

QSAR and Molecular Docking of Echinopsine Derivatives Containing Acylhydrazone Moiety Against Tobacco Mosaic Virus

Nosrat Madadi Mahani¹, Maryam Bagherizadeh¹

Department of Chemistry, Payame Noor University, 19395-4697, Tehran, Iran

Received: 5 May 2024 Accepted: 11 June 2024

DOI: [10.30473/ijac.2024.71494.1298](https://doi.org/10.30473/ijac.2024.71494.1298)

Abstract

Tobacco mosaic virus causes great economic damage to tobacco, pepper, cucumber and ornamental flowers all over the world. In the current work, the relationship between the structure and activity of novel series echinopsin derivatives containing acylhydrazone fragments as antiviral activity against tobacco mosaic virus (TMV) was studied using quantitative structure-activity relationship (QSAR) calculations and molecular docking analysis. Molecular docking analysis of echinopsin derivatives with tobacco mosaic virus (2OM3) protein was done using AutoDock software and descriptors such as binding energy, electrostatic energy and hydrogen bond energy were calculated. The negative values of the binding energy illustrated that the binding nature of these derivatives, as the ligand with the 2OM3 protein is strong. For QSAR model first, the dataset was divided into two groups of training and test sets. Then, descriptors were calculated using quantum mechanics, molecular docking and molecular *descriptors*. Then, modeling was done by multiple linear regression (MLR) method. It was found that of the lowest unoccupied molecular orbital (LUMO) and Gibbs free energy changes play a role in the model. Also, the descriptors of the total energy of van der Waals interactions, hydrogen bond and energy of vdW + H-bond + desolvation of the molecular docking descriptors have an effect in the regression model. This study can play an important role to design anti-TMV inhibitors.

Keywords

Antiviral activity; Multiple linear regression; Lowest unoccupied molecular orbital (LUMO); Gibbs free energy; 2OM3 protein

1. INTRODUCTION

Plant virus diseases can be reducing grain production and can cause many economic losses [1]. Tobacco mosaic virus (TMV) can infect leaves of plants such as tobacco, cucumber, tomato, and pepper [2]. Commonly, antimicrobial agents like Ningnanmycin and Ribavirin are used against TMV [3, 4]. Also, plant virus inhibitor derived from natural products such as purine nucleoside derivative [5], phenanthroindolizidine alkaloid [6], and natural alkaloids derivatives [7] due to many advantages are widely used to control TMV. But only a limited number has been used in agriculture.

Recently, a series of echinopsine derivatives containing acylhydrazone moieties have been synthesized by Cui et al. and their antiviral activities against tobacco mosaic virus (TMV) have been investigated [8]. An important approach for designing novel drugs without the need to

synthesize is Quantitative structure-activity relationship (QSAR) [9]. Also, this approach could be predicting physicochemical properties of novel drugs [10]. On the other hand, molecular docking is one of the most important methods to evaluate the binding compatibility between the ligands and the active residues of protein [11]. Study of Reactivity, molecular docking of antifone analogues as anti-tobacco mosaic virus has been performed by Abdulhassan and coworkers [12]. Xiao et.al have been designed and synthesized a series of novel pyrazole amide derivatives. They have been investigated activity of Anti-Tobacco Mosaic Virus (TMV) and study of molecular docking of these compounds with the binding sites of 2OM3 protein has been performed [13]. Also, Ren and coworkers have been synthesized a series of ferulic acid ester-containing sulfonamid moieties and their anti-TMV activity have been

* Corresponding author:

N. Madadi Mahani; E-mail: nmmadady@pnu.ac.ir, nmmadady@gmail.com

investigated by molecular docking [14]. A series of novel 1, 4-pentadien-3-one derivatives containing triazine moieties have been synthesized and their antibacterial activities against TMV have been evaluated by molecular docking [15]. Long and coworkers have performed QSAR study on A series of novel 1,4-pentadien-3-one derivatives containing 4-thioquinazoline moiety as anti-tobacco mosaic virus which they themselves had designed and synthesized[16].

The aim of this research was to develop QSAR model to predict the activity of of echinopsine derivatives containing acylhydrazone moiety and to investigate the interaction between the echinopsine derivatives and tobacco mosaic virus target site (2OM3).

2.EXPERIMENTAL

27 molecules of echinopsine derivatives containing acylhydrazone moieties with good antiviral activities against tobacco mosaic virus (TMV) with experimental EC_{50} (*Half maximal effective concentration*) values were obtained from the literature [8]. The structures of different groups of studied compounds along with different substituents are shown in Table 1. The chemical structures of the molecules were drawn with ChemDraw Ultra. All compounds were optimized using the GAUSSIAN 09 suite program [17]. The

computation was performed using with density functional theory (DFT) utilizing the B3LYP and 6-31G basis set [18, 19] and quantum descriptors are calculated. The single-crystal x-ray structure of tobacco mosaic virus (TMV) protein was retrieved from Protein Data Bank (PDB ID: 2OM3) [20], and structures of echinopsine derivatives are obtained of the output from DFT optimization.

The preparation of the ligand and the protein files was done by AutoDock Tools. After that, hydrogen was added to the protein and the ligands, and all the water molecules were deleted from the protein structure. The grid box size and center were set at $60 \times 60 \times 60 \text{ \AA}^3$ and 40.52, 57.592, 30.349 for x-, y-, and z-coordinates to allow the ligand to rotate freely. AutoGrid4 and AutoDock4 [21, 22], implemented in MGL tools 1.5.4, were utilized for docking ligands with 2om3 protein.

The structure with the lowest free energy of binding in the highest-populated cluster was chosen as the optimal docking pose, and interactions of the ligand with active sites of protein were analyzed automatically using Ligplus version 2.2.5 [23]. Also, docking descriptors are calculated from docking output. On the other hand, the optimized geometries were loaded into Dragon software [24] to calculate 1481 descriptors in 18 different classes.

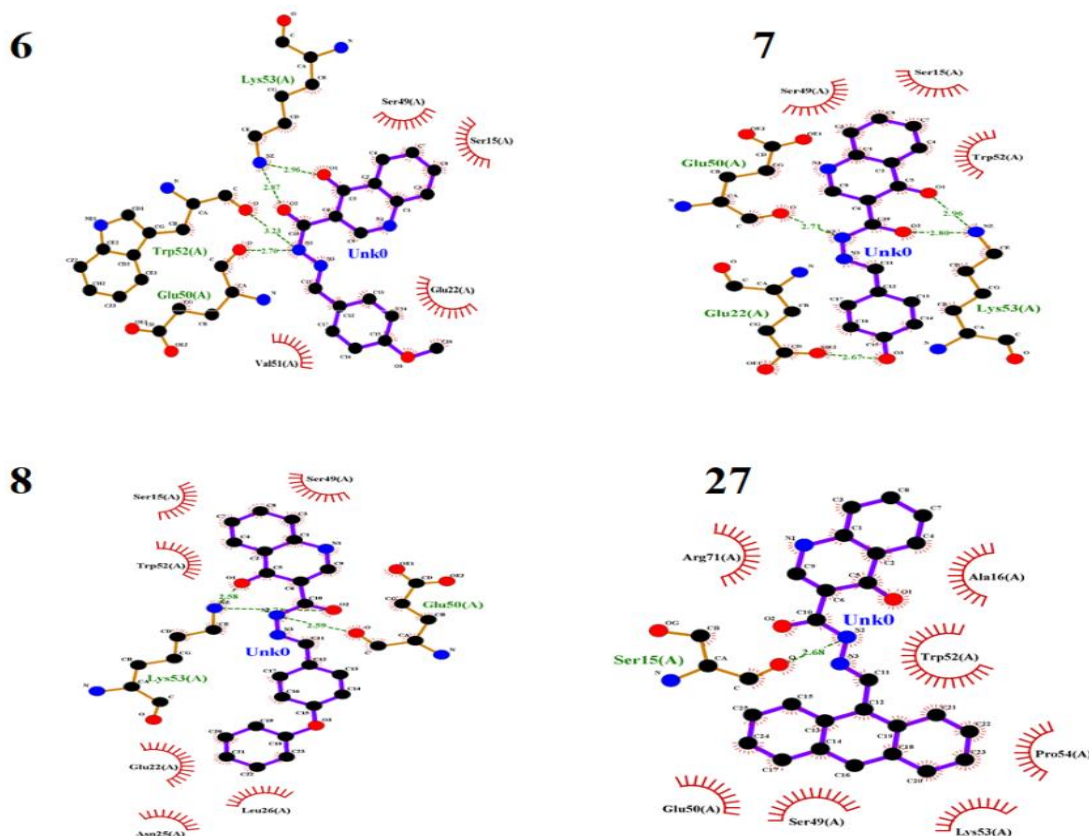


Fig. 1. The docked protein–ligand complexes with hydrogen bond interactions(6,7,8, and 27)

Table 1. Structural features of echinopsine derivatives and EC₅₀(experimental) [8]

Entry	R	EC ₅₀	Entry	R	EC ₅₀	Entry	R	EC ₅₀
1		14.60	10		56.10	19		19.50
2		64.60	11		11.00	20		13.40
3		62.20	12		35.40	21		14.60
4		12.40	13		74.40	22		63.40
5		13.40	14		50.00	23		56.10
6		25.60	15		13.40	24		26.80
7		89.00	16		37.80	25		17.10
8		53.70	17		17.10	26		11.00
9		12.20	18		39.00	27		18.30

3.RESULTS AND DISCUSSION

3.1.Molecular Docking

The molecular docking calculations were performed for systems of 2OM3 protein and 27 molecules of echinopsine derivatives. Docking

properties for the best configuration of 2om3 protein-ligand, such as binding energy, inhibition constant (K_i), intermolecular, and Energy of Vdm+Hbond +desolve. for the best-docked structures, have been listed in Table 2. The negative values of the binding energy illustrated that the binding nature of these derivatives, as the ligand with the 2om3 protein is strong, and compounds of 8, 25, 26, and 27 have the highest binding energy. Also, compounds of 8, 25, 26, and 27 have the smallest inhibition constant (K_i).

Two dimensions structures the best configurations of protein-ligand have generated by using LigPlot program. For instance, the docked protein–ligand complexes with hydrogen bond interactions of compound 6, compound 7, compound 8, and compound 27 illustrated in Fig. 1. Compound 6 forms hydrogen bonds with Lys53, Trp52, and Glu50 residues. Compound 7 forms hydrogen bonds with Glu50, Glu22, and Lys53 residues. Compound 8 forms hydrogen bonds with Lys53 and Glu50 residues. Finally, compound 27 forms hydrogen bonds with Ser15 residues. The hydrogen bonds are created through atoms O and N of echinopsine derivatives. Also, the number of

hydrogen bonds in the complex of compounds 6 and 7 with the 2OM3 protein is more than others.

3.2. QSAR study

A QSAR study was performed for echinopsine derivatives as anti-TMV agents, for characterizing a quantitative relationship between structure chemical and anti-virus 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 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 rely on the training set selection. The data set ($n=27$) was divided casually into two groups: train set ($n = 20$) and test set ($n=7$). The Pearson correlation coefficients are listed in the following table 3. 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.

Table 2. Docking properties for the best docked structure (Energies are in kcal mol⁻¹).

Entry	Binding Energy	No. of Hydrogen bond	K_i (μ M)	Energy of Vdm+Hbond +desolve.	Intermolecular Energy
1	-5.18	0	129.07	-6.19	-6.20
2	-5.59	2	47.02	-6.35	-6.80
3	-5.32	2	96.88	-6.28	-6.37
4	-5.85	1	38.16	-7.29	-7.22
5	-5.54	0	47.77	-6.88	-7.09
6	-5.83	4	46.80	-6.76	-7.10
7	-6.08	4	32.91	-6.68	-7.31
8	-6.43	3	10.95	-7.66	-8.26
9	-5.21	1	73.76	-6.73	-6.83
10	-5.86	3	26.50	-7.31	-7.44
11	-5.04	1	148.18	-5.98	-6.12
12	-5.67	3	63.73	-6.47	-6.61
13	-5.30	1	123.40	-6.24	-6.23
14	-4.75	1	284.80	-6.35	-6.33
15	-5.51	3	80.26	-6.10	-6.48
16	-5.63	3	40.54	-6.51	-6.89
17	-5.69	3	59.33	-6.27	-6.66
18	-5.58	3	67.56	-6.02	-6.58
19	-5.39	3	97.99	-5.93	-6.36
20	-3.34	0	3.22	-4.29	-4.29
21	-5.40	3	95.08	-6.33	-6.38
22	-5.45	2	60.00	-6.62	-6.65
23	-6.18	2	22.15	-7.14	-7.24
24	-5.61	2	66.75	-6.48	-6.59
25	-6.42	1	11.73	-7.76	-7.92
26	-6.31	1	11.43	-7.26	-7.64
27	-6.75	1	10.53	-7.68	-7.69

We used the EC₅₀ of the echinopsine derivatives as the dependent variable. This model with

acceptable statistical quality ($R^2=0.787$) indicated that the anti-virus activity of compounds is

influenced by topological parameters AROM, GATS8p, and R2m, and quantum chemical descriptors of LUMO and ΔG and also, energy of Vdm+Hbond+ desolve. as docking molecular descriptor. Among these descriptors, AROM has the highest average positive effect and energy of vdW+Hbond+desolv has the highest average negative effect. In general, an increase in the amount of descriptors that have an average negative effect leads to a decrease in the anti-virus TMV *activity* of the compounds. In other words, compounds with smaller amounts of these

descriptors will have lower EC_{50} values and more effective molecules in the antiviral activity, and on the contrary, an increase in the amount of descriptors that have an average positive effect leads to an increase in the antiviral activity of the compounds. Therefore, an increase in the values of AROM, R2m, LUMO and ΔG descriptors increases the antiviral properties of compounds and an increase in the values of GATS8p, R2e+ and energy of Vdm+Hbond+desolve. descriptors decreases the activity and antiviral properties of compounds.

Table 3. Correlation coefficient (R^2) matrix for descriptors represented in multiple linear regressions

	AROM	GATS8p	R2m	R2e ⁺	LUMO	ΔG	Energy of Vdm+Hbond+ desolve.
AROM	1						
GATS8p	-0.070	1					
R2m	0.292	0.324	1				
R2e ⁺	0.266	0.113	0.610	1			
LUMO	0.204	0.020	0.125	0.047	1		
ΔG	0.078	0.143	0.348	0.076	-0.054	1	
Energy of Vdm+Hbon d+ desolve.	-0.104	-0.267	-0.205	-0.181	0.265	-0.736	1

Table 4. List of molecular descriptors involved in MLR model

Descriptor	symbol	Coefficient regression	Standard division
Geometrical	AROM	-455.635	180.837
2D autocorrelations	GATS8p	76.658	26.742
Geometry, Topology, and Atom-Weights Assembly (GETAWAY)	R2m	-44.964	22.239
Geometry, Topology, and Atom-Weights Assembly (GETAWAY)	R2e ⁺	1005.623	237.978
Quantum chemistry	LUMO	1.739	0.450
Quantum chemistry	ΔG	-286.887	126.769
Docking molecular	Energy of Vdm+Hbond+ desolve.	-14.422	6.497
-----	Constant	525.405	179.037

The EC₅₀ predicted of echinopsine derivatives by this model is partly like that observed. Fig. 2 displays a very orderly distribution of EC₅₀ values based on the observed values. The leave-one-out (LOO) approach was applied to carry out the cross-validated analysis. Q² (cross-validated coefficient) is computed using the following equation [25, 26]:

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

Where y_i is the i_{th} experimental EC₅₀ value, y_{ipred} is the i_{th} predicted EC₅₀ and y_{imean} is the mean of the experimental EC₅₀. The accuracy of the model was mostly estimated by Root Mean Square Error (RMSE) that calculated using the following equation:

$$\text{RMSE} = \sqrt{\frac{\sum(y_i - y_{ipred})^2}{n}} \quad (\text{Eq. 2})$$

Where n = number of compounds, y_i = experimental value, y_{ipred} = predicted value [27]. The RMSE and Q² of the calibration using MLR method were obtained as 0.419 and 0.781, respectively. Q² is used as a criterion of both validity and predictive ability of the model.

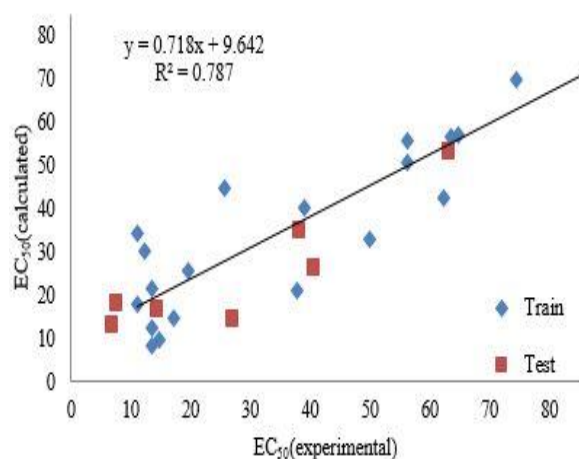


Fig.2. Correlation of EC₅₀(calculated) vs. EC₅₀(experimental).

Table 5. The statistical parameters of different created QSAR models

	Number of series	R ²	SE	F	RMSE	Q ²
Train	20	0.787	5.665	8.523	0.419	0.781
Test	7	0.821	3.746	22.857	0.810	

4. CONCLUSION

Molecular docking as silico approach is performed to explore detailed information about systems of the echinopsine derivatives – 2OM3 protein. The

active sites of 27 echinopsine derivatives in the 2OM3 protein are specified by molecular docking. The residues of Ser15, Arg41, Glu50, Gln45, Lys53, and Gln45 have an essential role in binding with many of echinopsine derivatives. The negative values of the binding energy illustrated that the binding nature of these derivatives, as the ligand with the 2OM3 protein is strong, and compounds of 8, 25, 26, and 27 have the highest binding energy. Also, compounds of 8, 25, 26, and 27 have the smallest inhibition constant (K_i). In this work, a QSAR study was performed for echinopsine derivatives as anti-TMV agents. With considering the error, the prediction of the EC₅₀ values was quite satisfactory and the performance of the QSAR model to predict EC₅₀ value was also calculated using the internal cross-validation method. By defining the molecular descriptors in the regression model, we finalize that an increase in the values of AROM, R2m, LUMO and ΔG descriptors increases the antiviral properties of compounds and an increase in the values of GATS8p, R2e+ and energy of Vdm+Hbond+desolve. descriptors decreases the activity and antiviral properties of compounds. This study can provide clear directions to design anti-TMV inhibitors.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support of this work by Payame Noor University Research

REFERENCES

- [1] K. B. G. Scholthof, S. Adkins, H. Czosnek, P. Palukaitis, E. Jacquot, T. Hohn, B. Hohn, K. Saunders, T. Candresse, P. Ahlquist, C. Hemenway, and G. D. Foster, Top 10 plant viruses in molecular plant pathology, *Mol. Plant. Pathol.* 12 (2011) 938-954.
- [2] M. Chen, S. Su, Q. Zhou, X. Tang, T. Liu, F. Peng, M. He, H. Luo, and W. Xue, Antibacterial and antiviral activities and action mechanism of favonoid derivatives with a benzimidazole moiety, *J. Saudi Chem. Soc.* 25 (2021) 101194-101203.
- [3] D. Wang, M. Huang, D. Gao, K. Chen, Xinxie, W. Xu, and X. Li, Screening anti-TMV agents targeting tobacco mosaic virus helicase protein, *Pestic. Biochem. Physiol.* 166 (2020) 104449-104459.
- [4] W. O. Dawson, and H. Lozoya-Saldana, Examination of the mode of action of ribavirin against tobacco mosaic virus, *Intervirology* 22 (1984) 77-84.
- [5] J. Zhang, F. He, J. Chen, Y. Wang, Y. Yang, D. Hu, and B. Song, Purine nucleoside derivatives containing a sulfa ethylamine

- moiety: Design, synthesis, antiviral activity, and mechanism, *J. Agric. Food Chem.* 69 (2021) 5575-5582.
- [6]L. Li, J. Zou, C. Xu, S. You, Y. Li, and Q. Wang, Synthesis and anti-tobacco mosaic virus/fungicidal/ insecticidal/ antitumor bioactivities of natural product hemigossypol and its derivatives, *J. Agric. Food Chem.* 69 (2021) 1224-1233.
- [7]F. Peng, T. Liu, Q. Wang, F. Liu, X. Cao, J. Yang, L. Liu, C. Xie, and W. Xue, Antibacterial and antiviral activities of 1,3,4-oxadiazole thioether 4H-chromen-4-one derivatives, *J. Agric. Food Chem.* 69 (2021)11085-11094.
- [8]P. Cui , M. Cai, Y. Meng, Y. Yang, H. Song , Y. Liu, and Q. Wang, Design, synthesis and biological activities of echinopsine derivatives containing acylhydrazone moiety, *Sci. Rep.* 12 (2022) 2935-2945.
- [9]K. Y. Wong, A. G. Mercader, L. M. Saavedra, B. Honarparvar, G. P. Romanelli, and P. R. Duchowicz, QSAR analysis on tacrin-related acetylcholinesterase inhibitors, *J. Biomed. Sci.* 21 (2014) 84-91.
- [10]M. Larif, S. Chřita, A. Adad, R. Hmamouchi, M. Bouachrine, and T. Lakhliif, Predicting biological activity of Anticancer Molecules 3-ary 1-4-hydroxyquinolin-2-(1H)-one by DFT-QSAR models, *Inter. J. Clin. Exp. Med.* 3 (2013) 32-42.
- [11]M. A. Lill, and M. L. Danielson, Computer-aided drug design platform using PyMOL, *J. Comput. Aid Mol. Des.* 25 (2011) 13-19.
- [12]H. A. Abdulhassan, B. A. Saleh, D. Harkati, H. Khelfaoui, N. L. Hewitt, G. A. El-Hiti, In Silico Pesticide Discovery for New Anti-Tobacco Mosaic Virus Agents: Reactivity, Molecular Docking, and Molecular Dynamics Simulations, *App. Sci.* 12 (2022) 2818-2827.
- [13]V. Nagalakshamma, M. Venkataswamy, C. Pasala, A. Umamaheswari, K. Thyagaraju, C. Nagaraju, P. V. Chalapathi, Design, synthesis, anti-tobacco mosaic viral and molecule docking simulations of urea/thiourea derivatives of 2-(piperazine-1-yl)-pyrimidine and 1-(4-Fluoro/4-Chloro phenyl)-piperazine and 1-(4-Chloro phenyl)-piperazine - A study, *Bioorg. Chem.* 102 (2020)104084-104095.
- [14]X. Ren, X. Li, L. Yin, D. Jiang, D. Hu, Design, Synthesis, Antiviral Bioactivity, and Mechanism of the Ferulic Acid Ester-Containing Sulfonamide Moiety, *ACS Omega* 5 (2020) 19721-19730.
- [15]X. Tang, M. Chen, J. He, S. Su, R. Xia, T. Guo, S. Jiang, L. Liu, W. Xue, Synthesis and biological activity of 1,4-pentadien-3-one derivatives containing triazine scaffolds, *Peer J. Org. Chem.* 2 (2020) 1-9.
- [16]C. Long , P. Li, M. Chen, L. Dong, D. Hu , and B. Song, Synthesis, anti-tobacco mosaic virus and cucumber mosaic virus activity, and 3D-QSAR study of novel 1,4-pentadien-3-one derivatives containing 4-thioquinazoline moiety, *Eur. J. Med. Chem.* 102 (2015) 639-647.
- [17]M. Frisch, G. Trucks, H. B. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci and G. Petersson, 2009, Gaussian 09, revision a. 02 (Vol. 200, p. 28). Gaussian Inc.
- [18]A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *Chem. Phys.* 58 (1993)5648-5652.
- [19]C. Lee, W. Yang, and 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.
- [20]C. Sachse, J.Z. Chen, P.D. Coureux, M. E. Stroupe, M. Fandrich, and N. Grigorieff, High-resolution electron microscopy of helical specimens: a fresh look at tobacco mosaic virus, *J. Mol. Biol.* 371 (2007) 812-835.
- [21]G. M. Morris, R. Huey, and A. J. Olson, Using AutoDock for ligand-receptor docking, *Curr. Prot. Bioinform.* 24 (2008) 14-19.
- [22]G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785-2791.
- [23]A. C. Wallace, R. A. Laskowski, and J. M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Eng.* 8 (1996) 127.
- [24]A. Mauri, V. Consonni, M. Pavan, and R. Todeschini, Dragon software: an easy approach to molecular descriptor calculations, *MATCH Commun. Math. Comput. Chem.* 56 (2006) 237-248.
- [25]V. Consonni, D. Ballabio, and R. Todeschini, Evaluation of model predictive ability by external validation techniques, *J. Chemometrics* 24 (2010) 194-201.
- [26]P.Y. Lee and C.Y. Chen, Toxicity and quantitative structure-activity relationships of benzoic acids to *Pseudokirchneriella subcapitata*, *J. Hazard. Mater.* 165 (2009) 156-161.

[27]G. Jing, Z. Zhou, and 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.



COPYRIGHTS

© 2022 by the authors. Licensee PNU, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY4.0) (<http://creativecommons.org/licenses/by/4.0>)

رابطه کمی فعالیت - ساختار و اتصال مولکولی مشتقات اکینوپسین حاوی اسیل مطالعه هیدرازون بر علیه ویروس موزاییک تنباکو

نصرت مددی ماهانی^{1*}، مریم باقرزاده¹

بخش شیمی، دانشگاه پیام نور، تهران، ایران

* E-mail: nmmadady@pnu.ac.ir, nmmadady@gmail.com

تاریخ پذیرش: ۲۲ خرداد ماه ۱۴۰۳

تاریخ دریافت: ۱۶ اردیبهشت ۱۴۰۳

چکیده

ویروس موزاییک تنباکو، زیان اقتصادی زیادی به تنباکو، فلفل، خیار و گل‌های زینتی در سراسر جهان وارد می‌نماید. در این مطالعه، یک سری از مشتقات جدید اکینوپسین حاوی اسیل هیدرازون با فعالیت آنتی ویروسی بر علیه ویروس موزاییک تنباکو (TMV) با استفاده از محاسبات رابطه فعالیت کمی-ساختار و اتصال مولکولی مورد بررسی قرار گرفت. آنالیز اتصال مولکولی مشتقات اکینوپسین با پروتئین ویروس موزاییک توتون (2OM3) با استفاده از نرم افزار اتوداک انجام شد و توصیف‌کننده‌هایی مانند انرژی اتصال، انرژی الکترواستاتیک و انرژی پیوند هیدروژنی محاسبه گردید. مقادیر منفی انرژی اتصال نشان داد که ماهیت اتصال این مشتقات، به عنوان لیگاند با پروتئین 2OM3 قوی است. برای مدل QSAR، ابتدا مجموعه داده به دو گروه آموزش و مجموعه تست تقسیم شد. سپس توصیفگرها با استفاده از مکانیک کوانتومی، اتصال مولکولی و توصیفگرهای مولکولی محاسبه شدند. سپس مدلسازی به روش رگرسیون خطی چندگانه (MLR) انجام شد. مشخص شد که انرژی کمترین اوربیتال مولکولی اشغال نشده (LUMO) و تغییرات انرژی آزاد گیبس در مدل نقش دارند. همچنین توصیفگرهای داکینگ مولکولی مانند توصیفگرهای انرژی کل برهمکنش‌های واندروالس، پیوند هیدروژنی و انرژی مجموع واندروالس، هیدروژنی و حل‌شوندگی در مدل رگرسیونی تأثیر دارند. این مطالعه می‌تواند نقش مهمی در طراحی مهارکننده‌های ضد ویروس TMV داشته باشد.

کلید واژه‌ها

فعالیت ضدویروسی، رگرسیون چند متغیره خطی، انرژی آزاد گیبس، انرژی پایین ترین اوربیتال مولکولی اشغال نشده، پروتئین 2OM3