How to Cite: Asgaonkar, K.D., Patil, S.M., Chitre, T.S., Wani, S.D. & Singh, M.T. (2022) QSAR tool for optimization of nitrobenzamide pharmacophore for antitubercular activity. *Bulletin of the University of Karaganda – Chemistry, 105(1), 60-68.* https://doi.org/10.31489/2022Ch1/60-68

Article

UDC 544.42+519.242.7 https://doi.org/10.31489/2022Ch1/60-68

K.D. Asgaonkar*, S.M. Patil, T.S. Chitre, S.D. Wani, M.T. Singh

AISSMS College of Pharmacy, Pune, Maharashtra, India
*(*Corresponding author's e-mail: kalyani_a@aissmscop.com)*

QSAR tool for optimization of nitrobenzamide pharmacophore for antitubercular activity

Tuberculosis (TB) is a leading cause of death worldwide from a single infectious agent, *Mycobacterium tuberculosis* (MTB), especially due to the development of resistant strains and its co-infections in HIV. Quantitative-structure activity relationship (QSAR) studies aid rapid drug discovery. In this work, 2D and 3D QSAR studies were carried out on a series of nitrobenzamide derivatives to design newer analogues for antitubercular activity. 2D QSAR was performed using MLR on a data set showing antitubercular activity. The 3D-QSAR studies were performed by kNN–MFA using simulated annealing variable selection method. Alignment of given set of molecules was carried out by the template-based alignment method and then was used to build the 3D-QSAR model. Robustness and predictive ability of the models were evaluated by using various traditional validating parameters. Different physiochemical, alignment-based, topological, electrostatic, and steric descriptors were generated, which indicated the key structural requirements for optimizing the pharmacophore for better antitubercular activity. For 2D QSAR, the best statistical model was generated using SA-MLR method (r^2 = 0.892, q^2 = 0.819) while 3D QSAR model was derived using the SA KNN method (q^2 = 0.722). The positively contributing descriptors can be incorporated to design new chemical entities for future study.

Keywords: tuberculosis, 2D QSAR, 3D QSAR, nitrobenzamide, SA-MLR, SA-kNN, pharmacophore, antitubercular activity.

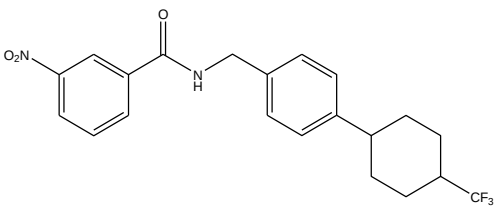
Introduction

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* (MTB). It has emerged as a global health menace due to drug resistance strains, such as multidrug-resistant, totally drug-resistant, and extremely drug-resistant TB. Also, its co-existence with HIV makes it even more challenging to treat [1]. Moreover, COVID–19 pandemic threatens the progress in reducing the global burden of TB disease [2]. This necessitates rapid drug development in this area. One way to achieve this is by applying statistical analytical methods as quantitative structure–activity relationship (QSAR). This technique is valuable as it helps to narrow down a library of molecules to effective potential inhibitors by predicting biological activities [3–5]. In the present study, 2D and 3D QSAR studies were carried out on nitrobenzamide derivatives to optimise the pharmacophore for antitubercular activity.

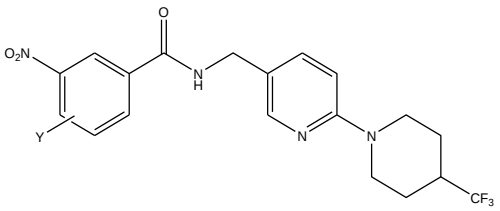
Experimental

All QSAR studies were performed using V-Life sciences MDS Version 4.3 [6].

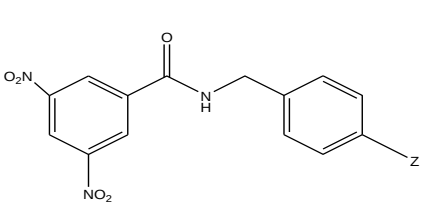
Data set: A data set of 24 nitrobenzamide derivatives with chemical and biological variation processing antitubercular activity reported by Wang H. et al. was used for the QSAR study (Table 1) [7]. Biological activity expressed as Minimum Inhibitory concentration (MIC, μM) values was converted into pMIC values [pMIC = -log (MIC)]. QSAR structures of the compounds were drawn using the ChemDraw tool and converted into 3D structures (.mol2) using the V life MDS software. Geometry optimisation of the structures was carried out using the standard Merck Molecular Force Field (MMFF).



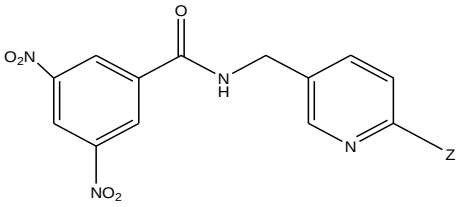
Compound 1-2



Compound 7-8



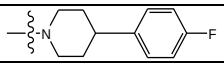
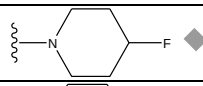
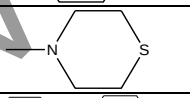
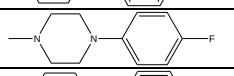
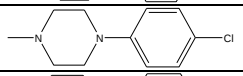
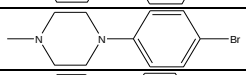
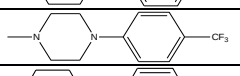
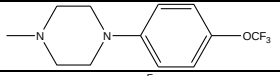
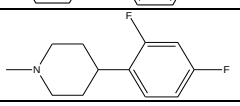
Compound 3-6



Compound 9-24

T a b l e 1

Data set of 24 nitrobenzamide derivatives for the QSAR study

Compound No.	Y-	Z-	MIC(μM)	pMIC
1	5-F	-	1.357	-0.599
2*	5-Br	-	0.459	0.055
3	-		0.060	0.903
4*	-	-F	0.120	-0.574
5	-	$-CF_3$	0.059	-0.201
6	-	$-OCF_3$	0.033	-1.346
7	$5-NO_2$	-	0.059	0.886
8	5-Br	-	0.944	-0.288
9	-		0.094	0.632
10	-		0.030	1.146
11			0.030	1.189
12	-		0.108	0.591
13*	-		0.059	0.835
14	-		0.119	0.603
15	-		0.452	0.040
16	-		0.235	0.361
17	-		0.480	0.042
18			1.255	-0.361
19	-		0.210	0.359
20	-		0.178	0.446
21*	-		0.233	0.366
22*	-		0.491	0.033
23	-		0.973	-0.250
24	-		0.143	0.542
*Test set				

Generation of training and test set: Entire data set of 24 compounds was distributed as training set (19 molecules) and test set (5 molecules) using the sphere exclusion method. The selection of the test compounds was made based on their biological activity, structural diversity, and activity distribution plot. Unicolumn statistics for both training and test set was applied to check rightness of selection criteria for training and test set molecules. The mean of the test set was higher than the mean in the training set, indicating the presence of relatively more active molecules than inactive ones. A higher standard deviation in the training set indicated a wide distribution of the molecules' activity compared to the test set molecules.

QSAR studies: 2D and 3D QSAR were computed using various statistical models. Robustness and predictive ability of the models was evaluated by using various traditional validating parameters for internal validation–correlation coefficient (r^2), cross-validated correlation coefficient (q^2) and external validation (pred_r^2) [3,4,5,8-11].

2D QSAR studies: V life MDS software can calculate various 2D descriptors such as physicochemical and alignment-based. While calculating the physicochemical descriptors, dipole moment, distance-based topological indices, electrostatic, hydrophobic descriptors were deselected as they are 3D descriptors. A molecular descriptor based upon a counting statistic of the topological distance matrix is used in QSAR studies. Thus Baumann alignment independent topological descriptors with attributes 2, T, C, N, O, F, S, and Cl, were selected. These topological descriptors provide an idea about the desired 2D pharmacophoric features. Correlation matrix was applied to select the predominant descriptors influencing the antitubercular activity of the analogues taking each descriptor as independent and pMIC as dependent variable. Descriptors showing the highest correlation with pMIC were selected for generation of the QSAR model using multiple linear regression (MLR), Partial Least Square (PLS), Principal component regression (PCR). Regression methods were performed by selecting Set Cross-Correlation Limit as 0.5, Number of variables in final Equation as 10, Term Selection criteria as r^2. Various models were generated and were analysed using the fitness plot, contribution plot, and statistical parameter compliance.

3D QSAR Studies: The 3D-QSAR studies were performed by kNN–MFA using simulated annealing variable selection method. KNN–MFA method requires suitable alignment of a given set of molecules. After optimization, alignment was carried out by the template-based alignment method. This was followed by generating common rectangular grid around the molecules (Figure 1). The resulting set of aligned molecules was then used to build 3D-QSAR models and information generated was used to predict activity of those designed molecules that have a similar template or set of atoms. Steric, hydrophobic and electrostatic interaction energies were computed at the lattice points of the grid.

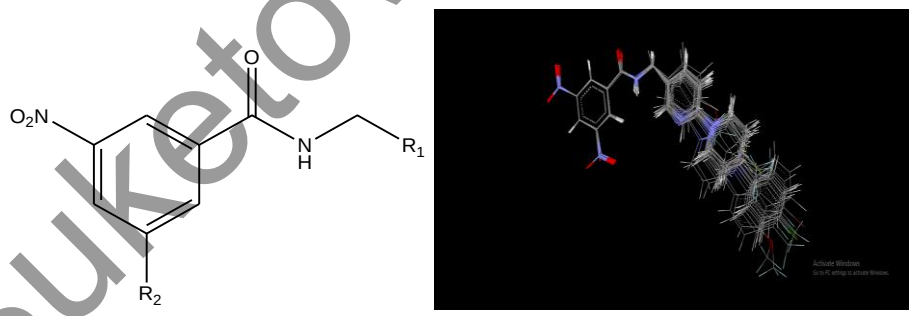


Figure 1. Common template used and alignment of nitrobenzamide derivatives

Results and Discussion

2D QSAR: Amongst the various 2D QSAR methods developed, SA-MLR method demonstrated the best results as given in Equation (1):

$$\text{pMIC} = -0.2140\ (\pm 0.0249)\ (\text{T\_T\_F\_5}) + 0.3571\ (\pm 0.0455)\ (\text{SssCH2count}) - -0.2377\ (\pm 0.0455)\ (\text{SaasCE-index}) \quad (1)$$

This model produced a correlation coefficient r^2=0.8922, cross-validated correlation coefficient q^2 = 0.8197 and pred_r^2 = 0.7356 (Table 2). The observed and predicted activities of the test and training sets are shown in Table 3. Contribution plot of descriptors is depicted in Figure 2. This plot describes the extent (percentage) to which different descriptors influence biological activity. Also, the plot of actual versus predicted activity of training set and test set is shown in Figure 3. Positive descriptors favor biological activity,

whereas negative descriptors would have a detrimental effect on biological activity. Hence while designing new chemical entities, positively contributing descriptors are favored, and negative descriptors should be avoided. In the present study, the positively contributing descriptors SssCH2count indicated that a total number of $–CH_2$ group connected with two single bonds would increase activity. Negatively contributing descriptors T_T_F_5 indicated that any atom separated from fluorine by five bond distance would result in decrease of activity. Negative Saas CE-index indicated electrotopological state indices for a number of carbon atoms connected with one single bond along with two aromatic bonds would decrease the antitubercular activity.

Table 2

Statistical parameters of 2D-QSAR model

Statistical parameter	Regression method SA-MLR
N	19
r^2	0.892
q^2	0.819
Pred $_r^2$	0.735
Pred $_r^2{}_{se}$	0.336
F test	0.336
$r^2{}_{se}$	0.225
$q^2{}_{se}$	0.292
Best Rand r^2	0.447
Best Rand q^2	0.227
Z score R^2	8.317
Z score Q^2	5.618
Alpha Rand R^2	0.000
Alpha Rand Q^2	0.000
Descriptors	T_T_F_5 SssCH2count SaasCE-index
Coefficients	-0.214(±0.024) 0.357(±0.045) -0.237(±0.045)

Table 3

Observed, predicted and residual values for training set and test set

Compound no.	Observed activity (pMIC)	Predicted activity	Residual activity
1	2	3	4
1	-0.599	0.049	-0.648
2*	0.055	-0.507	0.562
3	0.903	0.514	0.388
4*	-0.574	-0.521	-0.052
5	-0.201	0.493	0.694
6	-1.346	0.0507	-1.853
7	0.886	0.814	0.071
8	-0.288	-0.371	0.083
9	0.632	0.496	0.139
10	1.146	0.915	0.231
11	1.189	0.958	0.230
12	0.591	0.664	-0.072
13*	0.835	-0.156	0.991
14	0.603	0.559	0.442
15	0.040	0.593	-0.553
16	0.361	0.456	-0.094

Continuation of Table 3

1	2	3	4
17	0.042	0.335	-0.293
18	-0.361	0.226	-0.587
19	0.359	0.443	-0.084
20	0.446	0.411	0.035
21*	0.366	0.226	0.139
22*	0.033	-0.452	0.485
23	-0.250	0.336	-0.586
24	0.542	0.522	0.02
		RMSE	0.558

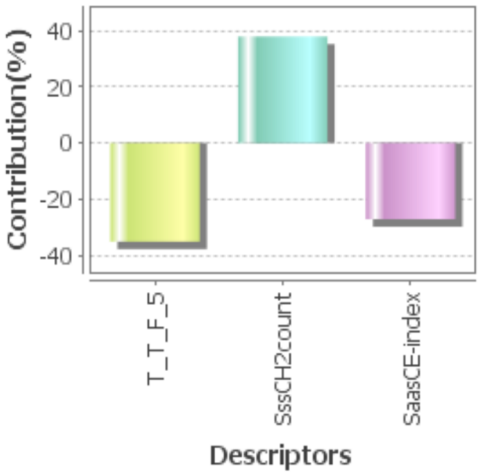


Figure 2. Contribution of descriptors

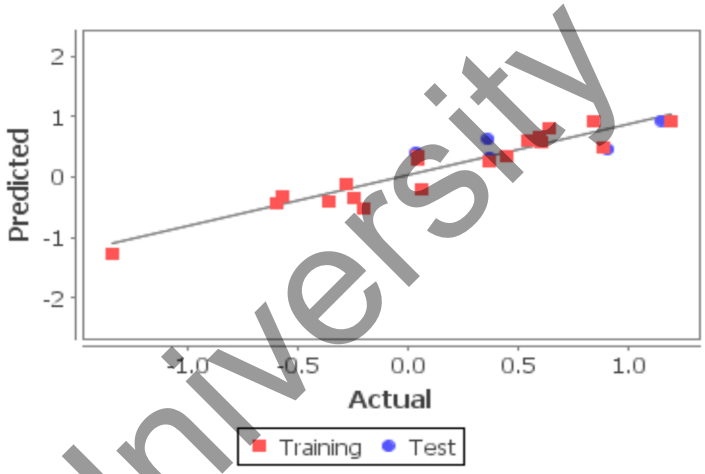


Figure 3. Plot of actual versus predicted Activity of training set and test set of SA-MLR method

3D QSAR: In continuation to 2D QSAR studies, 3D QSAR SA kNN MFA models were also commuted. The statistical results generated by SA-kNN MFA methods are depicted in Table 4. The q^2, pred_r^2, pred_r^2 se and K values of model were found to be statistically significant hence model was considered for designing of NCE's. The 3D data point descriptors were generated in rectangular grid according to the range of contribution mentioned in parenthesis using SA kNN-MFA are depicted in Figure 4. Experimental and predicted activities are shown in Table 5. In model, residuals obtained are near to zero indicating a good predicting ability of the model. The plots of observed vs. predicted activity for the optimal cross-validated kNN-QSAR model are depicted in Figure 5.

Table 4

Statistical results of 3D QSAR generated by SA kNN-MFA methods

Statistical Parameters	SA-kNN MFA
k Nearest Neighbour	4
N	19
Degree of freedom	14
q^2	0.722
q^2_ se	0.330
pred_r^2	0.879
pred_r^2se	0.227
Contributing descriptors	E_671 -0.0228 -0.0198 S_943 -6.0242 -3.0694 E_580 1.1626 4.0054 S_901 -0.8031 -0.4015

To visualise the information contained in the 3D-QSAR models, grid was generated. The electrostatic and steric descriptors are shown in Figure 4. Points generated in SA kNN-MFA 3D-QSAR model were E_671 (-0.0228-0.0198), S_943(-6.0242-3.0694), E_580 (1.1626 4.0054), S_901(-0.8031 -0.4015) i.e., electrostatic and steric interaction at lattice points 671,580 and 943,901, respectively. Negative values in electrostatic field descriptors indicated that negative electronic potential is required to increase activity and more electronegative substituent is preferred on the aryl group. Similarly, negative values of steric descriptors revealed that less sterically bulky aryl groups are favorable for maximum activity.

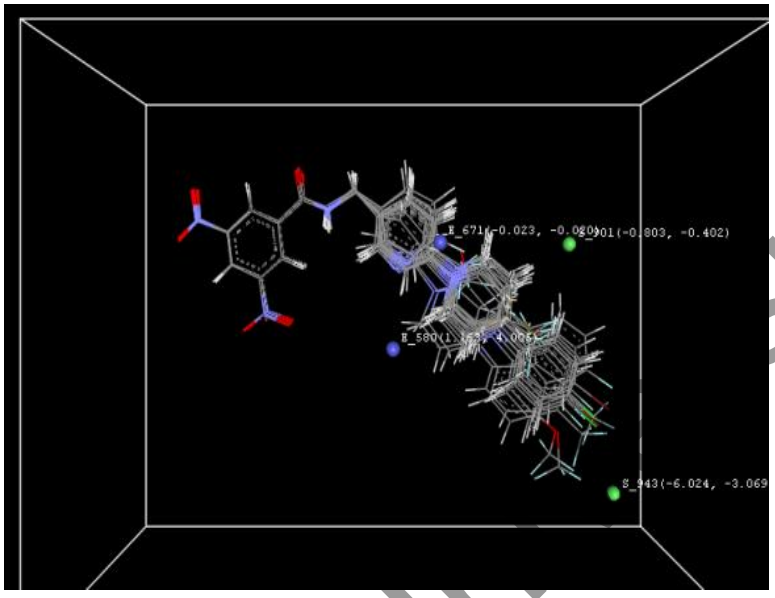


Figure 4. Generated data point

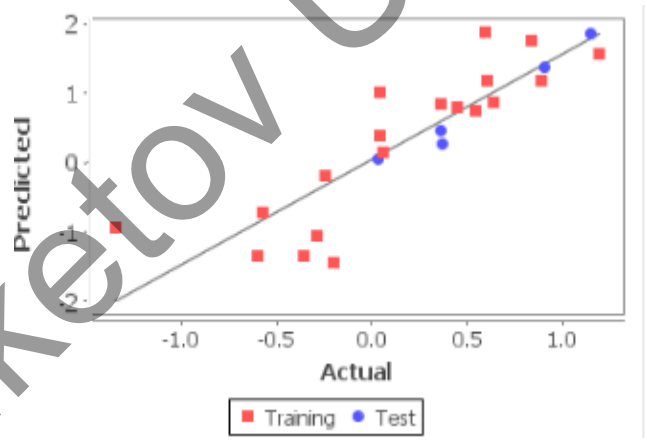


Figure 5. Plot of actual versus predicted activity of training set and test set of SA-kNN MFA

Table 5

Experimental and predicted activities

Compound no.	Actual activity (pMIC)	Model	
		Predicted activity	Residual activity
1	2	3	4
1	-0.599	-0.609	0.010
2	0.055	-0.529	0.585
3*	0.903	0.780	0.122
4	-0.574	-0.625	0.151
5	-0.201	-0.215	0.014

Continuation of Table 5

1	2	3	4
6	-1.346	-1.001	-0.344
7	0.886	0.165	0.720
8	-0.288	0.529	-0.817
9	0.632	0.448	0.184
10*	1.146	1.137	0.009
11	1.189	1.132	0.057
12	0.591	-0.184	0.775
13	0.835	1.239	-0.404
14	0.603	0.443	0.160
15	0.040	0.586	-0.546
16	0.361	0.588	-0.227
17	0.042	0.259	-0.217
18	-0.361	-0.648	0.287
19*	0.359	0.043	0.315
20	0.446	0.187	0.258
21*	0.366	0.189	0.176
22*	0.0330	-0.140	0.173
23	-0.250	-1.179	0.929
24	0.542	-0.195	0.737
		RMSE	0.438

Conclusions

In order to optimise pharmacophore for antitubercular activity, a data set of nitrobenzamide derivatives was selected to perform QSAR studies. 2D and 3D QSAR were performed using MLR and SA kNN methods, respectively. Statistically significant models were used for interpretation. The study indicated a positive contribution of 2D descriptors (SssCH2count), 3D descriptors (more electronegative substituent on the aryl group and less sterically bulky aryl groups) are favorable for maximum antitubercular activity (Fig. 6). Considering different descriptors generated from 2D and 3D QSAR, new chemical entities can be designed for further studies.

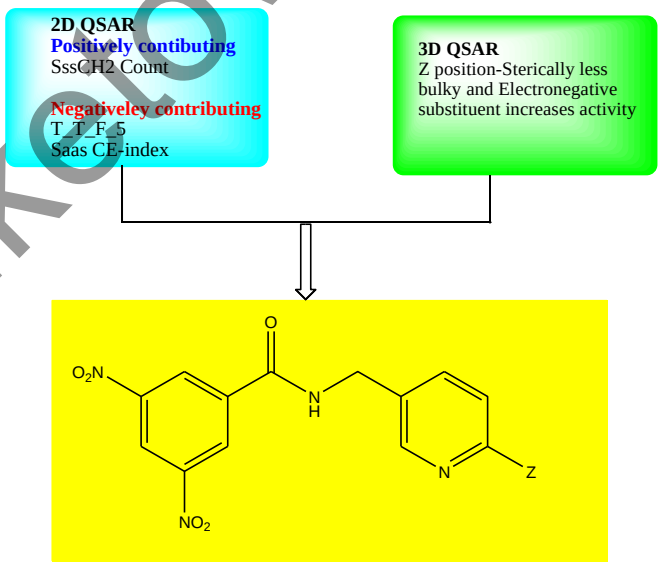


Figure 6. Optimised Pharmacophore

Acknowledgments

The authors are grateful to Dr. Ashwini R. Madgulkar, Principal of AISSMS College of Pharmacy, for continuous motivation, support, and providing the necessary infrastructure to carry out this work.

References

1 Bhat, Z.S., Rather, M.A., Maqbool, M., & Ahmad, Z. (2018). Drug Targets Exploited in Mycobacterium Tuberculosis: Pitfalls and Promises on the Horizon. *Biomedicine & Pharmacotherapy*, *103*, 1733–1747. https://doi.org/10.1016/j.biopha.2018.04.176

2 Global Tuberculosis Report 2020. World Health Organization https://www.who.int/publications/i/item/9789240013131

3 Yuanita, E., Sudirman, Dharmayani NKT., Ulfa, M., & Syahri, J. (2020). Quantitative structure–activity relationship (QSAR) and molecular docking of xanthone derivatives as anti-tuberculosis agents. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, *21*, 100203. https://doi.org/10.1016/j.jctube.2020.100203

4 Bhadoriya, K.S., Kumawata, N.K., Bhavthankara, S.V., Avchara, M.H., Dhumala, D.M., Patil, S.D. et al. (2016). Exploring 2D and 3D QSARs of benzimidazole derivatives as transient receptor potential melastatin 8 (TRPM8) antagonists using MLR and kNN-MFA methodology. *Journal of Saudi Chemical Society*, *20*, S256–S270. https://doi.org/10.1016/j.jscs.2012.11.001

5 Qiao, K., Fu, W., Jiang, Y., Chen, L., Li, S., Ye, Q., & Gui W. (2020). QSAR models for the acute toxicity of 1,2,4-triazole fungicides to zebrafish (Daniorerio) embryos. *Environmental Pollution*, *265*, 114837. https://doi.org/10.1016/j.envpol.2020.114837

6 Vlife MDS SUITS manual version 4.3

7 Wang, H., Lva, K., Lia, X., Wang, B., Wang, A., Tao, A et al. (2018). Design, synthesis and antimycobacterial activity of novel nitrobenzamide derivatives. *Chinese Chemical Letters; 30(2):* https://doi.org/10.1016/j.cclet.2018.08.005

8 Zambre, V.P., Patil, R.B., Rajesh, B., Sangshetti, J.N., & Sawant, S.D. (2019).Comprehensive QSAR studies reveal structural insights into the NR2B subtype selective benzazepine derivatives as N-Methyl-D-Aspartate receptor antagonists. *Journal of Molecular Structure*, *1197*, 617–627. https://doi.org/10.1016/j.molstruc.2019.07.068

9 Kumar, A., Narasimhan, B., & Kumar, D. (2007). Synthesis, antimicrobial, and QSAR studies of substituted benzamides. *Bioorganic & Medicinal Chemistry*, *15 (12)*, 4113–4124.

10 Chitre, T.S., Asgaonkar, K.D., Patil, S.M., Kumar, S., Khedkar, V.M., & Garud, D.R. (2017). QSAR, docking studies of 1,3-thiazinan-3-yl isonicotinamide derivatives for antitubercular activity. *Computational Biology and Chemistry*, *68*, 211–218. https://doi.org/10.1016/j.compbiolchem.2017.03.015

11 Chen, J.C., Qian, L., Wu, W.J., Chen, L.M., & Zheng, K.C. (2005). QSAR study of substituted benzophenazines as potential anticancer agents. *Journal of Molecular Structure THEOCHEM*, *756*, 167–172. https://doi.org/10.1016/j.theochem.2005.09.010

К.Д. Асгаонкар, С.М. Патил, Т.С. Читре, С.Д. Вани, М.Т. Сингх

Туберкулезгеқарсы белсенділік үшін нитробензамид фармакофорасының QSAR оңтайландыруы

Туберкулез (ТБ) бүкіл әлемде *Mycobacterium tuberculosis (MTB)* сияқты инфекциялық агенттен, әсіресе төзімді штамдардың дамуына және оның АИТВ-мен бірге жұқтыруына байланысты жұқпалы қоздырғышынан өлімнің негізгі себебі болып табылады. Химиялық қосылыстардың құрылымы мен белсенділігі (QSAR) арасындағы байланысты сандық зерттеу жаңа препараттың даму процесін едәуір жылдамдатуға көмектеседі. Осы жұмыста туберкулезгеқарсы белсенділігі бар жаңа аналогтарды әзірлеу мақсатында нитробензамид туындыларының бірқатарына 2D және 3D QSAR-зерттеулер жүргізілді. 2D QSAR туберкулезгеқарсы белсенділікті көрсететін мәліметтер жиынтығында MLR әдісін қолдану арқылы жасалды. 3D-QSAR зерттеулері kNN-MFA алгоритмін қолданып, күйдіру процесін модельдейтін айнымалыларды таңдау әдісі арқылы орындалды. Берілген молекулалар жиынтығын тегістеу шаблонға негізделген тегістеу алгоритмі арқылы жүргізілді, содан кейін 3D-QSAR моделін құру үшін қолданылды. Модельдердің сенімділігі мен болжау қабілеті әртүрлі дәстүрлі тексеру параметрлерін қолдана отырып бағаланды. Теңестіруге негізделген әртүрлі физика-химиялық, топологиялық, электростатикалық және стерильді дескрипторлар анықталды, олар туберкулезгеқарсы белсенділігі жоғарылаған фармакофорды оңтайландыру үшін негізгі құрылымдық талаптарды көрсетті. 2D QSAR үшін ең жақсы статистикалық модель SA-MLR (r^2 = 0.892, q^2 = 0.819) әдісін қолдана отырып жасалды, ал 3D qsar моделі SA KNN (q^2 = 0.722) алгоритмін қолдана отырып алынды. Анықталған дескрипторларды әрі қарайғы зерттеулерде жаңа химиялық туындыларды әзірлеу үшін пайдалануға болады.

Кілт сөздер: туберкулез, 2D QSAR, 3D QSAR, нитробензамид, SA-MLR, SA-kNN, фармакофор, туберкулезгеқарсы белсенділік.

К.Д. Асгаонкар, С.М. Патил, Т.С. Читре, С.Д. Вани, М.Т. Сингх

QSAR-оптимизации фармакофора нитробензамида для противотуберкулезной активности

Туберкулез является ведущей причиной смерти во всем мире от такого инфекционного агента, как *Mycobacterium tuberculosis*, особенно вследствие развития резистентных штаммов и его коинфекции при ВИЧ. Количественные исследования взаимосвязи структуры и активности (QSAR) химических соединений помогают значительно ускорить процесс разработки нового лекарства. В статье проведены 2D и 3D QSAR-исследования ряда производных нитробензамида с целью разработки новых аналогов с противотуберкулезной активностью. 2D QSAR был выполнен с использованием MLR-метода на наборе данных, показывающих противотуберкулезную активность. Исследования 3D-QSAR были выполнены с помощью kNN-MFA-алгоритма с использованием метода выбора переменных, моделирующих отжиг. Выравнивание заданного набора молекул проводилось с помощью алгоритма выравнивания на основе шаблона, а затем использовалось для построения модели 3D-QSAR. Надежность и прогностическая способность моделей оценивались с помощью различных традиционных параметров проверки. Были выделены различные физико-химические, основанные на выравнивании, топологические, электростатические и стерические дескрипторы, которые указывали на ключевые структурные требования для оптимизации фармакофора с повышенной противотуберкулезной активностью. Для 2D QSAR наилучшая статистическая модель была создана с использованием метода SA-MLR (r^2 = 0,892, q^2 = 0,819), тогда как модель 3D QSAR была получена с применением алгоритма SA KNN (q^2 = 0,722). Выявленные дескрипторы могут быть полезны для разработки новых химических производных в дальнейших исследованиях.

Ключевые слова: туберкулез, 2D QSAR, 3D QSAR, нитробензамид, SA-MLR, SA-kNN, фармакофор, противотуберкулезная активность.

Information about authors

Asgaonkar, Kalyani D. — Assistant Professor in Dept. of Pharmaceutical chemistry. AISSMS College of Pharmacy, Kennedy Road, Pune, Maharashtra, India. e-mail: kalyani_a@aissmscop.com. https://orcid.org/0000-0003-3780-5323

Patil, Shital M. — Assistant Professor in Dept. of Pharmaceutical chemistry, AISSMS College of Pharmacy, Kennedy Road, Pune, Maharashtra, India. e-mail: shital_patil@aissmscop.com. https://orcid.org/0000-0003-2692-2380

Chitre, Trupti S. — Associate Professor in Dept. of Pharmaceutical chemistry, AISSMS College of Pharmacy, Kennedy Road, Pune, Maharashtra, India. e-mail: drugdesign2@gmail.com. https://orcid.org/0000-0002-1722-7215

Wani, Sadeecha D. — M.Pharm student, Dept. of Pharmaceutical chemistry, AISSMS College of Pharmacy, kennedy Road, Pune, Maharashtra, India. e-mail: sadeecha.wani@gmail.com. https://orcid.org/0000-0003-2568-2338

Singh, Muskan T. — M.Pharm student, Dept. of Pharmaceutical chemistry, AISSMS College of Pharmacy, kennedy Road, Pune, Maharashtra, India. e-mail: muskansingh5735@gmail.com. https://orcid.org/0000-0002-8130-6401