APTEFF, 44, 1-321 (2013)
DOI: 10.2298/APT1344249K
UDC: 547.78:615.282.84:004.41
BIBLID: 1450-7188 (2013) 44, 249-258
Original scientific paper

NEURAL NETWORK MODELLING OF ANTIFUNGAL ACTIVITY OF A SERIES OF OXAZOLE DERIVATIVES BASED ON *IN SILICO* PHARMACOKINETIC PARAMETERS

Strahinja Z. Kovačević, Sanja O. Podunavac Kuzmanović, Lidija R. Jevrić and Nataša D. Kalajdžija

University of Novi Sad, Faculty of Technology Novi Sad, Bulevar cara Lazara 1, 21000 Novi Sad, Serbia

In the present paper, the antifungal activity of a series of benzoxazole and oxazolo[4,5-b]pyridine derivatives was evaluated against Candida albicans by using quantitative structure–activity relationships chemometric methodology with artificial neural network (ANN) regression approach. In vitro antifungal activity of the tested compounds was presented by minimum inhibitory concentration expressed as log(1/c_{MIC}). In silico pharmacokinetic parameters related to absorption, distribution, metabolism and excretion (ADME) were calculated for all studied compounds by using PreADMET software. A feedforward back-propagation ANN with gradient descent learning algorithm was applied for modelling of the relationship between ADME descriptors (blood-brain barrier penetration, plasma protein binding, Madin-Darby cell permeability and Caco-2 cell permeability) and experimental log(1/c_{MIC}) values. A 4-6-1 ANN was developed with the optimum momentum and learning rates of 0.3 and 0.05, respectively. An excellent correlation between experimental antifungal activity and values predicted by the ANN was obtained with a correlation coefficient of 0.9536.

KEY WORDS: benzoxazoles, oxazolo[4,5-*b*]pyridines, *Candida albicans*, artificial neural networks, *in silico* ADME properties

INTRODUCTION

Mycoses are a consequence of infection of organism by fungi of pathogenic potential. Most fungi have developed complete resistance towards antimicrobial therapeutics, therefore the medical treatment of these infections can be long-lasting, and in some cases unsuccessful. Most of compounds, that exert an inhibitory effect on the fungi pathogenic to man, are relatively toxic, so their usage as therapeutic agents is limited (1-3). The need for new and more efficient antifungal drugs is becoming critical because of the increasing number of detected cases of systemic mycoses in patients suffering from immunocompromising diseases (AIDS, diabetes, leukemia, etc). *Candidiasis* is one of the common

* Corresponding author: Strahinja Z. Kovačević, University of Novi Sad, Faculty of Technology Novi Sad, Bulevar cara Lazara 1, 21000 Novi Sad, Serbia. E-mail: strahko@uns.ac.rs

fungal diseases caused by *Candida* species. *Candida albicans* is commonly responsible for infection in humans (2,3). This opportunistic pathogen is resident of normal flora of the gastrointestinal and respiratory mucosis, and can be also identified on vaginal mucosis. The infection develops when the balance between normal bacterial and fungi flora is disrupted. The medical treatment of *Candidiasis* often includes amphotericin B, nystatin, ketoconazole and fluconazole (2). Since the number of aforementioned therapeutics is limited, synthesis and analysis of new drugs with fungistatic activity are very desirable.

Prediction of the antifungal activity of chemical compounds based on experimental data and *in silico* molecular descriptors can provide basic guidelines in the synthesis of new efficient antifungal drugs. Molecular structure and biological reaction can be correlated by using chemometric quantitative structure-activity relationship (QSAR) approach, which enables assessment of the newly-synthesized and unsynthesized compounds (4,5). The QSAR methodology provides the possibility to analyze a large number of molecules for a short time and with minimal costs (6-8). It includes a number of statistical methods, such as multiple linear regression (MLR), principal component regression (PCR), partial least squares regression (PLS), artificial neural networks (ANN), etc.

Benzoxazoles and oxazolo[4,5-*b*]pyridines belong to the group of well-known antifungal agents with antioxidant, antiallergic, antitumoral and antiparasitic activity (1,9). The previous success in the examination of these molecules has stimulated new research based on the synthesis of some new oxazole derivatives.

ANNs are a versatile and flexibile tool for modelling complex relationships between variables. The application of ANN method in QSAR analysis has been presented in earlier studies (10-12). In the present study, the main aim was to develop an ANN model for prediction of antifungal activity of studied benzoxazoles and oxazolo[4,5-*b*]pyridines against *Candida albicans*. In our previous work (9), we have already studied the influence of some molecular descriptors of benzoxazoles on their *in vitro* antifungal activity against *Candida albicans* using MLR. Therefore, the novelties in this study are the series of compounds extended to oxazolo[4,5-*b*]pyridines, the application of ANN on this class of compounds, and prediction based on *in silico* descriptors related to absorption, distribution, metabolism and excretion (ADME).

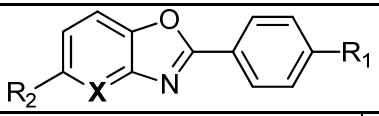
EXPERIMENTAL

Benzoxazoles and oxazolo[4,5-*b*]pyridines

Molecular structures and IUPAC names of 24 studied compounds are presented in Table 1. The results of their *in vitro* antifungal activity against *Candida albicans* (MTCC 183) were taken from the literature (3). Minimum inhibitory concentration (MIC) of tested molecules is defined as the lowest concentration of the compound at which there is no growth of the strain (9). The logarithm of molar MICs, $\log(1/c_{\mathrm{MIC}})$, was calculated and used for further calculations.

APTEFF, 44, 1-321 (2013) UDC: 547.78:615.282.84:004.41
DOI: 10.2298/APT1344249K BIBLID: 1450-7188 (2013) 44, 249-258
Original scientific paper

Table 1. Chemical structures and IUPAC names of studied compounds



No.	IUPAC name	X	R_1	R_2
1	2-phenyl-1,3-benzoxazole	CH	H	H
2	2-(4-*tert*-butylphenyl)-1,3-benzoxazole	CH	$C(CH_3)_3$	H
3	4-(1,3-benzoxazol-2-yl)aniline	CH	NH_2	H
4	4-(1,3-benzoxazol-2-yl)-N-methylaniline	CH	$NHCH_3$	H
5	5-chloro-2-(4-ethylphenyl)-1,3-benzoxazole	CH	C_2H_5	Cl
6	N-[4-(5-chloro-1,3-benzoxazol-2-yl)phenyl]acetamide	CH	$NHCOCH_3$	Cl
7	4-(5-chloro-1,3-benzoxazol-2-yl)-N-methylaniline	CH	$NHCH_3$	Cl
8	5-chloro-2-(4-chlorophenyl)-1,3-benzoxazole	CH	Cl	Cl
9	5-chloro-2-(4-nitrophenyl)-1,3-benzoxazole	CH	NO_2	Cl
10	2-(4-ethylphenyl)-1,3-benzoxazol-5-amine	CH	C_2H_5	NH_2
11	2-(4-fluorophenyl)-1,3-benzoxazol-5-amine	CH	F	NH_2
12	N,N-dimethyl-4-(5-methyl-1,3-benzoxazol-2-yl)aniline	CH	$N(CH_3)_2$	CH_3
13	5-methyl-2-(4-methylphenyl)-1,3-benzoxazole	CH	CH_3	CH_3
14	2-(4-ethylphenyl)-5-methyl-1,3-benzoxazole	CH	C_2H_5	CH_3
15	2-(4-methoxyphenyl)-5-methyl-1,3-benzoxazole	CH	OCH_3	CH_3
16	2-(4-fluorophenyl)-5-methyl-1,3-benzoxazole	CH	F	CH_3
17	N-[4-(5-methyl-1,3-benzoxazol-2-yl)phenyl]acetamide	CH	$NHCOCH_3$	CH_3
18	N-methyl-4-(5-methyl-1,3-benzoxazol-2-yl)aniline	CH	$NHCH_3$	CH_3
19	2-(4-methylphenyl)-[1,3]oxazolo[4,5-*b*]pyridine	N	CH_3	H
20	2-(4-ethylphenyl)-[1,3]oxazolo[4,5-*b*]pyridine	N	C_2H_5	H
21	2-(4-methoxyphenyl)-[1,3]oxazolo[4,5-*b*]pyridine	N	OCH_3	H
22	2-(4-ethoxyphenyl)-[1,3]oxazolo[4,5-*b*]pyridine	N	OC_2H_5	H
23	4-{[1,3]oxazolo[4,5-*b*]pyridin-2-yl}aniline	N	NH_2	H
24	2-(4-nitrophenyl)-[1,3]oxazolo[4,5-*b*]pyridine	N	NO_2	H

Calculation of ADME descriptors

On the basis of the two-dimensional structures, the ADME properties of studied molecules were calculated by the PreADMET online program for drug discovery (13). The calculated ADME properties of examined oxazoles included the parameters of oral absorption (Madin-Darby cells permeability (MDCK) and Caco-2 cells permeability (Caco-2)), parameter of plasma protein binding (PPB) and parameter of blood-brain barrier (BBB) penetration. These descriptors are very important factors prior to the synthesis of new drug molecules because they give an insight into the metabolism of the substance and its pharmacological potential before its application (14-17). The calculated values of ADME properties are presented in Table 2.

Table 2. *In silico* ADME properties of studied molecules

Compound	Caco-2 (nm/sec)	MDCK (nm/sec)	PPB (%)	BBB (c_{brain}/c_{blood})
1^T	54.058	20.169	100.000	3.475
2^T	56.264	0.687	100.000	0.737
3^T	9.308	49.510	100.000	1.523
4^E	42.785	21.218	98.507	0.408
5^T	55.729	2.041	94.060	0.573
6^T	26.914	9.185	90.834	0.047
7^T	46.766	3.733	90.388	0.695
8^E	46.255	3.730	93.143	1.814
9^E	6.039	0.520	86.501	0.013
10^V	37.232	6.076	100.000	0.080
11^T	4.900	40.535	97.348	0.204
12^V	53.316	0.292	94.927	0.172
13^T	55.456	10.481	100.000	2.197
14^T	54.530	3.251	100.000	0.464
15^T	53.570	3.966	96.979	0.079
16^E	54.272	9.007	100.000	0.247
17^V	46.889	12.191	85.064	0.046
18^E	48.769	5.011	97.501	1.032
19^V	53.144	197.978	91.416	3.417
20^V	54.091	183.702	96.707	1.366
21^E	57.363	55.938	86.787	0.131
22^V	46.553	60.464	90.444	0.019
23^T	24.535	44.142	80.879	0.754
24^T	0.727	40.413	71.096	0.339

V – validation set, E – external test set, T – training set

Application of artificial neural network

For the ANN modelling, Statistica software version 10.0 (18) was applied. The whole set of molecules was divided in three subsets: training, validation, and external test set. The training set contains 50% of studied molecules (12 molecules), while the external test set and validation set contain 25% each (6 + 6 molecules). Before any calculations were made, the data set was normalized by *min-max* normalization method by applying the following equation:

$$y_{norm} = (1 - \Delta^U - \Delta^L) \cdot ((y - y_{min})/(y_{max} - y_{min})) + \Delta^L \quad [1]$$

where y_{norm}, y_{max} and y_{min} are the normalized, maximum and minimum value of dependent variable y, and Δ^U, Δ^L are the values of margins which limit extrapolation ability of the network ($\Delta^U = \Delta^L = 0.01$) (19). The normalization process makes the data suitable for the

training. Namely, without normalization, training process would have been very slow. It is especially useful for modelling where the inputs are generally on widely different scales. Normalized values of input data were within the range from 0.01 to 0.99. Afterwards, the optimal structure of ANN was searched. In order to determine the optimal number of nodes in the hidden layer, the root mean square error (*RMSE*) was plotted as a function of the number of neurons in the hidden layer. The normalized data were processed by feedforward back-propagation neural network with gradient descent learning algorithm. The vaules of weights of neurons were predetermined by the applied software (18).

The output data obtained by ANN were correlated with the experimental values. The correlation between them is characterized by Pearson's correlation coefficient (*r*) and predictive ability by the following parameters: *RMSE*, predicted residual sum of squares (*PRESS*) and relative error of prediction (*RE*).

RESULTS AND DISCUSSION

Optimal neural network architecture

In the first stage of evaluation of neural network, the number of neurons in the hidden layer was determined by trial and error method based on minimum *RMSE* value. From Figure 1, it is evident that 6 neurons in the hidden layer are sufficient to obtain a low *RMSE* value and that a further increase in the number of neurons does not bring any improvement. Figure 2 presents the optimal architecture, which implies 4 input variables, 6 hidden neurons and 1 output variable (4-6-1) of ANN applied in this study.

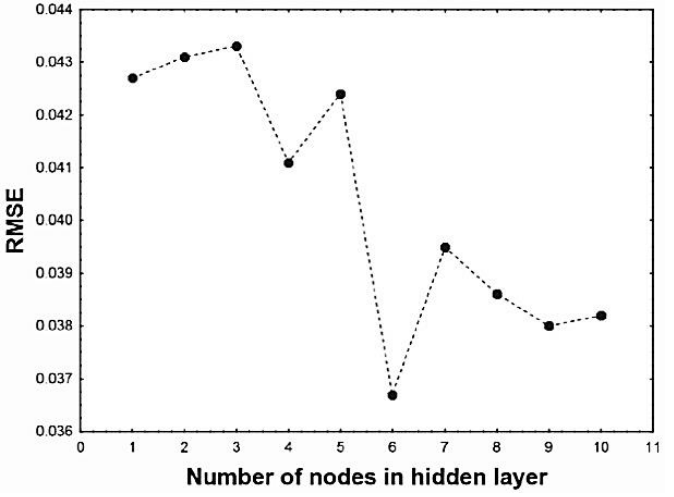


Figure 1. *RMSE* as a function of the number of neurons in the hidden layer

UDC: 547.78:615.282.84:004.41
BIBLID: 1450-7188 (2013) 44, 249-258
Original scientific paper

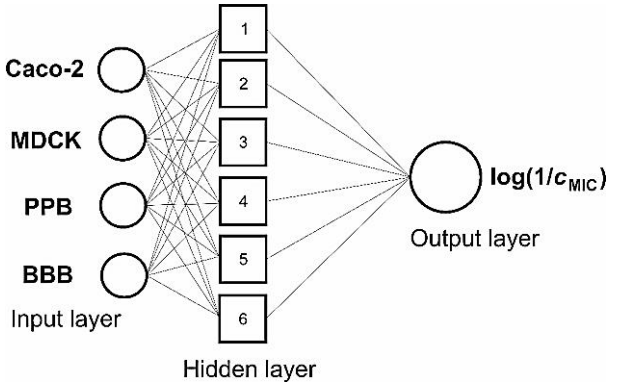


Figure 2. Optimal architecture of the established neural network

It was important to define the training conditions, which take into account the learning rate and the momentum value (20). These parameters were estimated on the basis of the minimum *RMSE* value. The dependences between the learning rate and the *RMSE*, and between the momentum and the *RMSE* are shown in Figure 3.

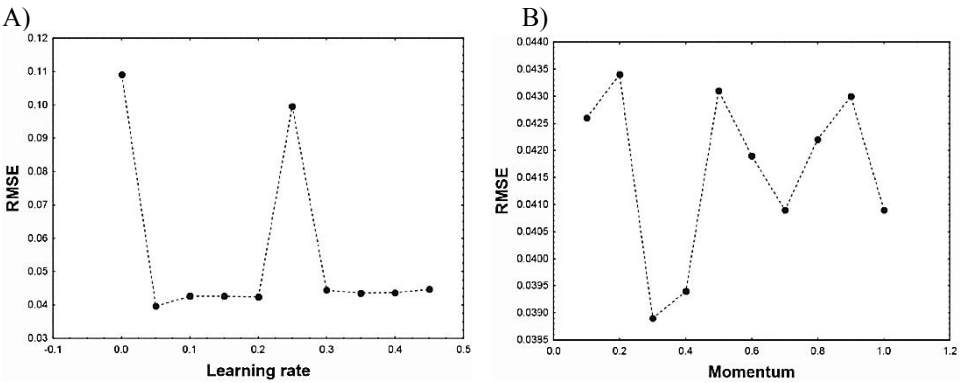


Figure 3. *RMSE* as a function of the learning rate (A) and momentum coefficient (B)

As it can be seen from Figure 3, judging from the minimum values of *RMSE*, the optimal values for learning rate and momentum are 0.05 and 0.3, respectively. In order to avoid over-training, the performance of the ANN was tested every 100 epochs during the training and the weightings for the minimum *RMSE* for the learning and test set were recorded ($RMSE_{training}$ = 0.0002, $RMSE_{test}$ = 0.0019) (Figure 4).

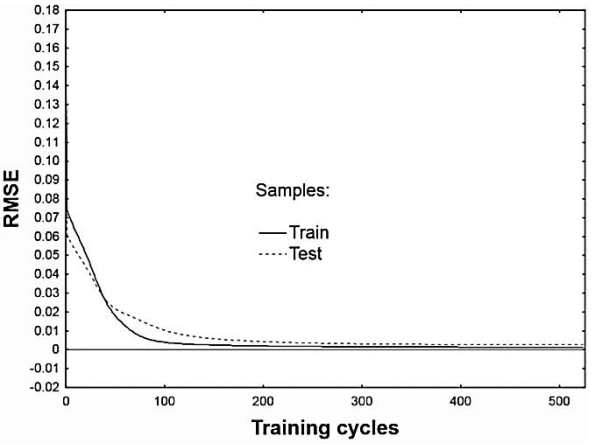


Figure 4. *RMSE versus* training epochs

Prediction of the minimum inhibitory concentration by optimized neural network

After optimization, the actual predictive performance of the trained network was evaluated using the external data set. A comparison between experimental and predicted $\log(1/c_{MIC})$ values for the whole data set is presented in Figure 5.

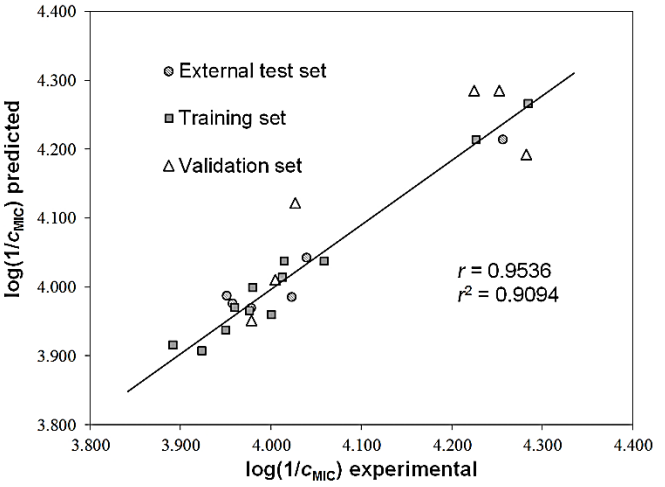


Figure 5. Experimental *versus* predicted $\log(1/c_{MIC})$ values

The low scattering of points around the linear relationship, significant slope (>0.90) and intercept close to zero (<0.03), indicate a very good concurrence between the experimental and predicted data. Statistical measures of the obtained results indicate very good predictivity and accuracy of the ANN model: r = 0.9536, *RMSE* = 0.0369, *PRESS* = 0.0227 and *RE* = 0.6815%. The absolute maximum value of the residuals was 0.095,

therefore the maximum individual percentage deviation (IPD%) was 2.13%. On the basis of the magnitude of the residues and IPD%, there is a close agreement between the experimental and estimated antifungal activities. From the presented results it can be concluded that the established neural network has an excellent predictive power and very high accuracy in the observed range of the $\log(1/c_{MIC})$ values.

The influence of every ADME parameter in the input layer on the variations is estimated by global sensitivity analysis. Sensitivity analysis is used to determine how much 'sensitive' a network is to the changes in the value of the parameters of the model and to the changes in the structure of the network model. The sensitivity coefficients describe the change in the network's outputs due to variations in the parameters that affect the network. A large sensitivity of input parameter suggests that the network's performance can significantly change with small variation in the parameter (21). Conversely, a small value of the sensitivity index suggests a small change in the parameter. The values of the sensitivity index for the input variables are as follows: MDCK (6.5795), PPB (5.6532), Caco-2 (1.4998) and BBB (1.3310). In this case, the largest influence on variability of the $\log(1/c_{MIC})$ values have the MDCK and PPB pharmacokinetic parameters.

CONCLUSION

The presented results indicate that the ANN regression method can be a very useful tool for prediction of antifungal activity of studied benzoxazoles and oxazolo[4,5-*b*]pyridines against *Candida albicans*. They demonstrate a very good concurrence between the experimental and predicted values of the minimum inhibitory concentration of the analyzed compounds. With the selection of suitable learning parameters and optimal neural network topology (4-6-1), we achieved an excellent approximation of experimental results with the estimated ones. This was confirmed by the values of the basic statistical parameters: high correlation coefficient and low values of *RMSE*, *RE* and *PRESS*. Besides, the importance of this paper also lies in successful correlation between the *in silico* ADME properties and antifungal activity of studied compounds. This study could give necessary guidelines for the analyses of new antifungal therapeutics and facilitate new research in drug discovery processes.

Acknowledgements

This paper was performed within the framework of the research projects No.172012 and No.172014, supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia and the project No. 114-451-2373/2011, financially supported by the Provincial Secretariat for Science and Technological Development of Vojvodina.

REFERENCES

1. Şener, E., Yalçin, İ. and Sungur, E.: QSAR of some antifungal benzoxazoles and oxazolo(4,5-b)pyridines against C. Albicans. Quant. Struct.-Act. Relat. **10** (1991) 223-228.

2. Ursu, O., Costescu, A., Diudea, M.V. and Parv, B.: QSARs of some novel isosteric heterocyclic with antifungal activity. Carpatian J. Math. **20** (2004) 267-274.
3. Ursu, O., Costescu, A., Diudea, M.V. and B. Parv.: QSAR modeling of antifungal activity of some heterocyclic compounds. Croat. Chem. Acta **79** (2006) 483-488.
4. Podunavac Kuzmanović, S.O. and Cvetković, D.D.: Lipophilicity and antifungal activity of some 2-substituted benzimidazole derivatives. Chem. Ind. Chem. Eng. Q. **17** (2011) 9-15.
5. Podunavac Kuzmanović, S.O, Cvetković, D.D. and Gadžurić, S.B.: Prediction of the inhibitory activity of benzimidazole derivatives against Bacillus spp. APTEFF **42** (2011) 251-261.
6. Satpathy, R., Guru, R.K. and Behera, R.: Computational QSAR analysis of some physicochemical and topological descriptors of curcumin derivatives by using different statistical methods. J. Chem. Pharm. Res. **2** (2010) 344-350.
7. Podunavac Kuzmanović, S.O, Cvetković, D.D. and Barna, D.J.: The effect of lipophilicity on the antibacterial activity of some 1-benzylbenzimidazole derivatives. J. Serb. Chem. Soc. **73** (2008) 967-978.
8. Podunavac Kuzmanović, S.O. and Cvetković, D.D.: QSAR modeling of antibacterial activity of some benzimidazole derivatives. Chem. Ind. Chem. Eng. Q. **17** (2011) 33-38.
9. Podunavac Kuzmanović, S.O., Jevrić, L.R., Kovačević, S.Z. and Kalajdžija, N.D.: A chemometric approach for prediction of antifungal activity of some benzoxazole derivatives against Candida albicans. APTEFF **43** (2012) 273-282.
10. Wnuk, M., Marszałł, M.P., Zapęcka, A., Nowaczyk, A., Krysiński, J., Romaszko, J., Kawczak, P., Bączek, T. and Buciński, A.: Prediction of antimicrobial activity of imidazole derivatives by artificial neural networks. Cent. Eur. J. Med. **8** (2013) 1-15.
11. Buciński, A., Socha, A., Wnuk, M., Baczek, T., Nowaczyk, A.; Krysiński, J., Goryński, K. and Koba, M.: Artificial neural networks in prediction of antifungal activity of a series of pyridine derivatives against *Candida albicans*. J. Microbiol. Methods **76** (2009) 25-29.
12. Hasegawa, K., Deushi, T., Yaegashi, O., Miyashita, Y. and Sasaki, S.: Artificial neural network studies in quantitative structure-activity relationships of antifungal azoxy compounds. Eur. J. Med. Chem. **30** (1995) 569-574.
13. PreADMET Software, http://preadmet.bmdrc.org/
14. Hou, T.J., Zhang, W., Xia, K., Qiao, X.B. and Xu, X.J.: ADME evaluation in drug discovery. 5. Correlation of Caco-2 permeation with simple molecular properties. J. Chem. Inf. Comput. Sci. **44** (2004) 1585-1600.
15. Irvine, J. D., Takahashi, L., Lockhart, K., Cheong, J., Tolan, J.W., Selick, H.E. and Grove, R.: MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening. J. Pharm. Sci. **88** (1999) 28-33.
16. Thomas, G.: Fundamentals of medicinal chemistry, John Wiley and Sons, New York (2003).
17. Hou, T.J. and Xu, X.J. ADME evaluation in drug discovery. 3. Modeling blood-brain barrier partitioning using simple molecular descriptors. J. Chem. Inf. Comput. Sci. **43** (2003) 2137-2152.
18. StatSoft Inc., 2300 East 14th Street, Tulsa, Oklahoma, USA: http://www.statsoft.com/

19. Jayalakshmi, T. and Santhakumaran, A.: Statistical normalization and back propagation for classification. IJCTE **3** (2011) 89-93.
20. Jokić, A.I., Grahovac, J.A., Dodić, J.M., Dodić, S.N., Popov, S.D., Zavargo, Z.Z. and Vučurović, D.G.: Artificial neural network approach to modelling of alcoholic fermentation of thick juice from sugar beet processing. Hem. Ind. **66** (2012) 211-221.
21. Shojaeefard, M.H., Akbari, M., Tahani, M. And Farhani, F.: Sensitivity analysis of the artificial neural network outputs in friction stir lap joining of aluminium to brass. Adv. Mater. Sci. Eng. (2013) 7 pg. doi:10.1155/2013/574914

МОДЕЛОВАЊЕ АНТИФУНГАЛНЕ АКТИВНОСТИ ДЕРИВАТА ОКСАЗОЛА НА ОСНОВУ СОФТВЕРСКИ ИЗРАЧУНАТИХ ФАРМАКОКИНЕТИЧКИХ ПАРАМЕТАРА ПОМОЋУ РЕГРЕСИОНЕ МЕТОДЕ ВЕШТАЧКИХ НЕУРОНСКИХ МРЕЖА

Страхиња З. Ковачевић, Сања О. Подунавац Кузмановић, Лидија Р. Јеврић и Наташа Д. Калајџија*

Универзитет у Новом Саду, Технолошки факултет Нови Сад, Булевар цара Лазара 1, 21000 Нови Сад, Србија

Применом методе вештачких неуронских мрежа испитана је могућност предвиђања минималне инхибиторне концентрације, $\log(1/c_{MIC})$, једињења бензоксазола и оксазоло[4,5-*b*]пиридина према *Candida albicans* на основу софтверски израчунатих дескриптора везаних за апсорпцију, дистрибуцију, метаболизам и екскрецију (АДМЕ). Оптимализована нерекурентна неуронска мрежа са једним скривеним слојем састављеним од 6 неурона, улазним слојем од 4 и излазним слојем од 1 неурона, предвидела је вредности за $\log(1/c_{MIC})$ које су у веома јакој корелацији са вредностима добијеним експериментално. Статистичка спецификација примењене мреже указује на успешно предвиђање антифунгалне активности на основу софтверски моделованих АДМЕ параметара испитиваних једињења употребом вештачких неуронских мрежа, што може да представља добру смерницу код дизајнирања нових антифунгалних супстанци.

Кључне речи: бензоксазоли, оксазоло[4,5-*b*]пиридини, *Candida albicans*, вештачке неуронске мреже, АДМЕ карактеристике

Received: 27 May 2013.
Accepted: 01 July 2013.