Indian Journal of Chemistry
Vol. 59B, December 2020, pp 1881-1886

QSAR of 1,3,5-triazine compounds towards inhibition of toxoplasmosis utilizing computed molecular descriptors

Tanya & Sisir Nandi*

Department of Pharmaceutical Chemistry, Global Institute of Pharmaceutical Education and Research, Affiliated to Uttarakhand Technical University, Kashipur 244 713, India

E-mail: sisir.iicb@gmail.com

Received 25 June 2020; accepted (revised) 11 November 2020

Growing attention has been focussed towards a considerable amount of experimental and theoretical study of anti toxoplasmosis compounds having inhibitory activities against toxoplasmosis. Toxoplasmosis is a dangerous disease in both urban as well as rural areas. It is caused by *Toxoplasma gondii*. The life cycle involves feline species such as both domestic and wild cats, and other felids such as lions, *etc*. There are various targets for developing anti toxoplasmosis agents. One of the most promising targets is dihydrofolate reductase (DHFR). 1,3,5-Triazine compounds have been reported to inhibit the *T. gondii*. But there is hardly any formulation of quantitative structure-activity relationship (QSAR) involving 1,3,5-triazine inhibitors to date. Therefore, it is our target in the present study to develop QSAR models based on the computed theoretical molecular descriptors to scale the essential features responsible for the DHFR inhibition. These screened features of the selected compounds will help to design the potent congeneric series.

Keywords: Toxoplasmosis, *Toxoplasma gondii*, 1,3,5-triazine compounds, calculated theoretical molecular descriptors, QSAR modeling

Toxoplasmosis is a major zoonotic parasitic disease that transmits from animals to human beings. In the life cycle of the disease, warm-blooded animals like feline cat and intermediate host human beings are involved. The infection is transmitted through the eating of infected meat and vegetables and also by drinking infected water, unpasteurized milk, and blood transfusion[1]. The microorganism responsible for the toxoplasmosis is *Toxoplasma gondii* which approximately infects a third of the world's population *i.e*. one-third population is serum positive. This is the most dangerous parasite[2]. The deadly impact of infection during pregnancy and its reactivation in immune-compromised patients are well known for a long time although progress has been made in the field of its diagnosis, epidemiology, and understanding of the relationship between parasite and cell[3]. In expectant women and patients with a compromised immune system, *T. gondii* causes grave damage and fetus loss. The graveness of infection is dependent on the gestation stage. The first trimester and the last trimester have low and high transmission risks respectively[4].

The introduction of antiretroviral therapy in high-income countries cause the reduction of HIV-associated *Toxoplasma* but still not in low-income countries[5]. The therapy targets a different *T. gondii* enzymes such as farnesyl diphosphate/geranylgeranyl-diphosphate, bumped kinase, fatty acid synthase II (FAS II), histone acetyltransferase/histone deacetylase, and dihydrofolate reductase (DHFR). Dihydrofolate is converted to tetrahydrofolate by the DHFR enzyme which is present in living organisms such as bacteria, plants, and mammals. Inhibition of DHFR enzyme causes disturbance in the synthesis of protein, RNA, DNA because the level of THF becomes lower which is responsible for the synthesis. So, the disease caused by these pathogens can weaken our immune system and DHFR inhibition is the most important and effective treatment for it[6].

Pyrimethamine along with sulfadiazine which targets DHFR is the first-line treatment for toxoplasmosis. Pyrimethamine is a folic acid antagonist and can cause dose-related suppression of the bone marrow, which is mitigated by concurrent administration of folinic acid (leucovorin). Leucovorin is added to the above combination to prevent hematologic toxicity. Pyrimethamine-sulfadiazine was later discontinued owing to toxicity and severe side effects in 62% and 44% of patients respectively. The condition of the above trials were toxoplasma encephalitis. Clindamycin was an able substitute for sulfadiazine for

patients susceptible to sulfa drugs. Clindamycin however couldn't prevent relapse and the same level of toxicity. In patients resistant to sulfa susceptibility sulfadiazine was replaced with sulfamethoxazole. For patients who couldn't tolerate pyrimethamine or due to its unavailability, it was replaced with trimethoprim. Azithromycin and atovaquone can be used in place of sulfadiazines or pyrimethamine in the event of contraindications. The combination of pyrimethamine and sulfadiazine is also known to cause serious side effects such as Stevens-Johnson syndrome, liver necrosis, toxic necrosis, *etc*. It is widely known that congenital infection, ocular infection, and accidental infection require a prolonged treatment regime of 4-6 weeks to at least one year to kill the cysts formed in tissue. Atovaquone is a promising candidate and is known to act against tissue cysts in *T. gondii*. At the same time, atovaquone has failed to prevent relapse in the cases of ocular infection and encephalitis[7-11]. Therefore it requires the development of new drugs with reduced toxicity and improved efficacy towards the prevention of infection. As DHFR is a potential target of *T. gondii*, it has been tried to design and discover new DHFR inhibitors. The potent class of drugs acting against DHFR is dihydrotriazines. The dihydrotriazines have been effective against both malaria and *T. gondii*. But there are hardly any studies reported on the design of this class of congeners using quantitative structure-activity relationships (QSAR). So it is our target in the present study to develop QSAR models using computed structural properties of these compounds. The modeled parameters can give a clue for the crucial features for further design of the potent congeneric compounds.

Computational Methods

Several 60 halogenated compounds based on 1,3,5-triazine[12] with varied potency toward *T. gondii* dihydrofolate reductase inhibition have been taken into consideration to form the activity dataset (Table I). Activity data is given in the following Table I. The structures of all compounds under study were drawn in the 2D ChemDraw window. The drawn structures were then imported into 3D modules using the default conversion procedure implemented in the CS Chem3D Ultra[13]. A further process is to minimize the energy of these compounds up to 0.01 by the MM2 molecular field. All the minimized ligands were browsed into PaDEL software[14] to calculate many 1875 theoretical molecular descriptors including electrostatic, topological, constitutional, geometrical, and physicochemical, useful for our purpose, and before model development, these were reduced to 672.

Descriptors with perfectly constant and highly inter-correlated descriptors were removed considering variance and correlation coefficient cut-off values of 0.0001 and 0.99 using the V-WSP algorithm[15] incorporated into the vWSP module of Nano BRIDGES software[16]. The reduced set of theoretical molecular descriptors was taken into consideration for the development of quantitative structure-activity relationship models.

As the number of structural predictors greatly exceeds the number of compounds, the selection of important predictors is necessary for QSAR modeling. A genetic algorithm coupled with multiple linear regression methodology has been applied in the present work to screen the class of descriptors having a higher impact on the anti Toxoplasmosis activities of 1,3,5 triazine compounds. The quality of the model is calculated by the fitness function by taking 100 different random combinations of the calculated molecular descriptors. The fitness function of each model is formulated in terms of Q^2_{Loo} or R^2 where Q^2_{Loo} represents cross-validated R^2. The values of Q^2_{Loo} and R^2 are calculated by the standard statistical equations[17,18].

Result and Discussion

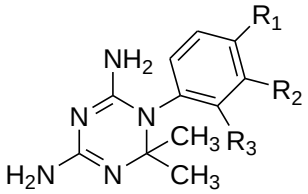
The total data set consists of 60 halogenated 1,3,5-triazine compounds. Several training and test set was randomly generated. Many training QSAR models have been developed utilizing various sets of computed molecular descriptors using GA-MLR methods of NanoBridges software. The best model is produced while the total data set was divided into training (75%) and test sets (25%) on a random basis. The best training QSAR model is given below.

PIC_{50} = -3.177(+/-1.412) + 6.957(+/-0.921) E1v + 3.662(+/-0.454) CIC2 + 0.853(+/-0.390) nHBint4 + 0.053(+/-0.011) AATSC6m - 0.174(+/-0.032) nAtomP + 0.760(+/-0.254) SHdsCH

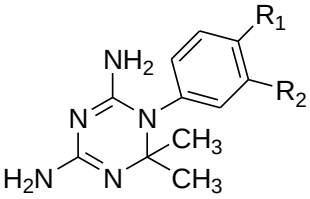
N = 49, R^2 = 0.791, $Q_{Loo}{}^2$ = 0.659, $R_{pred}{}^2$ = 0.548, PRESS = 17.928, SE=0.518 Eq. 1

In the above equation, N, R^2, $Q_{Loo}{}^2$, $R_{pred}{}^2$, PRESS, and SE represent the number of observations, squared regression coefficient, leave-one-out (LOO) cross-validated R^2, predictive R^2, the predictive sum of square deviations, and standard error of the regression model, respectively.

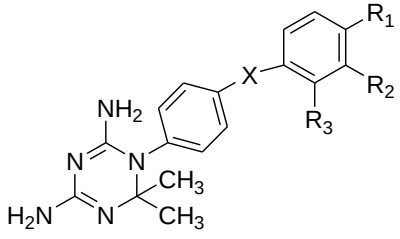
Table I — Biological activity data of 1,3,5-triazine compounds



S.No.	R_1	R_2	R_3	pIC_{50}
1*	Cl	H	H	0.107
2*	Cl	Cl	H	2.657
3	H	H	Cl	−1.927
4	H	Cl	H	0.657
5	H	Br	H	1.187
6	Cl	H	Cl	−1.748
7	F	CF_3	H	0.769
8	H	CF_3	H	0.236



S.No.	R_1	R_2	−	pIC_{50}
9	OCH_3	H	−	0.148
10	SCH_3	H	−	−0.181
11	SCH_2CH_3	H	−	0.602
12	$SO_2CH_2CH_3$	H	−	−0.949
13	OCH_2CH_3	H	−	−0.491
14*	$N(CH_2CH_2C1)_2$	H	−	0.481
15	NH_2	H	−	−0.158
16	CH_2NH_2	H	−	−0.919
17*	H	SO_2F	−	0.468
18*	H	CN	−	0.769

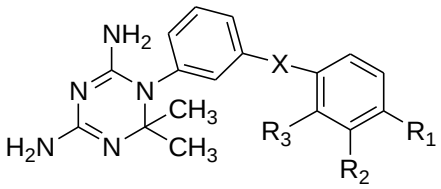


S.No.	R_1	R_2	R_3	X	pIC_{50}
19	Cl	H	Cl	$(CH_2)_4$	1.468
20	SO_2F	H	H	CH_2CH_2CONH	1.522
21*	SO_2F	H	H	CH_2CONH	0.721
22	SO_2F	CH_3	H	CH_2CH_2CONH	1.283
23	H	SO_2F	H	CONH	−1.838
24	H	SO_2F	H	CH_2CH_2CONH	1.173
25	SO_2F	H	H	$CH_2CH_2CONHSO_2$	0.823

(*Contd*.)

Table I — Biological activity data of 1,3,5-triazine compounds

26	H	SO_2F	H	OCH_2CONH	1.585
27*	NO_2	H	H	OCH_2CH_2O	1.568
28	NH_2	H	H	OCH_2CH_2O	1.494
29	$NHCOCH_2Br$	H	H	OCH_2CH_2O	1.958
30	H	H	H	$(CH_2)_4$	1.585
31	H	H	H	$(CH_2)_2$	1.920
32	H	CH_3	H	CH_2CH_2CONH	1.552
33*	H	H	H	$CH(CN)(CH_2)_3$	1.508
34	H	H	H	CO	1.602



S.No.	R_1	R_2	R_3	X	pIC_{50}
35*	Cl	H	Cl	$(CH_2)_4$	1.795
36	H	H	H	$(CH_2)_4$	1.638
37	H	$NHCOCH_2Br$	H	OCH_2	0.113
38	NH_2	H	H	$O(CH_2)_3O$	1.522
39*	$NHCOCH_2Br$	H	H	$O(CH_2)_3O$	1.251
40	H	NH_2	H	OCH_2	0.568
41	H	NH_2	H	$(CH2)_4$	1.823
42	H	NO_2	H	OCH_2	0.853
43	H	NO_2	H	$(CH=CH)_2$	1.000
44	H	NO_2	H	$O(CH_2)_2O$	1.397
45	H	$NHCOCH_2Br$	H	$O(CH_2)_2O$	0.721
46	H	$NHCOCH_2Br$	H	$O(CH_2)_3O$	0.346
47	H	NO_2	H	$O(CH_2)_3O$	1.481
48	NH_2	H	H	$O(CH_2)_2O$	1.958
49	NHCOCH=CHCOOH	H	H	$O(CH_2)_2O$	2.292
50	NHCOCH=CHCOOH	H	H	$O(CH_2)_3O$	2.376
51*	NO_2	H	H	$O(CH_2)_2O$	1.619
52	NO_2	H	H	$O(CH_2)_3O$	1.522
53	SO_2F	H	H	CH_2CONH	0.853
54	SO_2F	H	H	$(CH_2)_2CONH$	0.920
55*	SO_2F	H	H	CONH	−1.463
56	H	H	H	$(CH_2)_2$	2.000
57	H	H	H	$(CH_2)_4O$	2.337
58	SO_2F	Cl	H	$(CH_2)_4$	1.408
59	H	H	H	CH_2CONH	1.075
60	H	$CON(CH_3)_2$	H	$CH_2NHCONH$	0.966

* Test set

To validate the further predictive power of the models, QSAR model Eq. 1 is used to predict the biological activities of the corresponding test set compounds which are taken as 1, 2, 14, 17, 18, 21, 27, 33, 35, 39, 51, and 55 which were shown by an asterisk in Table I. The square correlation between observed and predicted activities has been calculated as 0.596 which stands as significant model validation (Figure 1). The modeled parameters are described in Table II.

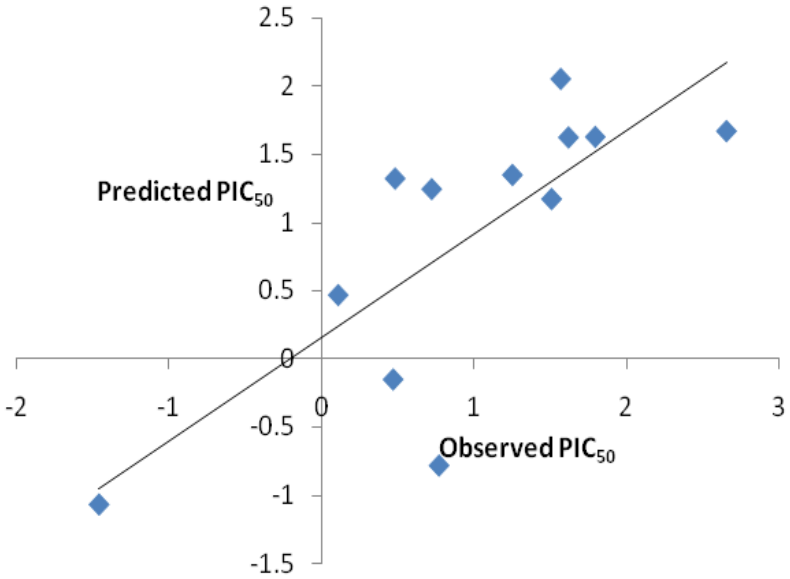


Figure 1 — Observed *vs*. predicted activity of test set compound (Eq. 1)

Table II — Physical interpretation of the modeled parameters

Name of the descriptors	Interpretation
E1v	1st component accessibility directional WHIM index / weighted by van der Waals volume
CIC2	Complementary Information Content index (neighborhood symmetry of 2-order)
nHBint4	Count of E-State descriptors of strength for potential hydrogen Bonds of path length 4
AATSC6m	Average centered Broto-Morau autocorrelation-lag 6/weight by mass
nAtomP	Number of atoms in the largest pi system
SHdsCH	Sum of atom-type hydrogen E-State: =CH-

Conclusion

The above QSAR studies predicted the crucial features such as E1v, CIC2, AATSC6m, nHBint4, and many =CH- group of the studied 1,3,5-triazine compounds are responsible for producing the inhibition of *T. gondii* DHFR. The groups having large van der Waals volume and greater molecular size encoding descriptors of E1v and CIC2 having a positive coefficient may give a positive contribution towards *T. gondii* DHFR inhibition. The descriptor AATSC6m indicates average centered Broto-Morau autocorrelation-lag 6/weight by mass and is favorable for the enzyme inhibition. The nHBint4 representing potential hydrogen bonding is favorable for DHFR inhibition. An increase in the sum of atom-type hydrogen E-State for a group of =CH- may inhibit *T. gondii* DHFR. It may impart nonpolar hydrophobicity. The decrease in molecular property such as nAtomP representing Pi-stacking interactions may increase the biological activity.

References

1 Hopper A T, Brockman A, Wise A, Gould J, Barks J, Radke J B, Sibley L D, Zou Y & Thomas S, *J Med Chem*, 62 (2019) 1562.

2 Montoya J G & Liesenfeld O, *Lancet*, 363 (2004) 1965.

3 Torgerson P R & Mastroiacovo P, *Bulletin of the World Health Organization*, 91 (2013) 501.

4 Negero J, Yohannes M, Woldemichael K & Tegegne D, *International Journal of Infectious Diseases*, 57 (2017) 44.

5 Konstantinovic N, Guegan H, Stäjner T, Belaz S & Robert-Gangneux F, *Food and Waterborne Parasitology*, 15 (2019) 00036.

6 Mattioni B & Jurs P, *J Mol Graph Model*, 21 (2003) 391.

7 Foulon W, Villena I, Stray-Pedersen B, Decoster A, Lappalainen M & Pinon J M, *Am J Obstet Gynecol*, 180 (1999) 410.

8 Araujo F G, Shepard R M & Remington J S, *Eur J Clin Microbiol Infect Dis*, 10 (1991) 519.

9 Fichera M E, Bhopale M K & Roos D S, *Antimicrob Agents Chemother*, 39 (1995) 1530.

10 Montoya J G & Remington J S, 'Toxoplasma gondii' in *Principles and Practice of Infectious Disease*, 5th edn, edited

by Mandell G L, Benntt J E & Dolin R (Churchill Livingstone, Philadelphia), pp. 2858–88 (2000).

11 Rajapakse S, Shivanthan M C, Samaranayake N, Rodrigo C & Fernando S D, *Pathog Glob Health*, 107 (2013) 162.

12 Chio L C & Queener S F, *Antimicrob Agents Chemother*, 37 (1993) 1914.

13 Mills N, *J Am Chem Soc*, 128 (2006) 13649.

14 Yap C W, *Journal of Computational Chemistry*, 32(7) (2011) 1466.

15 Ballabio D, Conssoni V, Mauri A, Claeys-Bruno M, Sergant M, Todeschini R, *Chemometrics and Intelligent Laboratory Systems*, 136 (2014) 147.

16 Ambure A P, Gajewicz R B, Puzyn A T & Roy K, *Chemometr Intell Lab Syst*, 147 (2015) 1.

17 Broadhurst D, Goodacre R, Jones A, Rowland & Kell D B, *Anal Chim Acta*, 348 (1997) 71.

18 Hoffman B T, Kopajtic T, Katz J L & Newman A H, *J Med Chem*, 43 (2000) 4151.