Iranian Journal of Analytical Chemistry
Volume 8, Issue 2, September 2021 (25-31) Origonal Research Article

مجله ایرانی شیمی تجزیه
دوره۸، شماره ۲، مهرماه ۱۴۰۰ (۳۱-۲۵)

A Quantitative Structure-Activity Relationship Study of 2, 4, 6-s-Triazine Derivatives as Antimalarial Agents

Nosrat Madadi Mahani*, Maryam Alibeigi
Department of Chemistry, Payame Noor University, 19395-4697 Tehran, Iran
Received: 24 August 2021 **Accepted:** 20 September 2021
DOI: 10.30473/ijac.2021.61986.1211

Abstract
A quantitative Structure-Activity Relationship (QSAR) model was applied to the prediction of the antimicrobial activity of 22 derivatives 2, 4, 6-s-triazine as anti-malarial agents. The antimicrobial activity of 22 2, 4, 6-s-triazine derivatives were modeled with the descriptors of quantum-chemical calculations with density functional theory (DFT) method at B3LYP/6–31G level and topological descriptors. This study was conducted using the multiple linear regressions (MLR), the partial least square analysis (PLS) and the principal component regression (PCR) method. Results displayed that the MLR method predicted of antimicrobial activity good enough. The best model, with six descriptors was selected. Also it indicates very good consistency towards data variations for the validation methods. The predicted values of antimicrobial activity are in suitable agreement with the experimental results. The obtained results suggested that the PLS method could be more helpful to predict the antimicrobial activity of 2, 4, 6-s-triazine derivatives. This study to be usable to predict the activity of other derivatives in the same groups.

Keywords
Multiple Linear Regression; Partial Least Square; Density Functional Theory; Principal Component Regression.

1. INTRODUCTION

Malaria is a protozoan disease caused *Plasmodium* genus. Antifolate antimalarial drugs such as cycloguanil and pyrimethamine have been used in prevention and treatment of malaria [1]. 2-amino-4-(*p*-chloro-anilino)-6 6-dimethyl-5, 6-dihydro-1, 3, 5-triazine and *p*-chlorophenylbiguanide compounds were found to be inactive against *Plasmodium falciparum.* A large number of dihydrotriazines have been synthesized and the most of them demonstrated antimalarial activity [2]. A series of 1, 3, 5-triazine-substituted polyamines have been synthesized and have been investigated antiplasmodial activity biologically. Several triazine compounds have been synthesized and biologically evaluated for biochemical targets such as polyamine metabolism [3] and dihydrofolate reductase (DHFR) inhibition [4].

Besides it, *s*-triazine compounds were distinguished to be active against viruses, active as antitumorigenic agents, in chemotherapeutical preparations, protozoa, helminths, pharmacologically effective to therapy cardiovascular, neuropsychotic disorders, or inflammatory processes, diuretics, antidiabetic agents, etc. [5]. *S*-triazine derrivatives are the second class that inhibits Plasmodium falciparum-DHFR [6, 7]. DHFR has obtained significant consideration as it is the target of cycloguanil. That it is a triazine based antimalarial drug and other antifolates. Also, DHFR is used for the treatment and prophylaxis of Plasmodium falciparum infection [8].

Some of silico techniques, like molecular docking, pharmacophore models or quantitative structure–activity relationships (QSAR) significantly decrease the time and cost in the drug discovery process. Among the techniques, QSAR is considered a valuable tool that is applied extensively in rational drug design and plays a useful role for screening of new potential lead compounds.

The predictive QSAR model can be create a mathematical correlation between the structural properties of the compounds and their anti-malarial activities using different methodologies, such as multiple linear regression (MLR), partial least squares (PLS), artificial neural networks (ANN) and heuristic method (HM). Prediction of Quantitative structure-activity relationship on the anti-malarial activity a set of new imidazolopiperazines based on artificial neural networks has been studied [9].

The studies of 2D and 3D-QSAR molecular docking and ADMET properties of azaaurones as antimalarial agents have been performed by Hadni and coworker [10]. The predictive quantitative structure activity relationship (QSAR) models for anti-malarial activity of 4-aminoquinolines has been developed by Masand et al [11]. Sanchait et al. have been investigated QSAR of the antimalarial agent artemisinin and some of its

*Corresponding Author :nmmadady@gmail.com

derivatives with a DFT approach has been studied by [12]. Also, Nguyen et al. have been studied 2D-quantitative structure–activity relationships model using PLS method for anti-malarial activities of anti-haemozoin compounds [13]. The aim of this study was to develop quantitative structure–activity relationship models to determine the influences of physiochemical structures of 2, 4, 6-s-triazine derivatives as anti-malarial agents.

2. EXPERIMENTAL

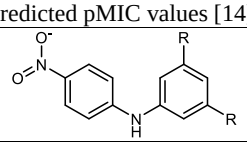
QSAR modelling was performed for 2, 4, 6-s-triazine derivatives as anti-malarial agents using the multiple linear regression (MLR), principal component regression (PCR) and, partial least square (PLS) methods. Database of 22 compounds possessing anti-malarial activity were used for building QSAR models.

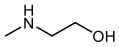
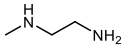
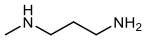
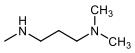
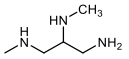
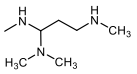
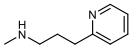
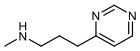
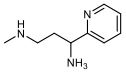
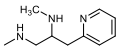
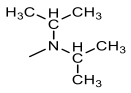
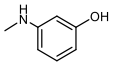
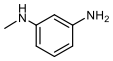
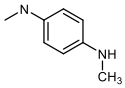
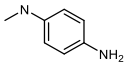
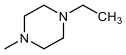
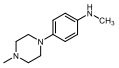
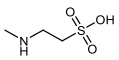
The antimicrobial activity of 4, 6-s-triazine derivatives was used. The chemical structures and minimum inhibitory concentration (MIC) of these compounds based on reference [14] are listed in Table 1. A complete geometry optimization was carried out with GAUSSIAN 09 program [15] taking the most general conformations as outset geometries. Density functional theory (DFT) calculations [16] of the structures were performed in B3LYP/6-31G level [17].

The molecular descriptors were obtained using HyperChem, Dragon and, GAUSSIAN Packages. Some of quantum chemical descriptors including dipole moment (DM), lowest unoccupied molecular orbital (LUMO) and highest occupied molecular orbital energy (ET), zero point energy (ZPE), hardness, electronic chemical potential, global electrophilicity index and mass. Some chemical parameters including molecular volume, molecular surface area (Sur), hydrophobicity (logP), polarizability, refractivity (Ref) and hydration energy (HE) were calculated using Hyperchem software.

The optimized geometries were loaded into Dragon software [18] to calculate 1481 descriptors in 18 different classes. The program contains scripts for generating descriptors of different types including: constitutional, topological, radialdistribution functions (RDF), geometry, topology and atoms weighted assembly (GETAWAY), functional groups, weighted holistic invariants (WHIM), Randic molecular profiles, 3D-molecular representation of structure based on electron diffraction (3D-Morse) etc [19]. For each compound in the training sets, the correlation equation was obtained with the same descriptors. Then, the obtained equation was used to predict pMIC values for the compounds from the corresponding test sets. Two programs including SPSS and Minitab were used for MLR, PCR and PLS.

Table 1. Structural features of 2, 4, 6-s-triazine and predicted pMIC values [14]**.**



No.	R group	pMIC
1		-4.47
2		-3.84
3		-3.19
4		-3.41
5		-2.61
6		-2.78
7		0.65
8		-0.19
9		-o.13
10		0.58
11		-2.76
12		-3.46
13		-0.90
14		-2.93
15		-1.85
16		-2.29
17		-3.23
18		0.44
19		-2.2
20		-3.88
21		-4.05
22		-6.8

3. RESULT AND DISCUSSION

A QSAR study was performed for a series of 2, 4, 6-s-triazine derivatives as anti-malarial agents, for

characterizing a quantitative relationship between structure chemical and antimicrobial activity. The multiple linear regression statistic method is used to study the relation between one dependent variable and several independent variables. Also, minimizes differences between experimental and predicted values. The MLR was acquired using the software SPSS to predict activities pMIC. The Linear Regression method related to a larger family of models called generalized linear models. The choice of the training set is one of the most significant stages in the QSAR modeling, since the confirmation and optimization of a QSAR model are based on this training set. Applicability and predictability of a QSAR model also relay on the training set selection. The data set (n = 22) was divided casually into two groups: train set (n = 16) and test set (n = 6). The Pearson correlation coefficients are listed in the following table 2. The correlation coefficient (R^2) matrix for the descriptors used in different MLR equations shows that no significant correlation exists between pairs of descriptors. The acquired matrix gives information on the positive or negative correlation between variables.

We used the pMIC of the 2, 4, 6-s-triazine derivatives as the dependent variable. This model with acceptable statistical quality (R^2=0.825) indicated that the antimicrobial activity of compounds is influenced by topological parameters IDDE, RDF095v, G2v and LUMO, DM and polar values as representative of quantum chemical descriptors. The positive relation of activity and "LUMO, RDF095v and, IDDE" displays that increasing of these descriptors cases increasing antimicrobial activity of compounds. The obtained descriptors demonstrating the electronic characteristic of the studied molecules listed on table 3.

The pMIC predicted of 2, 4, 6-s-triazine derivatives by this model is partly like that observed. Values of pMIC predicted in tables 4 and 5 are listed. Also, Fig. 1 displays a very orderly distribution of pMIC values based on the observed values. The leave–one–out (LOO) approach was applied to carry out the cross-validated analysis. Q^2 (cross–validated coefficient) is computed using the following equation [20, 21]:

$$Q^2 = 1 - \frac{\Sigma(y_i - y_{ipred})^2}{\Sigma(y_i - y_{imean})^2} \quad \text{(Eq. 1)}$$

Where y_i is the i_{th} experimental pMIC value, y_{ipred} is the i_{th} predicted pMIC and y_{mean} is the mean of the experimental pMIC.

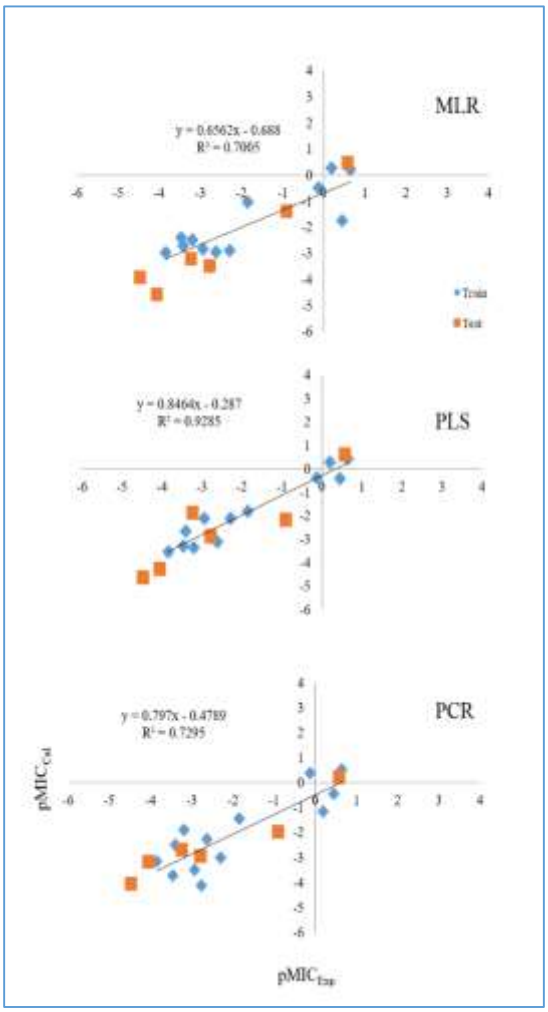


Fig. 1. Correlation of calculated vs. experimental pMIC.

Table 2. Correlation coefficient (R^2) matrix for descriptors represented in multiple linear regression.

	IDDE	RDF095v	G2v	LUMO	Dipole moment	Polarizability
IDDE	1					
RDF095v	0.326	1				
G2v	-0.532	-0.449	1			
LUMO	0.053	0.157	-0.207	1		
Dipole moment	0.433	0.306	-0.254	0.306	1	
Polarizability	0.457	0.536	-0.560	0.084	0.454	1

Table 3. List of molecular descriptors involved in MLR model.

Descriptor definition	Descriptor type	Symbol	Regression coefficient	Standard deviation
mean information content on the distance degree equality	Information indices	IDDE	4.882	1.842
Radial Distribution Function - 095 / weighted by van der Waals volume	RDF	RDF095v	0.319	0.108
2nd component symmetry directional WHIM index / weighted by van der Waals volume	WHIM	G2v	-272.008	105.597
The lowest unoccupied molecular	Quantum mechanic	LUMO	0.075	0.042
Dipole moment	Quantum mechanic	DM	-0.550	0.343
Polarizability	Quantum mechanic	Polar	-0.300	0.117
Constant	-	-	40.587	18.581

The accuracy of the model was mostly estimated by Root Mean Square Error (RMSE) that calculated using the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{y_i - y_{ipred}}{n}} \qquad \text{(Eq. 2)}$$

Where n = number of compounds, y_i = experimental value, y_{ipred} = predicted value [22, 23]. The RMSE and Q^2 of the calibration using MLR method were obtained as 0.399 and 0.781, respectively. Q^2 is used as a criterion of both validity and predictive ability of the model.

PLS technique is a generalization of regression, which can apply data with forcefully correlated and ⁄ or numerous independent variables. The linear PLS model detects new variables that are linear combination of the principal variables. To eschewing over fitting, a formidable test for the significance of each successive PLS component is essential and then pausing when the components are non-significant. The PLS have two purposes: to estimate the matrix *X* of molecular structure descriptors to the matrix Y of dependent variables and to maximize the relationship between them [24]. We presented the data matrix organized clearly from the descriptors offered by MLR corresponding to the molecules, to the PLS. This method used the coefficients *R*, R^2, and the SE values to choose the best regression performance.

The obtained parameters explaining the electronic aspect of the investigated molecules listed on Table 6. The resulted predictions of the pIC_{50} using PLS method in gas phase was given in table 6. $pMIC_{predicted}$ of 2,4,6-s-triazine derivatives by PLS method is little similar to that observed. Figure 1shows a normal distribution of pMIC values based on the observed values. The resulted parameters describing the electronic aspect of the studied molecules are: R^2=0.953and RMSE= 0.487. Cross-validation is a practical and validity method for testing the significance. PLS is usually used in merging with cross-validation to gain the optimum number of components.

Table 4. The predicted activity (by MLR, PLS and PCR) for pMIC of 2, 4, 6-s-triazine derivatives (training set)**.**

Name	pMIC	MLR	PLS	PCR
2	-3.84	-3.14	-3.53	-2.97
3	-3.19	-1.89	-3.38	-2.47
4	-3.41	-2.5	-2.68	-2.71
5	-2.62	-2.28	-3.12	-2.94
7	0.65	0.54	0.44	0.20
8	0.19	-1.17	0.29	0.29
9	0.13	0.38	-0.40	-0.49
11	-2.76	-4.13	-2.93	-3.48
12	-3.46	-3.72	-3.29	-2.41
14	-2.93	-3.49	-2.12	-1.76
15	-1.85	-1.45	-1.80	-1.19
16	-2.29	-3.02	-2.12	-3.23
18	0.44	-0.44	-0.42	-1.76
19	-2.20	-1.66	-1.95	-1.19
20	-3.88	-3.24	-3.64	-3.23
22	-6.80	-6.34	-7.06	-6.94

Table 5. The predicted activity (by MLR, PLS and PCR) for pMIC of 2, 4, 6-s-triazine derivatives (test set)

Name	pMIC	MLR	PLS	PCR
1	-4.47	-4.05	-4.46	-3.92
6	-2.78	-2.93	-2.90	-3.50
10	0.58	0.20	0.61	0.47
13	0.90	-1.98	-2.19	-1.38
17	-3.23	-2.67	-1.88	-3.20
21	-4.05	-3.19	-4.25	-4.56

Table 6. The statistical parameters of different created QSAR models.

		R^2	SE	F	RMSE	Q^2
MLR	Train	0.825	5.656	7.086	0.399	0.781
	Test	0.907	3.674	38.870	0.705	
PLS	Train	0.953	0.436	282.28	0.487	0.932
	Test	0.818	0.928	17.97	0.674	
PCR	Train	0.808	1.098	6.294	0.470	0.678
	Test	0.944	0.514	67.407	0.797	

The PCR is a helpful statistical technique for summing up all the information coded in the structures of compounds and very useful for identifying the link between the different variables [25]. The molecules of 2, 4, 6-s-triazine derivatives were studied by statistical method based on the PCR. The obtained parameters from PCR analysis of the studied molecules are listed on Table 5. The resulted predictions of the pMIC using PCR method in gas phase were given in Table 6. Values of pMIC predicted of 2, 4, 6-s-triazine derivatives by PCR method is almost similar to that observed. Figure1(c) shows a very adequate distribution of pMIC values based on the observed values. The obtained parameters defining the electronic aspect of the studied molecules are: R^2=0.808and RMSE= 0.47. It corroborates that the PCR results were the best to creating the quantitative structure activity relationship models.

4. CONCLUSION

In this work, we have studied the QSAR models to predict the activity of 2, 4, 6-s-triazine derivatives. The study of the MLR, PCR and PLR models show that the PLS method has substantially better predictive capability than the other methods. With considering the error, the prediction of the pMIC values was quite satisfactory and the performance of the QSAR model to predict pMIC value was also calculated using the internal cross-validation method. The sanity of the three created models used in this study has good consistency and great predictive power. By defining the molecular descriptors in the regression model, we finalize that the decreased LUMO, dipole moment and, polarization as well as the increased magnitude of IDDE, RDF095v, and G2v are reliable for the larger activity of the investigated compounds. Eventually, the accuracy and predictability of the suggested models were demonstrated by evaluating essential statistical indexes such, as Q^2, R2 and RMSE of different models using different statistical models and descriptors, as shown in Table 6.

ACKNOWLEDGEMENTS

We gratefully thank Payam Noor University for financial support. The authors would like to express their appreciations to department of chemistry this university for providing research facilities.

REFERENCES

[1] B. Blasco, D. Leroy, D.A. Fidock, Antimalarial drug resistance: linking Plasmodium falciparum parasite biology to the clinic, *Nat. Med.* 23 (8) (2017) 917–928.

[2] E.J. Modest, Chemical and biological studies on 1, 2-dihydro-s-triazines. II. Three-component synthesis, *J. Org. Chem.* 21 (1956) 1–13.

[3] B.Klenke, M.P. Barrett,R. Brun, I.H. Gilbert, Antiplasmodial activity of a Series of 1, 3, 5-triazine-substituted polyamines, *J. Antimicrob Chemother.* 52 (2003) 290–293.

[4] A. Garwal, K. Srivastava, S.K. Puri, P.M.S. Chauhan, Synthesis of 2, 4, 6-trisubstituted triazine as antimalarial agents, *Bioorg. Med. Chem. Lett.* 15 (2005) 531–533.

[5] L.N. Yakhontov, G.M. Vakhatova, Search for medicinal preparations in the series of 1, 3, 5-triazines, *Pharm. Chem. J.* 15 (1981) 546–561.

[6] P.M.S. Chauhan, Currents and potential flaricides, *Drugs Future.* 25 (2000) 481–488.

[7] P.M.S. Chauhan, S.K. Srivastava, Present trends and future strategy in chemotherapy of malaria, *Curr. Med. Chem.* 8(2001) 1535–1542.

[8] R.L. Blakley, S.J. Benkovic, Dihydrofolate reductase in folates and pteridines. Chemistry and biochemistry of folates: dihydrofolate reductase, vol 1. Wiley, New York (1984) 191–251.

[9] S.Yousefnejad, M. Mahboubifar, R. Eskandari, Quantitative structure-activity relationship to predict the anti-malarial activity in a set of new imidazolopiperazines based on artifcial neural networks, Malaria J. 18 (310) (2019) 1-17.

[10] H. Hadni, M. Elhallaoui, 2D and 3D-QSAR, molecular docking and ADMET properties in silico studies of azaaurones as antimalarial agents, *New J. Chem.* 44 (2020) 6553–6565.
[11] V.H. Masand , A.A. Toropov, A.P. Toropova, D.T. Mahajan, QSAR models for anti-malarial activity of 4-aminoquinolines, *Curr. Comput. Aided Drug Des*. 10 (1) (2014) 75-82.
[12] R. Sanchait, H. Iftikar, K.H. Kalyan, S. Pubalee, C.D. Ramesh, Quantitative Structure-Activity Relationships of the Antimalarial Agent Artemisinin and Some of its Derivatives – A DFT Approach, *Comb. Chem. High T. Scr*. 16 (8) (2013) 590-602.
[13] P.T.V. Nguyen, T. V. Dat, S. Mizukami, D. L. H. Nguyen, F. Mosaddeque, S. N. Kim, D. H. B. Nguyen, O. T. Đinh, T. L.Vo, G. L.T. Nguyen, C. Q. Duong, S. Mizuta, D. N. H. Tam, M. P. Truong, N. T. Huy, K. Hirayama, 2D-quantitative structure–activity relationships model using PLS method for anti-malarial activities of anti-haemozoin compounds, *Malaria J.* 20 (64) (2021)1-12.
[14] M. Pathak, H. Ojha, A.K. Tiwari, D. Sharma, M.Saini, R. Kakkar, Design, synthesis and biological evaluation of antimalarial activity of new derivatives of 2,4,6-s-triazine, *Chem. Cen. J.* 11 (2017) 132-143.
[15] M.J. Frisch, et al. Gaussian 09, Revision B,01, Gaussian, Inc, Pittsburgh, PA. 2009.
[16] R.G.Parr, R.G.Pearson, Absolute hardness: companion parameter to absolute electronegativity, *J. Am. Chem. Soc.* 105(1983) 7512-7516.
[17] C. Lee, W.Yang, R.G. Parr, Development of the Colle-Salvetti correlation energy formula into a functional of the electron density, *Phys. Rev. B* 37 (1988) 785-789.
[18] A. Mauri, V. Consonni, M. Pavan, R. Todeschini, Dragon software: an easy approach to molecular descriptor calculations, MATCH *Commun. Math. Comput. Chem.* 56 (2006) 237–248.
[19] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim (2000) pp. 667-690.
[20] V. Consonni, D. Ballabioand, R. Todeschini, Comments on the definition of the Q2 parameter for QSAR validation, *J. Chem. Inf. Model* 49 (2009) 1669-1678.
[21] V. Consonni, D. Ballabioand,R. Todeschini, Evaluation of model predictive ability by external validation techniques, *J. Chemometrics* 24 (2010)194-201.
[22] P.Y. Lee, C.Y. Chen, Toxicity and quantitative structure–activity relationships of benzoic acids to Pseudokirchneriella subcapitata, J. *Hazard. Mater*. 165 (2009) 156-161.
[23] G. Jing, Z. Zhou, J. Zhuo, Quantitative structure-activity relationship (QSAR) study of toxicity of quaternary ammonium compounds on Chlorella pyrenoidosa and Scenedesmus quadricauda, *Chemosphere* 86 (2012) 76-82.
[24] S. Wold S, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics. *Chemomet Intell Lab.* 58 (2001) 109–130.
[25] L. Saghaie, H. Sakhi, H. Sabzyan, M.Shahlaei, D. Shamshirian, Stepwise MLR and PCR QSAR study of the pharmaceutical activities of antimalarial 3-hydroxypyridinone agents using B3LYP/6-311++G** descriptors. *Med. Chem. Res*. 22 (2013) 1679–1688.

مطالعه رابطه کمی فعالیت – ساختار مشتقات۲،٤،٦-S-تری آزین به عنوان عامل‌های ضد مالاریا

نصرت مددی ماهانی*، مریم علی بیگی

گروه شیمی، دانشکده علوم، دانشگاه پیام نور، تهران، ایران

تاریخ دریافت: ۲ شهریور ۱٤۰۰ تاریخ پذیرش: ۲۹ شهریور ۱٤۰۰

چکیده

یک مدل کمی رابطه ساختار–فعالیت (QSAR) برای پیش‌بینی فعالیت ضد میکروبی مشتقات ۲، ۴، ۶–s–تری آزین به عنوان عامل¬های ضد مالاریا استفاده شد. فعالیت ضد میکروبی ۲۲ مشتقات ۲، ۴، ۶–s–تریازین با توصیفگرهای محاسبات شیمیایی کوانتومی با روش نظریه تابعیت چگالی (DFT) در سطح B3LYP/6-31G و توصیفگرهای توپولوژیکی مدل‌سازی شد. این مطالعه با استفاده از رگرسیون خطی چندگانه (MLR)، تحلیل حداقل مربعات جزئی (PLS) و روش رگرسیون مؤلفه اصلی (PCR) انجام شد. نتایج نشان داد که روش MLR فعالیت ضد میکروبی را به اندازه کافی خوب پیش‌بینی می‌کند. بهترین مدل با شش توصیفگر انتخاب شد. همچنین نشان‌دهنده سازگاری بسیار خوبی با تغییرات داده‌ها برای روش‌های اعتبارسنجی است. مقادیر پیش‌بینی‌شده فعالیت ضد میکروبی مطابقت مناسبی با نتایج تجربی دارد. نتایج به‌دست‌آمده نشان می‌دهد که روش PLS می‌تواند برای پیش‌بینی فعالیت ضد میکروبی مشتقات ۲، ۴، ۶ –s تری آزین مفیدتر باشد. این مطالعه برای پیش‌بینی فعالیت سایر مشتقات این در گروه قابل استفاده می‌باشد.

واژه‌های کلیدی

رگرسیون خطی چند متغیره؛ نظریه تابعیت چگالی؛ حداقل مربع جزئی؛ رگرسیون مولفه اصلی.