

**DESIGN OF LOW CYTOTOXICITY DIARYLANILINE DERIVATIVES  
BASED ON QSAR RESULTS: AN APPLICATION OF ARTIFICIAL NEURAL  
NETWORK MODELLING**

**PERANCANGAN TURUNAN DIARILANILINA DENGAN SITOTOKSISITAS  
RENDAH BERDASARKAN HASIL HKSA:  
PENGUNAAN PEMODELAN JARINGAN SYARAF TIRUAN**

**Ihsanul Arief<sup>1\*</sup>, Ria Armunanto<sup>2</sup>, Bambang Setiaji<sup>2</sup>, Muhammad Fachrie<sup>3</sup>**

<sup>1</sup>Akademi Farmasi Yarsi Pontianak, Kalimantan Barat, Indonesia

<sup>2</sup>Austrian-Indonesian Centre (AIC) for Computational Chemistry, Department of Chemistry,  
Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia

<sup>3</sup>Informatics Engineering Major, Faculty of Business and Information Technology  
Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia

\*email : ihsanul.arief@gmail.com

*Received 1 June 2016; Accepted 12 October 2016; Available online 29 November 2016*

**ABSTRACT**

Study on cytotoxicity of diarylaniline derivatives by using quantitative structure-activity relationship (QSAR) has been done. The structures and cytotoxicities of diarylaniline derivatives were obtained from the literature. Calculation of molecular and electronic parameters was conducted using Austin Model 1 (AM1), Parameterized Model 3 (PM3), Hartree-Fock (HF), and density functional theory (DFT) methods. Artificial neural networks (ANN) analysis used to produce the best equation with configuration of input data-hidden node-output data = 5-8-1, value of  $r^2 = 0.913$ ; PRESS = 0.069. The best equation used to design and predict new diarylaniline derivatives. The result shows that compound N1-(4'-Cyanophenyl)-5-(4"-cyanovinyl-2",6"-dimethyl-phenoxy)-4-dimethylether benzene-1,2-diamine) is the best-proposed compound with cytotoxicity value ( $CC_{50}$ ) of 93.037  $\mu$ M.

**Keywords** : diarylaniline, cytotoxicity, QSAR, ANN

**ABSTRAK**

Kajian terhadap sitotoksitas turunan diarilanilina menggunakan hubungan kuantitatif struktur-aktifitas (HKSA) telah dilakukan. Struktur dan sitotoksitas dari turunan diarilanilina diperoleh dari literatur. Perhitungan parameter molekuler dan elektronik dilakukan dengan metode *Austin Model 1* (AM1), *Parameterized Model 3* (PM3), Hartree-Fock (HF), dan *density functional theory* (DFT). Analisis jaringan syaraf tiruan (JST) digunakan untuk menghasilkan persamaan terbaik dengan konfigurasi dari *input data-hidden node-output data* = 5-8-1, nilai  $r^2 = 0,913$ ; PRESS = 0,069. Persamaan terbaik tersebut kemudian digunakan untuk merancang dan memprediksi senyawa-senyawa turunan diarilanilina yang baru. Hasil yang diperoleh menunjukkan bahwa senyawa N<sup>1</sup>-(4'-sianofenil)-5-(4"-sianofinil-2",6"-dimetil-fenoksi)-4-dimetileter benzena-1,2-diamina) adalah senyawa usulan terbaik dengan nilai sitotoksitas ( $CC_{50}$ ) sebesar 93,037  $\mu$ M.

**Kata kunci** : diarilanilina, sitotoksitas, HKSA, JST

**INTRODUCTION**

Toxicity is still a major safety concern for drug withdrawal, the 'black box warning', and the discontinuation of clinical trials (such as Pfizer's hypercholesterolemia drug torcetrapib

withdrawal from Phase III). An analysis of the first-in-human registration for ten big pharmacy companies demonstrated only 10% of total success rate leading to the final FDA approval. The failure rate becomes even higher when all drug candidates in preclinical research are

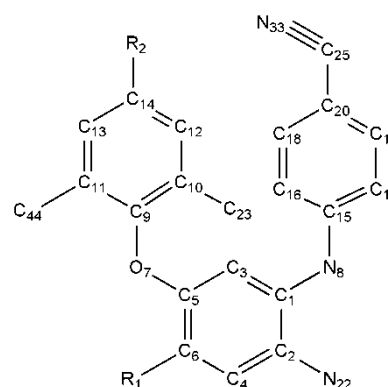
included in the statistics. The traditional drug safety testing approaches include in vivo animal models and in-vitro cell-based assays, and most recently in silico assessment is also introduced. QSAR-based expert systems are mainly used in early drug discovery to predict toxicological endpoints (Sun and Scott, 2010).

Currently, molecular modeling and computational chemistry is an indispensable part of the search and drug design. The computational method could save time and money to find new drugs. Among the computation methods in drug design, QSAR is the most widely used. QSAR method conducts a study of a relationship between molecular and electronic parameters of the activity or toxicity in the series of analog compounds. These parameters are obtained from calculations using quantum mechanical methods that have been developed (Hemmateenejad et al., 2009). In some earlier studies, multiple linear regression methods were not able to provide a good model, so that the artificial neural network method was used to produce the non-linear models.

Ekins and Williams (2012) mentioned that the potential to predict human toxicity directly from a molecular structure is feasible. By using the experimental properties of known compounds as the basis of predictive models, it is possible to develop structure activity relationships and resulting algorithms related to toxicity. Gacche and Jadhav (2012) reported the results of their research to make a model of the toxicity of the coumarin derivatives and their molecular parameters. Hosseini et al. (2013) have conducted a study by modeling cytotoxicities of the substituted amides of pyrazine $\beta$ -carboxylic acids versus their molecular parameters with the best model has the value of  $r^2 = 0.922$ . Low et al. (2011) reported their study on predicting drug-induced hepatotoxicity using QSAR and toxicogenomics

approaches and had external predictivity as 76%. Ruiz et al. (2012) had obtained a relatively good model of acute mammalian toxicity using QSAR method, with the best model has the  $r^2$  value of 0.929 (T.E.S.T model).

Sun et al. (2012) reported the results of their research were synthesized and tested the anti-HIV activities and cytotoxicities of twenty derivatives of diarylaniline (DAAN) as the parent structure which is showed in **Figure 1**. Their results are promising, which have the lower value of  $EC_{50}$  and higher value of  $CC_{50}$  than the control drug (rilpivirine). QSAR modeling study on anti-HIV activity of the DAAN derivatives has been done in our previous research (Arief et al., 2013). Thus, in this study, we performed QSAR modeling on cytotoxicities of DAAN derivatives to design the new compound with the lower cytotoxicity.



**Figure 1.** Parent structure of diarylaniline derivative

## COMPUTATIONAL METHODS

### Data Set

The total set of compounds (**Table 1**) was divided into a training set (15 compounds) for generating QSAR models and a test set (5 compounds) for validating the quality of the models. Selection of molecules in the training set and the test is a key and important feature of any QSAR model. Therefore, the care was taken in such a way that biological activities of all compounds in test lie within the maximum and minimum value range of biological activities of the training set of compounds.

**Table 1.** Molecular structures of compounds and their cytotoxicities (Sun et al., 2012)

No	R <sub>1</sub>	R <sub>2</sub>	CC <sub>50</sub> (μM)
13a <sup>a</sup>	CF <sub>3</sub>	CH=CHCOCH <sub>3</sub>	8.60
13b <sup>a</sup>	CF <sub>3</sub>	CH <sub>2</sub> CH <sub>2</sub> COCH <sub>3</sub>	11.10
13c <sup>b</sup>	CF <sub>3</sub>	CN	13.30
13d <sup>b</sup>	CF <sub>3</sub>	NH <sub>2</sub>	15.80
13e <sup>a</sup>	CH <sub>2</sub> OH	CH <sub>2</sub> OH	15.90
14a <sup>a</sup>	CF <sub>3</sub>	CH=CHCN	15.00
14b <sup>a</sup>	SO <sub>2</sub> NH <sub>2</sub>	CH=CHCN	21.80
14c <sup>a</sup>	COOCH <sub>3</sub>	CH=CHCN	22.80
14d <sup>a</sup>	COOH	CH=CHCN	14.90
14e <sup>b</sup>	CONH <sub>2</sub>	CH=CHCN	23.60
14f <sup>b</sup>	CONHCH <sub>3</sub>	CH=CHCN	9.20
14h <sup>a</sup>	CH <sub>2</sub> OH	CH=CHCN	16.30
15a <sup>a</sup>	CF <sub>3</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	7.10
15b <sup>a</sup>	SO <sub>2</sub> NH <sub>2</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	11.10
15c <sup>b</sup>	COOCH <sub>3</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	22.78
15d <sup>a</sup>	COOH	CH <sub>2</sub> CH <sub>2</sub> CN	14.30
15e <sup>a</sup>	CONH <sub>2</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	9.40
15f <sup>a</sup>	CONHCH <sub>3</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	9.10
15g <sup>a</sup>	CONHNH <sub>2</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	22.72
15h <sup>a</sup>	NH <sub>2</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	25.10
2	Rilpivirine		19.40

<sup>a</sup> Training set<sup>b</sup> Test set

The Uni-Column Statistics of test and training sets further reflected the correct selection of test and training sets.

The maximum and minimum value in training and test set were compared in a way that (Jain et al., 2012):

1. The maximum value of CC<sub>50</sub> of test set should be less than or equal to a maximum value of CC<sub>50</sub> of the training set.
2. The minimum value of CC<sub>50</sub> of test set should be higher than or equal to minimum value of CC<sub>50</sub> of training set.

*In vitro* cytotoxicity concentration (CC<sub>50</sub>) of the molecules were converted into corresponding log CC<sub>50</sub> values and used as dependent variables in QSAR calculations.

### Descriptor Calculation

The basis of energy minimization is that the drug binds to effectors/receptors in the most stable form, i.e., the minimum energy form. QSAR study requires the calculation of molecular descriptors. In

this study, the methods which have been used to optimize the structural geometries of data set are Austin Model 1 (AM1), Parameterized Model 3 (PM3), Hartree-Fock (HF), and density functional theory (DFT) on Gaussian 09W package (Frisch et al., 2009).

A large number of theoretical molecular descriptors such as surface area, volume, hydration energy, log P, refractivity, polarisability, molecular mass, HOMO-energy, LUMO-energy, ΔE (EHOMO-ELUMO) calculated by Hyperchem 8.0.10 package (HyperCube, Inc., 2011); electronic descriptors such as C1, C2, C3, C4, C5, C6, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C23, C25, C44, N8, N22, N33, O7 have been computed for these geometrically optimized structures from the chemical structures of the compounds referred to above with a view to develop structure-activity relationship of diarylaniline (DAAN) derivatives.

## Model Development

The QSAR model was generated previously by Multiple Linear Regression (MLR) Backward method by using SPSS Release 19.0.0 package (IBM, 2010). It relates the dependent variable  $\hat{y}$  (biological activity) to a number of independent variables  $x_i$  (molecular descriptor) by using linear equations. This regression method estimates the values of the regression coefficients by applying least square curve fitting method. MLR is the traditional and standard approach for multivariate data analysis. The best model was chosen based on some statistical parameters such as  $r^2$ , standard estimation of error (SEE), F-ratio between the variance of predicted and observed activity, and PRESS (Podunavac-Kuzmanović et al., 2009), where :

$$\text{PRESS} = (\text{predicted value} - \text{observed value})^2$$

Because of the linear analysis method produced a model that does not pass the validation test, then it proceeds with the analysis of the non-linear form of artificial neural networks (ANN) as performed by Deeb and Jawabreh (2012) using MATLAB package. On ANN analysis, examination of data outliers in advance. An examination carried out by plotting the first principal component values and the value of the second component, which is then observed the distribution of data.

If there are points that apart from most of the other, then the points are considered as a data outlier and not included in the ANN analysis. ANN analysis was conducted by normalizing parameter data, which is then calculated as the input data. A number of hidden nodes used in the range of 3-15, where the sigmoid used as the activation function.

The value of the mean square error (MSE) is  $1 \times 10^{-6}$ .

## Model Validation

The best model which has chosen then used to predict  $\log CC_{50}$  values of the test set. A model can be determined as validated if passed some of the criteria such as  $r^2_{\text{pred}} > 0.5$ ,  $r^2_m > 0.5$ , where (Hu et al., 2009):

$$r^2_m = r^2 \left( 1 - \sqrt{|r^2 - r^2_0|} \right)$$

$r^2$  and  $r^2_0$  between the observed and predicted values are calculated from the test set with and without intercept, respectively.

## Design and Toxicities Prediction of New Compounds

Based on the validated model, new compounds of diarylaniline derivatives has been designed by replacing the substituent. Those new compounds then optimized and calculated their descriptors. The descriptors which have been calculated used to predict  $CC_{50}$  values of new compounds using the validated model. The best new compound has chosen based on its  $CC_{50}$  value which is higher than Rilpivirine.

## RESULT AND DISCUSSION

### Training Set and Data Set

A Uni-Column statistics for the training set and test set were generated to check the correctness of selection criteria for training and test set molecules (**Table 2**). Training set and test set were checked using a uni-column statistics as listed in **Table 2**. **Table 2** shows that average and standard deviation values of training and test set are not different significantly, indicating a similar data distribution in both.

**Table 2.** Uni-column statistics of the training and test sets for QSAR models

Set	Number of data	Mean	Maximum	Minimum	St. Dev	Sum
Training	15	1.145	1.398	0.851	0.172	17.177
Test	5	1.203	1.373	0.964	0.170	6.015

**Table 3.** Statistical parameters of 4 QSAR models of diarylaniline derivatives

Model	Descriptor	$r^2$	SEE	$F_{cal}/F_{tab}$	PRESS
AM1	qC2, qC4, qC14, qC23, qN33	0.721	0.114	1.335	0.116
PM3	qC1, qC4, qC10, qC18, qC19	0.778	0.101	1.816	0.092
HF	qC16, qC23, qN22, qN33, qO7	0.326	0.176	0.250	0.280
DFT	qC3, qC6, qC10, qC20, qO7	0.452	0.159	0.427	0.228

$r^2$  = linearity coefficient  
 SEE = standard error of estimation  
 $F_{cal}/F_{tab}$  = ration between F-value calculated to F-value in F-table;  $df_1 = 5$ ,  $df_2 = 9$   
 PRESS = prediction sum of squares

### Model Development

The process used in statistical analysis to develop the model was the backward method. This method used a principle that, in the first step of the analysis, all of the descriptors were included in model development. On the next step, non-significant descriptors excluded from the model and then regression parameters was recalculated. This procedure has been done continuously until the simpler model obtained (with fewer descriptor), but still has approved significantly (less than 0.05).

Based on the values of  $r^2$  ( $>0.6$ ) and  $F_{cal}/F_{tab}$  ( $>1$ ) from each model, only AM1 and PM3 models are eligible to proceed to the model validation.

### Model Validation

The models that passed the requirement then used to predict the value of  $\log CC_{50}$  of the data in the test set. The result is shown in **Table 4**. The data showed in **Table 4**, then plotted to check the other parameter, the value of  $r^2_{pred}$ . Unfortunately, the model of AM1 and PM3 has very low values of  $r^2_{pred}$ , they

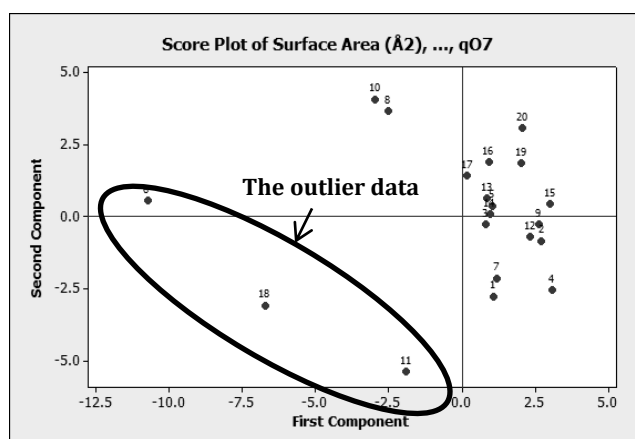
are 0.059 and 0.003 respectively. Those values were not qualified to express QSAR models to be valid ( $r^2_{pred} > 0.5$ ). This shows that both models resulted from the analysis by linear methods can not predict the compounds outside the training set well, so we need another analysis to gain a better model. In this research, advanced analysis for the study of toxicity in the form of non-linear analysis of artificial neural networks (ANN).

Analysis method with ANN occurred to the independent variables and the dependent variable obtained from the previous MLR analysis. ANN analysis conducted in the process of finding the value of the "weights" that describes the relationship of each layer so as to produce an output value. Furthermore, the back-propagation process generates the less error value continuously (iteration).

Before analyzing with ANN, the data were checked for the outliers of the twentieth data that would not use by principle component analysis (PCA) as done by Deeb and Drabh (2010), Deeb and Jawabreh (2012). For example, the result of data analysis for outliers in AM1 data showed in **Figure 2**.

**Table 4.** Comparison of observed and predicted values of  $\log CC_{50}$ 

Compound	Observed $\log CC_{50}$	Predicted $\log CC_{50}$	
		AM1	PM3
14f	0.964	1.165	1.112
13c	1.124	2.021	1.098
13d	1.199	2.427	0.696
15c	1.356	0.909	1.218
14e	1.373	1.194	1.043
	PRESS	2.586	0.403



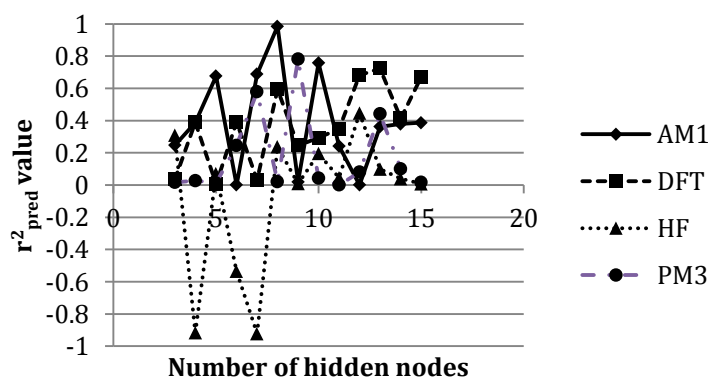
**Figure 2.** The result of outlier check by using principle component analysis

It can be observed that the data with the numbers 6, 11, and 18 on the order of compounds in the total series is regarded as outlier data, so it was not included in the later analysis of ANN. Examination of outliers data was also performed on three other models. Outliers data may interfere with the results of modeling because it has a different tendency with most of the other data. Exclusion of data outliers can improve the yield model, as reported by Eroglu et al. (2007) the MLR models that originally had  $r^2$  values of 0.837 (using 18 data) amounted to 0.943 (using 16 data). Similarly, the results of the study of Wang et al. (2012) where the value of  $q^2$  increased from 0.258 (using 17 data) into 0.582 (using 15 data).

On ANN analysis, the amount of nodes in the hidden layer was varied with ranges of 3-15 to find the best model. The comparison of the  $r^2_{\text{pred}}$  value of the test set to the number of nodes in the hidden layer is shown in **Figure 3**. **Figure**

**3** showed that each model has a different configuration to generate the highest  $r^2_{\text{pred}}$  values. In general, only the three models have the  $r^2_{\text{pred}}$  value which are qualified validity ( $> 0.5$ ). While the HF models do not meet, because it has a value of  $r^2_{\text{pred}}$  only 0.445. The highest  $r^2_{\text{pred}}$  value obtained on the AM1 model (0.985) with the configuration of the I-H-O = 5-8-1. The number of hidden nodes as much as 8 is the smallest value (the simplest configuration), in which the other models require a number of hidden nodes is 13 (DFT model), 9 (PM3 model), 12 (HF model). Therefore, the AM1 model regarded as the best model for predicting the toxicity of diarylaniline derivative compounds.

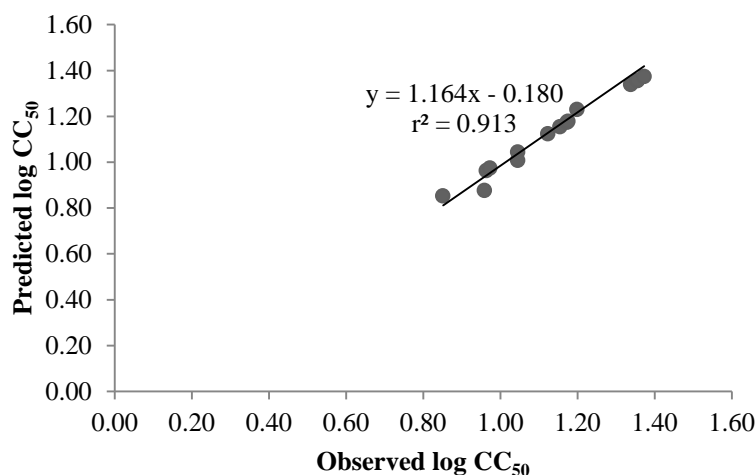
For additional validity test, it was performed the determination  $r^2_m$  and PRESS value for each model. The result of ANN analysis to four of the best models can be seen in **Table 5**.



**Figure 3.** Result of artificial neural network modeling

**Table 5.** Additional statistical parameter from models by ANN

Model	Configuration (I-H-O)	$r^2_{\text{pred}}$	$r^2_{\text{m}}$	PRESS
AM1	5-8-1	0.985	0.574	0.069
PM3	5-9-1	0.781	0.490	0.123
HF	5-12-1	0.445	0.326	0.416
DFT	5-13-1	0.724	0.701	0.189

**Figure 4.** Plot chart of predicted and observed log CC<sub>50</sub> values from AM1 model

**Table 5** shows that the  $r^2_{\text{m}}$  parameter of PM3 models do not meet the minimum threshold ( $> 0.5$ ). Whereas between AM1 and DFT models which are eligible the minimums, DFT models have the higher value of  $r^2_{\text{m}}$  (0.701), but the lower  $r^2_{\text{pred}}$  and higher PRESS value than the AM1 model.

Therefore, AM1 model was concluded as the best model. It supported also by the fact that in AM1 method, geometry optimization requires a shorter time than DFT. AM1 model regarded as the best model showed that in the toxicity studies, calculations involving energy with only the valence electrons can describe the relationship between the parameters and the toxicities of diarylaniline derivatives. While the method of calculation of the overall HF involving electrons (core and valence) and DFT methods are based on the function of the electron density, are less able to describe the relationship well. Results plot between log CC<sub>50</sub> values of observed and the model predictions using AM1 shown in **Figure 4**.

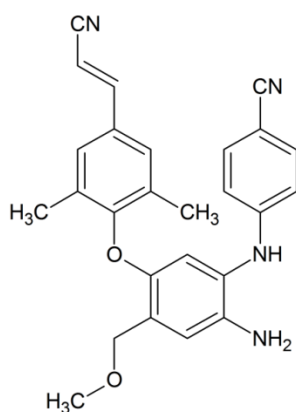
### Design and Activities Prediction of New Compounds

In the design of new compounds, substituent was replaced with other groups or chemical species. The selection of those new species or groups was based on the possibility to be synthesized and the materials availability. It was expected also, the step of the synthesis will be done only in single or double steps to keep the rendemen in a good term. Log CC<sub>50</sub> values of the compounds was determined by the model proposed AM1 with input data values of atomic charge qC2 , qC4 , qC14 , qC23 , qN33 (5-8-1 configuration). **Table 6** shows 10 designed compounds with their predicted log CC<sub>50</sub> and also predicted CC<sub>50</sub> values. In **Table 6**, it can be seen that generally, the best new compounds have substituent groups which potentially to make hydrogen bonding. In the other side, most of the best compounds have R2 groups with a double-bond. This indicates, the double bond in R2 also has a significant effect on cytotoxicity mechanism between DAAN derivatives to human liver microsomes.

**Table 6.** New designed diarylaniline derivative compounds and its predicted log CC<sub>50</sub> and CC<sub>50</sub> using AM1 model

No	R1	R2	Predicted log CC <sub>50</sub>	Predicted CC <sub>50</sub> (μM)
1	CH <sub>2</sub> OCH <sub>3</sub>	CH=CHCN	1.969	93.037
2	CH <sub>2</sub> OCH <sub>2</sub> CH <sub>3</sub>	CH=CHCN	1.967	92.695
3	COCH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	CH=CHCN	1.967	92.695
4	CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> OH	CH=CHCN	1.962	91.721
5	CHCl <sub>2</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	1.954	90.027
6	CH <sub>2</sub> F	CH <sub>2</sub> CH <sub>2</sub> CN	1.952	89.540
7	CONHCH <sub>2</sub> CH <sub>3</sub>	CH <sub>2</sub> CH <sub>2</sub> CN	1.951	89.331
8	CH <sub>2</sub> Cl	CH <sub>2</sub> CH <sub>2</sub> CN	1.948	88.681
9	COCl	CH <sub>2</sub> CH <sub>2</sub> CN	1.938	86.753
10	CH <sub>2</sub> Br	CH <sub>2</sub> CH <sub>2</sub> CN	1.935	86.069

Compound number 1 has predicted CC<sub>50</sub> value of 93.037 μM, which is higher than Rilpivirine's (24.4 nM). This compound assumed as more safe as a drug to be synthesized, and then to be tested in in vitro and in vivo tests. Compound number 1 has a systematic name is N1-(4'-Cyanophenyl)-5-(4''-cyanovinyl-2'',6''-dimethyl-phenoxy)-4-dimethylether benzene-1,2-diamine) and its molecular structure is shown in **Figure 5**.



**Figure 5.** The chosen designed diarylaniline derivative compound (predicted log CC<sub>50</sub> = 1.969)

## CONCLUSION

From the results of this study, it can be concluded that Austin Model 1 method give the best QSAR model using ANN to describes the relationship between descriptors of diarylaniline derivatives to its cytotoxicities. The best-

designed compound is N1-(4'-Cyanophenyl)-5-(4''-cyanovinyl-2'',6''-dimethyl-phenoxy)-4-dimethylether benzene-1,2-diamine) with predicted log CC<sub>50</sub> value of 1.969.

## REFERENCES

- Arief, I., Armunanto, R. and Setiaji, B., 2013, Study on Anti-Hiv Activity of Diarylaniline Derivatives using Quantitative Structure-Activity Relationship (QSAR), *Indonesian Journal of Chemistry* 13(2): 129-135.
- Deeb O. and Drabh, M., 2010, Exploring QSARs of Some Analgesic Compounds by PC-ANN, *Chemical Biology and Drug Design* 76: 255-262.
- Deeb, O., and Jawabreh, M., 2012, Exploring QSARs for Inhibitory Activity of Cyclic Urea and Nonpeptide-Cyclic Cyanoguanidine Derivatives HIV-1 Protease Inhibitors by Artificial Neural Network, *Advances in Chemical Engineering and Science* 2: 82-100.
- Ekins, S. and Williams, A.J., 2012, The Future of Computational Models for Predicting Human Toxicities, *Altex Proceedings*, 1/12, *Proceedings of WC8*.
- Eroğlu, E., Türkmen, H., Güler, S., Palaz, S., and Oltulu, O., 2007, A DFT-



- Based QSARs Study of Acetazolamide/Sulfanilamide Derivatives with Carbonic Anhydrase (CA-II) Isozyme Inhibitory Activity, *International Journal of Molecular Science* 8: 145-155.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., 2009, Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT.
- Gacche, R.N. and Jadhav, S.G., 2012, Antioxidant Activities and Cytotoxicity of Selected Coumarin Derivatives: Preliminary Results of a Structure-Activity Relationship Study Using Computational Tools, *Journal of Experimental and Clinical Medicine* 4(3): 165-169.
- Hemmateenejad, B., Javidnia, K., Nematollahi, M., and Elyasi, M., 2009, QSAR Studies on the Antiviral Compounds of Natural Origin, *Journal of Iranian Chemical Society* 6(2): 420-435.
- Hosseini, S., Monajjemi, M., Rajaeian, E., Haghgu, M., Salari, A. and Gholami, M.R., 2013, A Computational Study of Cytotoxicity of Substituted Amides of Pyrazine-2-carboxylic acids Using QSAR and DFT Based Molecular Surface Electrostatic Potential, *Iranian Journal of Pharmaceutical Research* 12(4): 745-750.
- Hu, R., Doucet, J., Delamar, M., and Zhang, R., 2009, QSAR Models for 2-Amino-6-Arylsulfonylbenzotriazoles and Congeners HIV-1 Reverse Transcriptase Inhibitors Based on Linear and Nonlinear Regression Methods, *European Journal of Medicinal Chemistry* 44: 2158–2171.
- HyperCube, Inc, 2011, HyperChem™ 8.0.10 for Windows, <http://www.hyper.com>.
- IBM Corp. Released 2010. IBM SPSS Statistics for Windows, Version 19.0. Armonk, NY: IBM Corp.
- Jain, S.V., Ghate, M., Bhadoriya, K.S., Bar, S.B., Chaudhari, A., and Borse, J.S., 2012, 2D, 3D-QSAR and Docking Studies of 1,2,3-Thiadiazole Thioacetanilides Analogues as Potent HIV-1 Non-Nucleoside Reverse Transcriptase Inhibitors, *Organic Medicinal Chemistry Letter* 2: 22-34.
- Low, Y., Uehara, T., Minowa, Y., Yamada, H., Ohno, Y., Urushidani, T., Sedykh, A., Muratov, E., Kuz'min, V., Fourches, D., Zhu, H., Rusyn, I. and Tropsha, A., 2011, Predicting Drug-Induced Hepatotoxicity Using QSAR and Toxicogenomics Approaches, *Chemical Research in Toxicology* 24: 1251-1262.

- MATLAB (Version 7.8.0.347 (R2009a), <http://www.mathworks.com>).
- Podunavac-Kuzmanović, S.O., Cvetković, D.D., and Barna, D.J., 2009, QSAR Analysis of 2-Amino or 2-Methyl-1-Substituted Benzimidazoles Against *Pseudomonas aeruginosa*, *International Journal of Molecular Sciences* 10 : 1670-1682.
- Ruiz, P., Begluitti, G., Tincher, T., Wheeler, J. and Mumtaz, M., 2012, Prediction of Acute Mammalian Toxicity Using QSAR Methods: A Case Study of Sulfur Mustard and Its Breakdown Products, *Molecules* 17: 8982-9001.
- Sun, H. and Scott, D.O., 2010, Structure-based Drug Metabolism Predictions for Drug Design, *Chemical Biology and Drug Design* 75: 3-17.
- Sun, L., Zhu, L., Qian, K., Qin, B., Huang, L., Chen, C.H., Lee, K., and Xie, L., 2012, Design, Synthesis, and Preclinical Evaluations of Novel 4-Substituted 1,5-Diarylanilines as Potent HIV-1 Non-Nucleoside Reverse Transcriptase Inhibitor (NNRTI) Drug Candidates, *Journal of Medicinal Chemistry* 55: 7219-7229.
- Wang, Z., Kai, Z., Beier, R.C., Shen, J. and Yang, X., 2012, Investigation of Antigen-Antibody Interactions of Sulfonamides with a Monoclonal Antibody in a Fluorescence Polarization Immunoassay Using 3D-QSAR Models, *International Journal of Molecular Sciences* 13: 6334-6351