Journal of Scientific & Industrial Research
Vol. 76, September 2017, pp. 550-555

QSAR Modeling of Antimicrobial Activity of some *p-substituted* Aromatic Hydrazones

V Dimova[1]* and M Jankulovska[2]

*[1]Faculty of Technology and Metallurgy, Ss Cyril and Methodius University, Ruger Boskovic 16, 1000 Skopje, Republic of Macedonia
[2]Faculty of Agricultural Sciences and Food, Ss Cyril and Methodius University, Blvd Aleksandar Makedonski bb, 1000 Skopje, Republic of Macedonia

Received 03 June 2016; revised 15 December 2016; accepted 04 July 2017

QSAR analysis of a series of previously synthesized *p*-substituted aromatic hydrazones tested for growth inhibitory activity against *Bacillus subtilis*, was performed using several physicochemical descriptors: Surface tension (ST), Molar Refraction (MR), Molar Volume (MV), Parachor (Pc), Index of Refractivity (η); Density (D) and Polarizability (α). Two-parameter models were obtained and validated by using several statistical parameters: R; R^2_{adj}; F-test; Sd; R_{ped}; PRESS/SSY; Q^2; S_{PRESS}; PSE and Q. Both the parameters (D and α) contributing to statistically best model (model 16) have positive input to the modeling of biological activity of selected hydrazones.

Keywords: QSAR, Physicochemical Descriptors, *p*-substituted Hydrazones

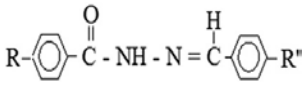
Introduction

Hydrazone derivatives represent one of the most active classes of compounds possessing a broad spectrum of biological activity[1]. The use of the hydrazones in medicine is due to their antiinflammatory, antimicrobial, antidepressant, antitumoral, analgesic, antiplatelet, anticonvulsant, antischistosomiasis and antiviral activity[2]. Hydrazone derivatives possessing an *azometine proton (*-NH-N=CH-) constitute a significant class of compounds for new drug development. The wide palette of useful medical properties has attracted considerable scientific interest for their synthesis[3]. Hydrazones and their metal complexes exhibit a wide spectrum of physiological and pharmacological activities[4]. Due to their physiological activity, they are used as herbicides, insecticides, and plant growth stimulants[3]. Some nitrophenyl hydrazone are effective for the treatment of amoebiasis caused by the protozoan parasite *Entamoeba hystolitica;* against *Mycobacterium tuberculosis* as well as in other arasitic diseases such as leishmaniasis and Chagas disease[5]. Furthermore, the hydrazones are also used in industry as plasticizers, polymer stabilizers, antioxidants, polymerization initiators[6]. Considering these applications some *p*-substituted aromatic hydrazones were synthesized and characterized Table 1, with the hope to get some interesting biological activity. The aim of this work was to derive some statistically significant QSAR models for series of *p*-substituted aromatic hydrazones with selected physicochemical descriptors and experimentally obtained results for *in vitro* inhibitory actvity against *Bacillus subtilis*.

Experimental

Structure of *p*-substituted aromatic hydrazones (H1 - H15)

A series of three *p*-substituted aromatic hydrazones have been synthesized by condensation of benzhydrazide or *p*-substituted benzhydrazides (CH_3, OCH_3, Cl and OH) with benzaldehyde or *p*-substituted benzaldehyde (OCH_3 and NO_2)[7]. Initially, *p*-substituted esters were prepared from benzoic, *p*-substituted benzoic acid and methanol. In the second step, *p*-substituted hydrazides were prepared from the previously synthesized esters and hydrazine hydrate. Finally, *p*-substituted aromatic hydrazones were obtained from hydrazides and benzaldehyde or *p*-substituted benzaldehyde. The identity of the synthesized hydrazones was confirmed by the



Scheme — Structure of *p*-substituted aromatic hydrazones (H1 - H15)

*Author for Correspondence
E-mail: vdimova@tmf.ukim.edu.mk

Table — 1 Structure of *p*-substituted aromatic hydrazones (H1 - H15) used in the present study; calculated log $1/c_{MIC}$ values for the active hydrazone derivatives and physicochemical parameters

Comp.	R	R'	log$1/c_{MIC}$	MR [cm^3]	MV [cm^3]	Pr [cm^3]	η	ST [dyne/cm]	D [g/cm^3]	α 10^{-24} [cm^3]
H1	H	H	5.3500	68.92	207.6	530.4	1.578	42.5	1.07	27.32
H2	CH_3	H	5.3766	73.35	222.9	561.5	1.572	40.2	1.06	29.07
H3	CH_3O	H	5.4048	74.74	229.4	580.7	1.565	41.0	1.10	29.62
H4	Cl	H	5.4116	73.52	216.9	559.3	1.592	44.1	1.19	29.14
H5	OH	H	5.3802	69.77	204.9	536.1	1.596	46.8	1.17	27.66
H6	H	CH_3O	5.4048	74.74	229.4	580.7	1.565	41.0	1.10	29.62
H7	CH_3	CH_3O	5.4280	79.16	244.6	611.8	1.560	39.1	1.09	31.38
H8	CH_3O	CH_3O	5.4533	80.55	251.1	631.0	1.554	39.8	1.13	31.93
H9	Cl	CH_3O	5.4594	79.34	238.7	609.5	1.579	42.5	1.20	31.45
H10	OH	CH_3O	5.4149	75.59	226.6	586.4	1.581	44.7	1.19	29.96
H11	H	NO_2	5.4297	74.58	213.0	575.9	1.617	53.3	1.26	29.56
H12	CH_3	NO_2	5.4517	79.01	228.2	607.0	1.608	50.0	1.24	31.32
H13	CH_3O	NO_2	5.4756	80.40	234.7	626.2	1.600	50.6	1.27	31.87
H14	Cl	NO_2	5.4814	79.18	222.3	604.7	1.630	54.7	1.36	31.39
H15	OH	NO_2	5.4548	75.43	210.3	581.6	1.636	58.4	1.35	29.90

MR - Molar Refraction; MV - Molar Volumen; Pr – Parachor; η - Index of Refracttivity; ST - Surface Tension; D - Density; α - Polarizability

following techniques: ^{1}H NMR, ^{13}C NMR, IR, UV and elemental analysis (CNH)[7]. The proposed method of synthesis resulted in excellent yield and purity of the prepared hydrazones. Structure of investigated *p*-substituted aromatic hydrazones is presented in Table 1.

Antibacterial investigation

All hydrazones were tested for their *in vitro* growth inhibitory activity against *Bacillus subtilis,* using filter paper disc method. Stock solutions of compounds were prepared in DMSO, as inert medium in 3 concentrations: 1, 5 and 10 mg/ml DMSO. A control disc using DMSO without any test compound was included and there was no inhibitory activity in those disks. The diameter of zone inhibition (mm) was measured. Every test was done in triplicate to confirm the findings.

Descriptors

To obtain the quantitative effect of the hydrazone derivatives structural parameters on their biological activities, QSAR analysis with the physicochemical descriptors was operated. The Molar Refraction (MR), Molar Volumen (MV), Parachor (Pr), Index of Refractivity (η); Surface Tension (ST), Density (D) and Polarizability (α), are the properties that can be calculated for each molecule.

These data were used to determine the QSAR models. Physicochemical parameters used in this study were calculated using ACD Labs software[8] Table 1.

Statistical analysis

The statistical evaluation of the data was performed using STATISTICA program package[9]. To test the quality of the regression models, the following statistical parameters were used: Correlation coefficient (R^2); Standard deviation of the estimate (Sd); Fisher test for significance of the equation (F-test); Adjusted R^2 (R^2_{adj}); Quality factor (Q); Predictive residual error (PRESS); Sum of Squares (SSY); Uncertainty of Prediction (S_{PRESS}); Predictive Square Error (PSE) and Cross-validation squared correlation coefficient (Q^2).

Results and Discussion

In the present study, an attempt has been made to find structural requirement for inhibition of *Bacillus subtilis* using QSAR approach on substituted hydrazone derivatives with several descriptors such as: Molar Refraction (MR), Molar Volumen (MV), Parachor (Pr), Index of Refractivity (η), Surface Tension (ST), Density (D) and Polarizability (α). Literature shows that no QSAR study for substituted hydrazones using previously mentioned physicochemical descriptors have been reported. In the first step for development of QSAR models, selected hydrazones were evaluated for *in vitro* antimicrobial activity against *Bacillus subtilis.* In the second step, efforts were focused on developing the QSAR models of compounds with antibacterial activity against

Bacillus subtilis. Inhibitory activity data determined as μg/ml were first transformed to the negative logarithms of molar MICs ($log1/c_{MIC}$), Table 1 which were used as a dependent variable in the QSAR study. According to the calculated values, $log1/c_{MIC}$ is lowest for *p*-benzoic substituted hydrazones H1, H2 and H5. Following sequence of antibacterial activity values were observed:

H1<H2<H5<H3=H6<H4<H10<H7<H11<H12<H8<H13<H14<H15

Calculated values of physicochemical descriptors presented in Table 1, were used as independent variables in correlation with antibacterial activity ($log1/c_{MIC}$).

QSAR study for antibacterial activity of *Bacillus subtilis*

After applying stepwise multiple linear regression method, using physicochemical descriptors Table 1 as independent and $log\ 1/c_{MIC}$ as dependent values, data shows that only MR, Pr and α correlate relatively well with the activity ($log1/c_{MIC}$) ($R^2 > 0.85$). Monoparametric correlations between $log\ 1/c_{MIC}$ and MR; Pr and α are presented by equations: $log\ 1/c_{MIC} = 0.009\ MR + 4.723$ ($R^2 = 0.8877$); $log\ 1/c_{MIC} = 0.01\ Pr + 4.760$ ($R^2 = 0.8556$) and $log\ 1/c_{MIC} = 0.023\ \alpha + 4.723$ ($R^2 = 0.8883$). From this it is concluded that single variable model is capable of modelling the activity and that only the refereed descriptors can be combined to obtain a statistically significant mono-parametric model for modelling the activity. Several bi-parametric models were obtained in the next step of investigation. Among them, best 16 were selected for further discussion. The selection was based on the preliminary statistical parameters (R; Sd; R^2_{adj}). The statistically significant results for bacteria inhibitory activity against *Bacillus subtilis* of hydrazones derivatives, using two descriptors, have been summarized in Table 2. It was important for further analysis to find correlation matrix for used descriptors and their correlation with the activity. The results show that some of them are high correlated. This high collinearity indicated that these parameters could not be combined to get multiple linear regression (MLR) models. If a combination of them is present in the regression expression, then the model may suffer from the defect due to collinearly. Also, it may result in change in signs of the coefficients, a change in the values of the previous coefficient, the change of a significant variable into an insignificant one or an increase in standard error of the estimate on addition of an additional parameter to the model. Having in mind this fact, model 1, 5, 6 and 10 were excluded from further analysis. The correlation among descriptors used in models 2-4, 7-9 and 11-16 are poorly correlated. Taking into consideration the above mentioned and preliminary conclusions of the statistical evaluation of the quality of the selected models, previously mentioned models can be used as relatively statistically significant Selected models (2-4, 7-9, 11-16) show that all physicochemical descriptors are positively correlated with activity, which indicates that the activity goes on increasing with the increasing value of descriptors. The correlation coefficients in all the cases were found to be good (0.938-0.988); R^2_{adj} above 0.92 and the standard deviation below 0.1 Table 2. An excellent correlation is obtained in Models: 4, 9, 13 and 16, where correlation coefficients are above 0.98. In order to confirm this findings, antimicrobial activity against *B. subtilis* ($log\ 1/c_{model}$) predicted by models 2-4, 7-9 and 11-16 Table 2 were compared with the corresponding observed values ($log\ 1/c_{exp} = log1/c_{MIC}$) reported in Table 1. A plot between the observed and predicted $log\ 1/c_{MIC}$ values is presented in Figure 1. Within experimental error, the values agree well. Parallelly, predictive correlation coefficients (Rpre) have been calculated Table 2. We have estimated the predictive correlation coefficients (R_{pre}) to examine the relative potential of the proposed models. The values R_{pre} are found > 0.93, for all models.

Statistical evaluation

Performing the multiple linear regression of a dependent variable (y; $log\ 1/c_{MIC}$) offers the possibility of choosing a large number of explanatory variables (x) and thus raises the question of significance in an acute form. Statistical quantities need to be calculated in order to assess the success of the correlation. Validation is a crucial aspect of any QSAR analysis and a cross-validation methodology was done for deciding predictive power of the proposed models. This is needed because a model with good statistics may not have good predictive potential. The predictivity of each model was measured by several cross-validation parameters: *Q*, PRESS/SSY; S_{PRESS}; *PSE* and Q^2 (Tab. 2). On the basis of those parameters, the selected two-parametric models can be ranked (from the best to the worst) in the following order:

- Quality factor (Q):
  9>16=4>13>7>15>3>11>2=14>12>8;

Table — 2 Regression parameters and the quality of correlation of log $1/c_{MIC}$ with chosen descriptors in multivariate regressions for substituted hydrazones derivatives (*B. subtilis*) for models 1-16 and Cross-validation parameters (Q, PRESS/SSY; S_{PRESS}; PSE and Q^2) for models 2-4, 7-9, 11-16

A. Two-parametric models: log $1/c_{MIC} = a + bX_1 + cX_2h$

Model	log $1/c_{MIC}$ =	R^2	Sd	R^2_{adj}	F-test
1	4.7229±0.047 + 0.0151±0.001MR - 0.0019±2.8x10^{-4}MV	0.9568	0.0086	0.9496	132.832
2	3.7332±0.144 + 0.0091±6.0x10^{-4}MR + 0.6266±0.087η	0.9602	0.0082	0.9536	144.8671
3	4.6380±0.046 + 0.0087±5.9x10^{-4}MR + 0.0026±3.5x10^{-4}ST	0.9620	0.0082	0.9555	151.362
4	4.6384±0.028 + 0.0073±3.8x10^{-4}MR + 0.1905±0.015D	0.9857	0.0049	0.9833	412.655
5	4.7550±0.101 - 0.7879±0.582MR + 2.0100±1.468α	0.8159	0.0177	0.7853	26.602
6	4.8020±0.052 - 0.0024±3.6x10-4MV + 0.0020±1.6x10-4Pr	0.9409	0.0100	0.9309	95.432
7	2.2238±0.191 + 0.0030±1.9x10^{-4}MV + 1.5811±0.102η	0.9630	0.0079	0.9569	156.337
8	4.5473±0.067 + 0.0026±2.3x10-4MV + 0.0060±15.1x10-4ST	0.9385	0.0102	0.9283	91.599
9	4.5700±0.028 + 0.0019±9.10x10^{-5}MV + 0.3560±0.013D	0.9882	0.0045	0.9862	501.249
10	4.7235±0.047 - 0.0019±2.8x10^{-4}MV + 0.0381±0.003α	0.9568	0.0086	0.9496	132.998
11	3.6252±0.147 + 0.0011±7.3x10^{-5}PR + 0.7209±0.087η	0.9603	0.0082	0.9537	154.100
12	4.6799±0.055 + 0.0010±8.8x10^{-5}PR + 0.0028±4.3x10^{-4}ST	0.9413	0.0100	0.9315	96.260
13	4.6653±0.031 + 8.6000x10^{-4}±5.2x10^{-5}PR + 0.2090±0.016D	0.9816	0.0056	0.9786	320.294
14	3.7350±0.144 + 0.6254±0.087η + 0.0231±0.001α	0.9602	0.0082	0.9536	144.931
15	4.6385±0.046 + 0.0026±3.5x10^{-4}ST + 0.0221±0.001α	0.9619	0.0080	0.9556	151.654
16	4.6388±0.028 + 0.1902±0.015D + 0.0186±9.7x10^{-4}α	0.9857	0.0049	0.9833	412.298

B. Cross-validation parameters for models 2-4, 7-9, 11-16

Model	log $1/c_{model} = a$ log $1/c_{exp} + b$			Q	PRESS/SSY	Spress	PSE	Q^2
	a	*b*	R_{pre}^2					
2	0.954	0.239	0.960	119.4998	0.0644	0.0104	0.0093	0.9356
3	0.979	0.243	0.961	119.6117	0.0800	0.0115	0.0103	0.9200
4	0.979	0.105	0.985	202.6172	0.0499	0.0091	0.0082	0.9501
7	0.955	0.231	0.962	124.2184	0.1623	0.0166	0.0148	0.8377
8	0.925	0.388	0.937	94.9767	0.2586	0.0210	0.0188	0.7414
9	0.985	0.072	0.987	220.9072	0.0297	0.0071	0.0063	0.9703
11	0.949	0.263	0.959	119.5060	0.1198	0.0141	0.0126	0.8802
12	0.904	0.488	0.940	97.0206	0.5777	0.0314	0.0280	0.4223
13	0.982	0.093	0.981	176.9209	0.0733	0.0112	0.0100	0.9267
14	0.960	0.215	0.960	119.4998	0.0417	0.0087	0.0077	0.9583
15	0.959	0.217	0.961	122.5956	0.0426	0.0087	0.0077	0.9574
16	0.985	0.077	0.985	202.6172	0.0158	0.0050	0.0045	0.9842

X_i - chosen descriptors; R^2 - Correlation coefficient; Sd - Standard deviation of the estimate; F-test - Fisher test for significance of the equation; R^2_{adj} - Adjusted; Q - Quality factor; PRESS - Predictive residual error; SSY - Sum of Squares; S_{PRESS} - Uncertainty of Prediction; PSE - Predictive Square Error; Q^2 - Cross-validation squared correlation coefficient

- Predictive residual error/Sum of Squares (PRESS/SSY): 16>9>14>15>4>2>13>3>11>7>8>12;
- Uncertainty of Prediction (S_{PRESS}): 16>9>14=15>4>2>13>3>11>7>8>12;
- Predictive Square Error (PSE): 16>9>14=15>4>2>13>3>11>7>8>12;
- Cross-validation squared correlation coefficient (Q^2): 16>9>14>15>4>2>13>3>11>7>8>12.

Since parameters in the quality factor definition $(Q=R/Sd)^{10}$ is not in any way connected to the prediction power of the model, use of Q factor ranking is excluded,

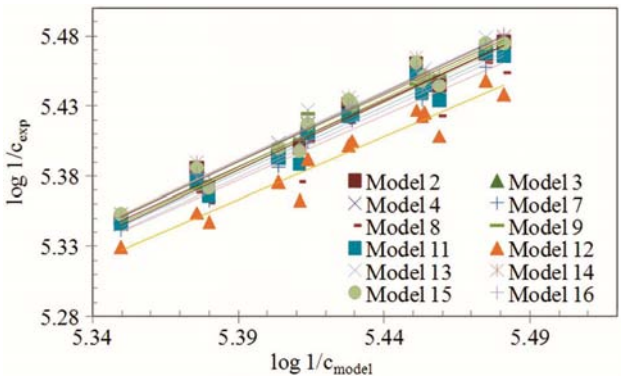


Fig. 1 — Correlation between the observed log $1/c_{MIC}$ (log $1/c_{exp}$) and predicted log $1/c_{MIC}$ (log $1/c_{model}$) values

in the next step of statistical evaluation ranking according the statistical parameters related to the prediction power (goodness of prediction) of the models, such as: PRESS/SSY; S_{PRESS}; *PSE* and Q^2. Models with PRESS/SSY < 0.4[11], as in this case (RESS/SSY ranges between 0.015 - 0.577), are reliable QSPR models. Good S_{PRESS} and *PSE* values, were obtained for all selected models (S_{PRESS} and *PSE* < 0.02 and 0.03, respectively) confirming the assumption that those models can be used as a tool for predicting the inhibition of *B. subtilis*[12]. The lowest value for PRESS/SSY, S_{PRESS} and *PSE*, for model 16 (PRESS/SSY = 0.015; S_{PRESS} = 0.005; *PSE* = 0.0045), compared to other models, supports its highest predictive potential. For a reliable model, the Cross-validation squared correlation coefficient values (Q^2) should be > 0.6[11], and the difference between R^2 and Q^2 should not be more than 0.3[12,13]. Hence, once again, the most appropriate model for modeling the activity is the model 16 ($Q^2_{Model\ 16}$ = 0.9842). The both parameters (D and α) contributing to model 16 have positive input to the modeling of biological activity of selected hydrazones. The higher influence associated with D (91.17 %), supported with lower influence related with α (8.83 %), indicate their positive role towards antibacterial activity. In this model, the coefficient of D and α are higher than their standard deviation, which is another confirmation for statistical significance of model 16. It can be also noticed that the t-value on the independent variables in model 16 is -12.863 for D and 19.034 for α, with corresponding p - value < 0.05 (D = $2.22x10^{-8}$; α = $2.48\ x10^{-10}$) indicating that they are statistically significant at the 95% or higher confidence level (Critical $F_{(0.05,2,12)}$=3.89; Critical $t_{(0.05,12)}$=2.178).

Conclusion

Spurred by the need of new antimicrobial agents and the fact that hydrazones derivatives are the active classes of compounds possessing a wide spectrum of biological activity, series of *p*-substituted aromatic hydrazones have been synthesized and their inhibitory activity against *Bacillus subtilis* was investigated. QSAR study was performed to estimate the quantitative effects of the physicochemical descriptors of selected hydrazones derivatives on their antibacterial activity. Physicochemical properties such Surface tension (ST), Molar Refraction (MR), Molar Volume (MV), Parachor (Pc), Index of Refractivity (η); Density (D) and Polarizability (α) were calculated for each molecule and a several two – parametric mathematical models relating the antimicrobial activity, $\log 1/c_{MIC}$, were defined. The statistical significance of each model was measured by a few cross-validation parameters *(Q, PRESS/SSY; S_{PRESS}; PSE* and *Q^2)*. Statistical evaluation of the data used to test the quality of the obtained models, indicated that model 16 is statistically significant when all parameters are summarized, as both parameters (D and α) contributing to statistically best model (model 16) have positive input to the modeling of biological activity of selected hydrazones.

References

1 Singh N, Ranjana R, Kumari M & Kumar B, A Review on Biological Activities of Hydrazone Derivatives, *Int J Pharm Clin Res*, **8** (2016) 162-166.

2 Venkatasatyanarayana G, Rao V Lakshmana, Laxminarayana E & Chary M Thirumala, Synthesis, characterization and antimicrobial evaluation of (E)-4-(substitutedphenyl)-N'-(1-phenylethylidene) cyclohexane carbohydrazide derivatives, *Ind J Chem, B*, **55B** (2016) 718-723.

3 Rajarajan M, Senbagam R, Vijayakumar R, Balaji S, Manikandan V, Vanangamudi G & Thirunarayanan G, Synthesis, spectral correlations and antimicrobial activities of substituted 4-((E)-2-benzylidenehydrazinyl) benzonitriles*, Ind J Chem, B*, **55B** (2016) 197-206.

4 Shakdofa M M E, Shtaiwi M H, Morsy N & Abdel-rassel T M A, Metal complexes of hydrazones and their biological, analytical and catalytic applications: A review, *Main Group Chemistry*, **13** (2014) 187-218.

5 Toledano-Magaña T, Carlos García-Ramos J, Navarro-Olivarria M, Flores-Alamo M, Manzanera-Estrada M, Ortiz-Frade L, Galindo-Murillo R, Ruiz-Azuara L, Ma Meléndrez-Luevano R & Cabrera-Vivas B, Potential amoebicidal activity of hydrazone derivatives: synthesis, characterization, electrochemical behavior, theoretical study and evaluation of the biological activity yanis, *Molecules*, **20** (2015) 9929-9948.

6 Shakir M & Abbasi A, Synthesis and spectral characterization of hydrazone Schiff base ligand, L derived from condensation of terephthalaldehyde and 2-furoic acid hydrazide and its binuclear complexes with Co(II), Ni(II), Cu(II) and Zn(II): Comparative in-vitro microbiological evaluation of L and its Zn(II) coordination complex*, J Chem Pharm Res*, **7** (2015) 375-382.

7 Jankulovska M, Colanceska-Ragenovic K, Dimova V, Spirevska I & Makreski P, Synthesis and characterization of new p-substituted aromatic hydrazones, *Org Chem*, *An Ind J*, **8** (2012) 326-334.

8 ACD-Lab software for calculationg the referred physicochemical properties, Chemsketh SO, [http://www.acdlabs.com].

9 STATISTICA program package [http://www.statsoft.com].

10 Sharma N, Dwivedi A, Srivastava A K & Singh A, QSAR modeling of HIV-1 reverse transcriptase inhibitor of aryluracil derivatives using ab initio and empirical calculations*, Ind J Chem, B*, **55B** (2016) 752-760.

11 Gharbi R, Almi Z, Belaidi S & Boughdiri S, Structure-Activity/Property Relationships and QSAR Modeling of

Antiamoebic activity of Citral Derived amides, *Res J Pharm, Biol Chem Sci*, **7** (2016) 651 – 657.

12 Shubha J, Kumar A A & Satish P, QSPR Models for Predicting Lipophilicity of Triazole Derivatives, *Re J Chem Sci*, **4** (2014) 1-6.

13 Yong-Hong Z, Zhi-Ning X, Li Y & Shu-Shen L, Prediction of Placental Barrier Permeability: A Model Based on Partial Least Squares Variable Selection Procedure, *Molecules*, **20** (2015) 8270-8286.