In this article, three-layer artificial neural network (input layer, hidden layer and output layer) with feedforward technique is used. Accordingly, each node of the preceding layer links with all nodes in the following layers and nodes in the same layer are not linked. The number of nodes entered by independent variable numbers will be used; the number of nodes of the output layeris a representative value of the activity of the molecule being investigated; the number of nodes in the hidden layer is set by the network regulator in such a way that all information in the data set is retrieved, but no over-processing occurs. This is also the most common type of ANN used in the QSAR model [15-18].

The operation diagram of ANN is illustrated in Fig. 2, where p is the input data vector; w is the weight typical for the link between data (signal) transmitted and data (signal) received; f is the transfer function. In the ANN network, the function f is usually a hard-limit, linear, log-sigmoid or tan-sigmoid function.

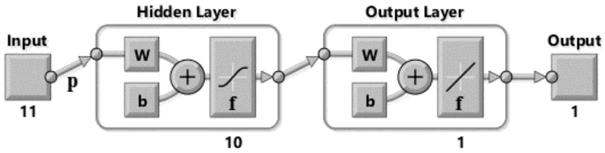


Fig. 2. Diagram of artificial neuron network

Рис. 2. Схематическая иллюстрация искусственных нейтронных сеток

However, not all of the calculated structural parameters are statistically significant for the model, therefore,selective evaluation of the potential variables to construct the structural data sets must be conducted. The process of selecting variables was performed by examining the correlation between structural parameters through the Pearson correlation matrix to find parameters that are closely related. Next, the set of structural parameters continues to be filtered through sensitivity about the mean procedure with the activity value. The nature of this operation is to investigate the variability of the activity according to the variability of an independent variable while keeping the values of other independent variables. These procedures are implemented through the NeuroSolution 6.0 program and the NeuroSolution for Excel add-in [11].

Testing the stability, predictability and generality of models a crucial step in QSAR research. The quality and predictability of the models is assessed by internal and external validation through statistical indicators [19, 20]. Cross-validation is a common method of referencing. Accordingly, 23 compounds were randomly divided into 3 sets: training set (70% of molecules); validation set (15% of molecules) and test set (15% of elements). The predictability of the model is evaluated through the leave-three-out procedure with the characteristic parameter as the generalization coefficient Q^2. The more Q^2 approaches 1, the more stable and predictable the model is. However, according to Tropsha and other experts in the field of QSAR research, internal modeling indicators do not guarantee the predictability for new compounds outside the model [19]. A quality and predictable QSAR model must be built with external references. The model's external references is made by using the model constructed from training set to predict the activity of the compounds in the test set, and then the predicted values are computed to calculate out-of-model indicators.

For ANN, the cessation criterion is that the Mean Square Error (MSE) of the validation indicates an increase; training algorithm is Levenberg-Marquardt; number of neurons on hidden layers is examined from 1 to 10; MSE values from 10 ANN models will be compared to select the optimum network for the smallest MSE.

RESULTS AND DISCUSSION

We calculated 50 characteristic parameters of 23 molecules. These parameters were compared with the corresponding parameters of BPA molecule calculated using the same method. The results indicate that there are 18 parameters that have zero descriptor. The remaining parameters were selected through the Pearson correlation coefficient matrix and their sensitivity about the mean values with activity values.

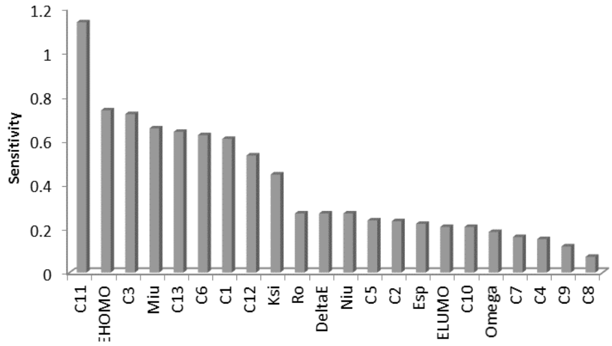


Fig. 3. Schema illustrating the sensitivity of parameters to activity
Рис. 3. Чувствительность дескрипторов на биоактивность

Based on the Pearson correlation matrix and the schema showing the sensitivity of the parameters with activity (Fig. 3), the selection of the 10 most sensitive parameters could be considered independent variables set up to build the model. Data on these parameters are given in Table 3.

Table 3

Structural parameters selected to build the model
***Таблица 3*. Дескрипторы, используемые для построения QSAR модели**

Compound	C11	E_{HOMO}	C3	μ	C13	C6	C12	ρ	C5	E_{SP}
DMB Bis A	-0.048	-5.92166	0.303	1.625279	-0.049	-0.209	-0.1	0.36884	-0.205	-849.63207
HPTE	-0.069	-6.23404	0.315	1.877432	-0.079	-0.201	-0.304	0.41546	-0.194	-2071.14844
MM4	-0.052	-5.98588	0.303	2.2785	-0.052	-0.218	-0.098	0.36794	-0.218	-810.31824
DM DMB Bis A	-0.04	-5.79295	0.308	0.922172	-0.04	-0.203	-0.09	0.36609	-0.206	-928.27007
HF Bis A	-0.095	-6.55186	0.319	1.5767	-0.095	-0.198	-0.032	0.35784	-0.198	-1327.16975
Bis B	-0.055	-5.96601	0.303	2.076508	-0.052	-0.214	-0.105	0.36975	-0.212	-771.00760
DM Bis A	-0.045	-5.77608	0.308	1.9216	-0.045	-0.215	-0.112	0.37007	-0.215	-810.33368
P Bis A	-0.067	-5.82533	0.349	1.30318	-0.066	-0.207	-0.077	0.37783	-0.203	-923.46071
MM2*	-0.063	-5.95268	0.303	2.048301	-0.061	-0.215	-0.282	0.37456	-0.21	-731.69810
Bis A	-0.052	-5.93281	0.302	1.7296	-0.052	-0.208	-0.113	0.36948	-0.208	-731.69781
PCP	-0.052	-6.06833	-0.248	1.545729	-0.021	-0.208	-0.115	0.35871	-0.228	-656.47670
TM Bis A	-0.037	-5.66941	0.312	1.3057	-0.037	-0.212	-0.111	0.36796	-0.212	-888.97147
MH MM1	-0.057	-6.07839	-0.247	1.535246	-0.028	-0.21	-0.29	0.35936	-0.225	-617.16711
o.p’-Bis A	-0.039	-5.90125	-0.228	1.766342	-0.068	-0.213	-0.111	0.36733	0.318	-731.69522
MH Bis F	-0.069	-6.11730	-0.247	1.445198	-0.039	-0.208	-0.479	0.36065	-0.229	-577.85490
MM1	-0.057	-5.95105	0.303	1.439828	-0.059	-0.211	-0.287	0.36842	-0.205	-692.38812
Bis F	-0.069	-5.98996	0.302	1.4703	-0.069	-0.209	-0.476	0.37013	-0.209	-653.07586
DM HPTE	-0.062	-6.07676	0.321	2.302561	-0.073	-0.197	-0.303	0.42385	-0.2	-2149.78590
1844-00-44*	-0.051	-5.92819	0.302	0.632798	-0.051	-0.212	-0.269	0.37605	-0.21	-771.01496
Mono MxyBis A	-0.066	-5.67513	0.335	0.943963	-0.063	-0.199	-0.07	0.36812	-0.203	-771.01928
TC Bis A	-0.03	-6.55785	0.283	3.5848	-0.03	-0.224	-0.1	0.37544	-0.224	-2570.07512
TB Bis A	-0.029	-6.51676	0.281	3.5184	-0.029	-0.221	-0.1	0.38069	-0.221	-11016.21701
MxyBis A*	-0.051	-5.79866	0.306	0.9166	-0.051	-0.212	-0.113	0.37028	-0.212	-810.30816

The network training start with the 10 independent variables identified above. The MSE values of the ANN models with the number of neurons on the hidden layer varies from 1 to 10 are analyzed. Compare the MSE values of the models, we could see that the ANN network with 9 neurons on the hidden layer provide the most appropriate MSE value for training set, testset and validation set (R^2 = 0.99, Q^2 = 0.98, R_{test}^2 = 0.98) (Fig. 4).

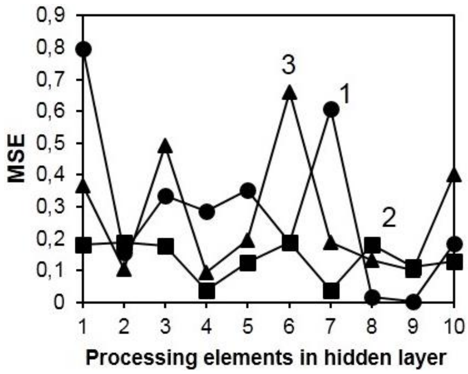


Fig. 4. Variabilities of MSE by number of neurons on the hidden layer
Рис. 4. Валидация MSE от числа элементов скрытого слоя

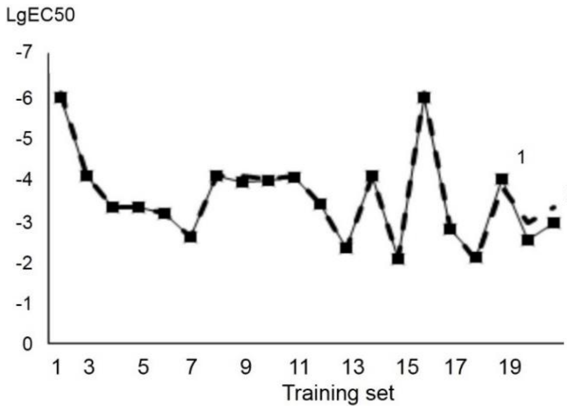
The calculated and predicted activity values for the ANN are given in Table 4.

Table 4
Data on the biological activity of the substances
***Таблица 4.* Данные о биологической активности исследованных веществ**

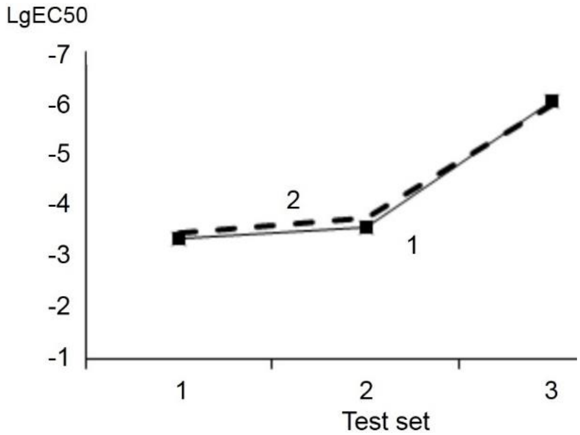
Compound	$LgEC_{50}$ (Observed) [10]	$LgEC_{50}$(Predicted)
DMB Bis A	-2.03	-2.05
HPTE	-3.37	-3.35
MM4	-2.28	-2.30
DM DMB Bis A	-1.99	-2.07
HF Bis A	-2.79	-2.77
Bis B	-3.28	-3.30
DM Bis A	-3.31	-3.29
P Bis A	-4.05	-4.03
MM2*	-3.57	-3.76
Bis A	-2.56	-2.57
PCP	-4.05	-4.04
TM Bis A	-3.8	-3.99
MH MM1	-4.05	-4.03
o.p’-Bis A	-3.96	-3.93
MH Bis F	-4.05	-4.04
MM1	-3.15	-3.16
Bis F	-3.28	-2.91
DM HPTE	-2.91	-2.50
1844-00-44*	-3.38	-3.48
Mono Mxy Bis A	-4.04	-3.90
TC Bis A	-6.04	-5.94
TB Bis A	-6.04	-5.96
Mxy Bis A*	-6.04	-5.96

* - Molecule in test set
* - Молекула в тестовом наборе

The correlation between the predicted, estimated and observed values are demonstrated on the graphs in Fig. 5.



a



b

Fig. 5. Correlation between predicted and observed values:
Рис. 5. Сопоставление прогнозированных значений биоактивности с экспериментальными данными:

The QSAR model built by ANN method has a good predictability with R^2_{test} = 0.98. This model not only allows the establishment of a quantitative relationship between the structure and estrogen activity of molecules, but also helps to identify the parameters that have a great effect on the activity. According to the calculation results these parameters are C11, E_{HOMO}, C3, μ, C13, C12 and C6. Furthermore, a structural change often entail a rigorous and general change of many quantum parameters. Therefore, the increase or decrease of a particular parameter in the structures of BPA and its derivatives does not reflect consistently in the change of estrogen activity. Based on the established QSAR models, it is necessary to analyze the change of the parameter of several molecules with high structural similarities. Thereby, it is possible to select optimal substituents on the molecular frame to establish new molecules with better biological response.

CONCLUSION

The QSAR study of BPA and its derivatives in this article is done through quantum chemistry calculations using B3LYP/6-31+G(d) method combined with modern data processing (ANN method). The survey results obtained a sustainable QSAR model with a determination coefficient $R^2 = 0.99$. The generalization and exogenous capabilities of this model are at a high level with a generalization coefficient $Q^2 = 0.98$; $R^2_{test} = 0.98$, so that this model can be applied in practice to predict the activity of BPA derivatives that haven't been studied.

REFERENCES
ЛИТЕРАТУРА

1. **Jalali-Heravi M.** Neural Networks in Analytical Chemistry. In: D. Livingstone (Ed.) Artificial Neural Networks. Humana Press. 2009. P. 78-118. DOI: 10.1007/978-1-60327-101-1_6.
2. **Morgan M.K., Jones P.A., Calafat A.M., Ye X., Croghan C.W., Chuang J.C., Wilson N.K., Clifton M.S., Figueroa Z., Sheldon L.S.** *Environ. Sci. Technol.* 2011. 45. P. 5309. DOI: 10.1021/es200537u.
3. **Schecter A., Malik N., Haffner D., Smith S., Harris T.R., Paepke O., Birnbaum L.** *Environ. Sci. Technol.* 2010. 44. P. 9425. DOI: 10.1021/es102785d.
4. **Vandenberg L.N., Hauser R., Marcus M., Olea N., Welshons W.V.** *Reproduct. Toxicol.* 2007. 24. P. 139–177. DOI: 10.1016/j.reprotox.2007.07.010.
5. **Krishnan A.V., Stathis P., Permuth S.F., Tokes L., Feldman D.** *Endocrine Soc.* 1993. V. 132. N 6. DOI: 10.1289/ehp.95103s7129.
6. **Rezg R, El-Fazaa S, Gharbi N, Mornagui B.** Bisphenol A and human chronic diseases: Current evidences, possible mechanisms, and future perspectives. *Envir. Internat.* 2014. 64. P. 83–90. DOI: 10.1016/j.envint.2013.12.007.
7. **Doerge D.R., Twaddle N.C., Vanlandingham M., Brown R.P., Fisher J.W.** *Toxicol. Appl. Pharmacol.* 2011. 255. P. 261. DOI: 10.1016/j.taap.2011.07.009.
8. **Ho S.M., Tang W.Y., Belmonte de Frausto J., Prins G.S.** Developmental exposure to estradiol and bisphenol A increases susceptibility to prostate carcinogenesis and epigenetically regulates phosphodiesterase type 4 variant 4. *Cancer Res.* 2006. 66 (11). P. 5624–32. DOI: 10.1158/0008-5472.CAN-06-0516.
9. **Rochester J.R., Bolden A.L.** Bisphenol S and F: A Systematic Review and Comparison of the Hormonal Activity of Bisphenol A Substitutes. *Environ. Health Perspect.* 2015. 123(7). P. 643-650. DOI: 10.1289/ehp.1408989.
10. **Kelly P.C., William A.T., Thomas E.W.** QSAR models of the in vitro estrogen activity of bisphenol A analogs. *QSAR Comb.Sci.* 2003. 22. P. 78—88. DOI: 10.1007/s11434-006-0287-1.
11. **Zhao Y., Truhlar D.** The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Account.* 2008. 120 (1-3). P. 215-241. DOI: 10.1007/s00214-007-0310-x.
12. **Frisch M.J.T.** Gaussian 09. Revision D.01. Gaussian, Inc. Wallingford CT. 2009.
13. **Hmamouchi R., Larif M., Adad A., Bouachrine M., Lakhlifi T.** QSAR analysis of the toxicity of phenols and thiophenols using MLR and ANNJ. *Comp. Meth. Mol. Des.* 2014. 4(3). P. 61-71. DOI: 10.1016/j.jtusci.2016.03.002.
14. **Andrienko G.A.** Chemcraft. https://www.chemcraftprog.com.
15. **Devillers J.** Strengths and Weaknesses of the Backpropagation Neural Network in QSAR and QSPR Studies. In: J. Devillers (Ed.) Neural Networks in QSAR and Drug Design. London: Academic Press. 1996. P. 1-46. DOI: 10.1016/B978-012213815-7/50002-9.
16. **Hornik K.** AppSroximation capabilities of multilayer feedforward networks. *Neural Networks*. 1991. 4. P. 251-257. DOI: 10.1016/0893-6080(91)90009-T.
17. **Balabin R.M., Safieva R.Z., Lomakina E.I.** Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. *Chemometr. Intell. Lab.* 2007. 88 (2). P. 183–188. DOI: 10.1016/j.chemolab.2007.04.006.
18. Tahmasebi P., Hezarkhani A. Application of a Modular Feedforward Neural Network for Grade Estimation. *Nat. Resour. Res.* 2011. 20. (1). P. 25–32. DOI: 10.1007/s11053-011-9135-3.
19. **Golbraikh A., Tropsha A.** Beware of q2. *J. Mol. Graphics Model.* 2002. 20. P. 269–276. DOI: 10.1023/A:1025386326946.
20. **Steyerberg E.W., Bleeker S.E., Moll H.A., Grobbee D.E., Moons K.G.** Internal and external validation of predictive models: A simulation study of bias and precision in small samples. *J. Clinic. Epidemiol.* 2003. V. 56. N 5. P. 441-447. DOI: 10.1016/S0895-4356(03)00047-7.

Поступила в редакцию 06.12.2018
Принята к опубликованию 07.03.2019

Received 06.12.2018
Accepted 07.03.2019