



ISSN 0975-413X  
CODEN (USA): PCHHAX

Der Pharma Chemica, 2016, 8(3):180-188  
(<http://derpharmachemica.com/archive.html>)

## Exploring the cytotoxicity of 1,3,5-triazines and triazine analogs against lung cancer by QSAR study using genetic function approximation

Marwa Fathy Balaha<sup>1\*</sup>, Mervat Hamed El-Hamamsy<sup>2</sup>  
and Nabawya Abd-El-Salam Sharaf El-Din<sup>3</sup>

<sup>1</sup>Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Kafrelsheikh University, Elgesh street, Kafrelsheikh, Egypt

<sup>2</sup>Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tanta University, Elgesh street, Tanta, Egypt

<sup>3</sup>Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tanta University, Elgesh street, Tanta, Egypt

### ABSTRACT

Recently several triazine derivatives were identified as excellent cytotoxic agents against non-small cell lung carcinoma by our group. QSAR model for prediction of biological activity of triazine derivatives against non-small cell lung carcinoma cell line (A549) is needed to construct selective inhibitors for lung cancer. Twenty four models were constructed using genetic function approximation algorithm. The best of these models was chosen based on its statistical validation parameters where the  $R^2$  value was found to be acceptable (0.98). The developed model was based on four molecular descriptors; two fast descriptors and two VAMP electrostatics descriptors. External validation of the model was governed by calculating the residual values for test set. Further external validation is investigated by calculating the biological activity of four new triazine derivatives synthesized by our group in a previous contribution from our laboratory. Our developed model is proved to have high predictive and diagnostic abilities and can distinguish different stereoisomers. The accuracy of the 3D structures used affects the model quality.

**Keywords:** antitumor, A549, genetic function approximation, non-small cell lung carcinoma, QSAR, triazine.

### INTRODUCTION

A major breakthrough in the field of quantitative structure activity relationship (QSAR) of triazines as anticancer agents was achieved by Hansch in 1975 [1]. He developed a QSAR model correlating the biological activity of a set of 256 compounds of triazines synthesized by Baker's group [2] to their chemical structures. Afterward several QSAR studies for triazines were developed [3-10] owing to the importance of this class of compounds. Triazines have a wide range of biological activities [11] including, anti-microbial [12], antifungal [13], antimalarial activity [14], antiviral activity [15] and cytotoxic activity [16-18]. Several studies based on the triazine scaffold toward antitumor activity have been carried out [19, 20] starting by Baker who studied active site-directed, irreversible inhibition of dihydrofolate reductase (DHFR) enzyme [2]. Hexamethylmelamine (HMM) is 1,3,5-triazine derivative and is used clinically as antitumor agent against lung, ovarian and breast cancers [21]. Hydroxymethyl pentamethylmelamine (HMPMM) is the hydroxylated metabolite and is the major active form of HMM [21].

Dihydrofolate reductase (DHFR) enzyme is responsible for synthesis of tetrahydrofolate (THF) which is a cofactor necessary for DNA synthesis. Repressed expression of DHFR induces cell cycle arrest in human cell lung cancer (A549) [22]. Methotrexate (MTX) is reported to be clinically useful DHFR inhibitor [23] and is frequently used in the treatment of cancer [24]. The antitumor activity represented as  $IC_{50}$  of MTX was determined in 6 different cancer cell lines and was in an extensively broad range from 6.05 nM to more than 1,000 nM. The osteosarcoma (Saos-2) ( $IC_{50} > 1,000$  nM) and breast cancer (MCF-7) ( $IC_{50} = 114.31$  nM) cells were the most resistant to MTX. In contrast, the

gastric cancer (AGS) and colon cancer (HCT-116) cells were highly sensitive to MTX with  $IC_{50}$  of 6.05nM and 13.56nM respectively. The two non-small cell lung cancer cell lines, (NCI-H23) and (A549) cells demonstrated similar sensitivity to MTX with  $IC_{50} = 38.25nM$  and 38.33 nM respectively[25].

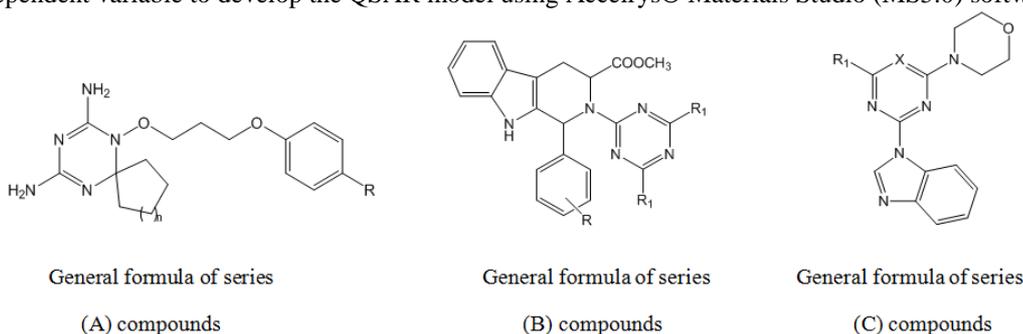
According to the world health organization (WHO) there are more than 100 types of cancers and any part of the body can be affected. Lung cancer is one of the most causes of the majority of cancer deaths all over the world. Worldwide, the five most common types of cancer in order of frequency that kill men are lung, stomach, liver, colorectal and esophagus. For women they are breast, lung, stomach, colorectal and cervical [26].

Recently, synthesis of new triazine derivatives with remarkable antitumor activity against non-small cell lung cancer was achieved in our laboratory[27].Consequently, we decided to explore the relationship between the chemical structure and cytotoxic activity for triazines and triazine analogs against lung cancer by constructing a new QSAR model to provide useful information on the structural requirements for anticancer activity against non-small cell lung carcinoma (A549) which could lead to potent drug candidates as well as better prediction of biological activity of novel non classical anticancer compounds.

## MATERIALS AND METHODS

### Biological activity data:

Our data set comprised of 43 compounds. The chemical structures of 1,3,5-triazines and triazine analogs are shown in Figure 1 while their anticancer activities expressed as  $IC_{50}$  against non-small cell lung carcinoma cell line (A549) are listed in Table 1 [19, 20, 28].The training set comprised of 32 compounds including methotrexate as a reference compound because binding of triazine derivatives to DHFR mimic that of MTX [29]. The remaining 11 compounds constituted the testing set including HMPMM. This selection considered the fact that the test molecules must represent a range of biological activities similar to that of the training set [30]. The logarithm of  $IC_{50}$  ( $\log_1/IC_{50}$ ) was used as dependent variable to develop the QSAR model using Accelrys® Materials Studio (MS5.0) software [31].



**Figure 1. Chemical structures of 1,3,5-triazines and triazine analogs that were extracted from three different literature sources after comprehensive and careful survey of literature to develop a statistically significant QSAR model**

**Table 1. Chemical and biological activity data of the training and test sets of series A, B and C compounds against non-small cell lung carcinoma cell line (A549) [19, 20, 28]**

Chemical and biological activity data of series (A) compounds [19]				
Entry	R	n	$IC_{50}$ ( $\mu M$ )	
1.	H	1	0.0402	
2.	F	1	0.0588	
3.	<sup>a</sup> Cl	1	0.0271	
4.	NO <sub>2</sub>	1	0.0658	
5.	Me	1	0.0481	
6.	t-Bu	1	0.1599	
7.	MeO	1	0.0591	
8.	<sup>a</sup> CN	1	0.0607	
9.	CH <sub>3</sub> CO	1	0.0598	
10.	H	2	0.0697	
11.	F	2	0.0592	
12.	Cl	2	0.0496	
13.	<sup>a</sup> NO <sub>2</sub>	2	0.3293	
14.	<sup>a</sup> Me	2	0.0833	
15.	t-Bu	2	0.4956	
16.	MeO	2	0.0516	
17.	CN	2	0.1164	
18.	<sup>a</sup> CH <sub>3</sub> CO	2	0.1448	
19.	SO <sub>2</sub> NH <sub>2</sub>	2	0.1664	
20.	Methotrexate (MTX)(reference compound)		0.0374	

Chemical and biological activity data of series (B) compounds [20]				
Entry	R	R <sub>1</sub>	isomer	IC <sub>50</sub> (μM)
21.	4-Methoxy	N-methylpiperazinyl	dl, cis	6.30
22.	<sup>a</sup> 4-Methoxy	N-methylpiperazinyl	dl, trans	5.83
23.	3,4,5-Trimethoxy	N-methylpiperazinyl	dl, trans	5.67
24.	3,4,5-Trimethoxy	N-methylpiperazinyl	dl, cis	9.84
25.	<sup>a</sup> 4-Methoxy	N-methylpiperazinyl	l, cis	3.13
26.	4-Methoxy	N-methylpiperazinyl	l, trans	2.1
27.	4-Isopropyl	N-methylpiperazinyl	l, cis	1.75
28.	4-Isopropyl	N-methylpiperazinyl	l, trans	1.76
29.	<sup>a</sup> 4-Methoxy	N-ethylpiperazinyl	l, trans	1.61
30.	<sup>a</sup> 4-Methoxy	N-propylpiperazinyl	l, trans	1.67
31.	4-Methoxy	N-methylpiperazinyl	d, cis	11.46
32.	4-Methoxy	N-methylpiperazinyl	d, trans	3.00
33.	-H	N-methylpiperazinyl	d, cis	8.62
34.	-H	N-methylpiperazinyl	d, trans	3.21
35.	<sup>a</sup> 4-Methyl	N-methylpiperazinyl	d, trans	2.34
36.	4-Chloro	N-methylpiperazinyl	dl, trans	6.55
Chemical and biological activity data of series (C) compounds[28]				
Entry	R <sub>1</sub>	X		IC <sub>50</sub> (μM)
37.	4-methylmorpholino	N		7.0
38.	2,2,4-trimethylmorpholino	N		6.5
39.	trans 2,3,4-trimethylmorpholino	N		6.8
40.	cis 2,3,4-trimethylmorpholino	N		4.1
41.	trans 2,3,4-trimethylmorpholino	C		5.4
42.	cis 2,3,4-trimethylmorpholino	C		3.1
43.	<sup>a</sup> Hydroxymethylpentamethylmelamine (HMPMM)			83

<sup>a</sup> Test compounds

### Geometry optimization:

3D Structures were drawn and geometry optimized using the ChemAxon® MarvinSketch 5.1.4 [32]. These structures were further geometry optimized using the Vienna ab-initio Molecular dynamics Package (VAMP) module [33]. Different algorithms were used to give the best output structure energy.

### Alignment of molecules:

The consensus flexible alignment was done in relation to X axis using the root mean square (RMS) with field fit method by employing a combination of steric and electrostatic field.

### Building a QSAR model:

Genetic function approximation (GFA)[34], a statistical modeling algorithm, was used to build the model using the most simplest fast descriptors which either one dimensional (1D) or two dimensional (2D) and the most complex three dimensional (3D) atomistic descriptors, VAMP electrostatics, spatial descriptors and for cite energetics [35]. GFA was employed to search for the best possible QSAR regression equation capable of correlating the variations in biological activities of the training compounds with variations in the generated descriptors, i.e., multiple linear regression modeling (MLR). GFA method was used to carry out both data reduction and parametric regression simultaneously. The equation length was set to make number of variables do not exceed one third to one fifth the number of data points. GFA parameters include population number and scoring function. Population was set to 500, maximum generations were set to 10000, number of top equations returned was set to 8 and constant equation length of 5, scoring function Friedman LOF, scaled LOF smoothness parameter of 0.5, mutation probability of 0.1 and using both linear, quadratic and spline functions to obtain the best equation.

### Validating the model:

All constructed models were validated using the reported validation parameters[36-39]. These parameters include Friedman lack of fit (LOF), the squared correlation coefficient ( $R^2$ ), adjusted  $R^2$ , cross validated  $R^2$  (CV) and significance of regression (SOR) F-values of the training set in addition to  $R^2$  value of the test set. Scaled LOF smoothness parameter was set to default of 0.5.

## RESULTS AND DISCUSSION

Development of statistically significant QSAR model depends mainly on careful selection of data set used in building the model [40, 41]. Consequently the most reliable and representative data set was extracted from three different literature sources [19, 20, 28] after comprehensive and careful survey of literature.

Some descriptors, such as the dipole moment components, depend on having all the molecules in the same alignment. The alignment of compounds involved in this study is achieved using flexible alignment.

GFA was selected to construct the QSAR models because this approach has a number of important advantages over other algorithms that utilize multiple linear regression (MLR), partial least squares (PLS) and neural network analysis [38]. These advantages include; building a multiple models rather than a single model, automatically selecting which features are to be used in the model (the most important step in QSAR studies), better at discovering combinations of features that take advantage of correlations between multiple features (incorporates Friedman's LOF error measure which estimates the most appropriate number of features), resists over fitting, and allows control over the smoothness of fit [38, 41, 42]. Moreover it can use a larger variety of equation term types in construction of its models (for example, splines, step functions, or high order polynomials). Finally, it provides, through study of the evolving models, additional information that is not available from standard regression analysis, such as the preferred model length and useful partitions of the data set [41].

Twenty four models each contain eight equations were constructed. No good model can be built using basic element fast descriptors only because they are either one dimensional (1D) or two dimensional (2D) and they are not geometry-dependent. The accuracy of the 3D structures used (i.e., the bond angles, etc.) will affect the model quality. Inclusion of 3D shape descriptors in the model lead to improved descriptions relating computed parameters to biological activity which revealed that stereochemical parameters have a remarkable effect on biological activity. The best generated model was based on four molecular descriptors, two fast descriptors and two VAMP Electrostatics descriptors. Equations and involved molecular descriptors and their physicochemical meaning are given in Table 2.

**Table 2. The developed QSAR model, equations and molecular descriptors**

	Equation		Equation
(1)	$Y = -31.658446473 * X37 + 0.003324167 * X74 + 0.108327040 * \text{ramp}(X13 - 30.996253036) - 0.866099800 * (\text{ramp}(X18 - 1.578054048))^2 + 740.952842524 * (\text{ramp}(X18 - 3.193861295))^2 + 2.638204064$	(5)	$Y = -31.718039989 * X37 + 0.003330155 * X74 + 0.108402634 * \text{ramp}(X13 - 30.919103344) - 0.759090066 * (\text{ramp}(X18 - 1.473600396))^2 + 765.668152326 * (\text{ramp}(X18 - 3.195387969))^2 + 2.638472161$
(2)	$Y = -31.658446473 * X37 + 0.003324167 * X74 + 0.108327040 * \text{ramp}(X13 - 30.996253036) + 740.952842524 * (\text{ramp}(X18 - 3.193861295))^2 - 0.866099800 * (\text{ramp}(X18 - 1.578054048))^2 + 2.638204064$	(6)	$Y = -31.718039989 * X37 + 0.003330155 * X74 + 0.108402634 * \text{ramp}(X13 - 30.919103344) + 765.668152326 * (\text{ramp}(X18 - 3.195387969))^2 - 0.759090066 * (\text{ramp}(X18 - 1.473600396))^2 + 2.638472161$
(3)	$Y = -31.658468036 * X37 + 0.108326746 * \text{ramp}(X13 - 30.996253036) + 0.003324232 * \text{ramp}(X74 + 197.695990383) + 740.958760509 * (\text{ramp}(X18 - 3.193861295))^2 - 0.866105350 * (\text{ramp}(X18 - 1.578054048))^2 + 1.981014867$	(7)	$Y = -31.674135306 * X37 + 0.108386588 * \text{ramp}(X13 - 30.996253036) + 0.003326166 * \text{ramp}(X74 + 197.695990383) - 0.758573593 * (\text{ramp}(X18 - 1.473600396))^2 + 765.258709111 * (\text{ramp}(X18 - 3.195387969))^2 + 2.001029408$
(4)	$Y = -31.658468036 * X37 + 0.108326746 * \text{ramp}(X13 - 30.996253036) + 0.003324232 * \text{ramp}(X74 + 197.695990383) - 0.866105350 * (\text{ramp}(X18 - 1.578054048))^2 + 740.958760509 * (\text{ramp}(X18 - 3.193861295))^2 + 1.981014867$	(8)	$Y = -31.674135306 * X37 + 0.108386588 * \text{ramp}(X13 - 30.996253036) + 0.003326166 * \text{ramp}(X74 + 197.695990383) + 765.258709111 * (\text{ramp}(X18 - 3.195387969))^2 - 0.758573593 * (\text{ramp}(X18 - 1.473600396))^2 + 2.001029408$

Where,  $Y$ :  $\log(1/IC_{50})$

$X37$ :  $N2(3)$ : Mulliken charge (VAMP Electrostatics)

$X13$ : Subgraph counts (1): path (Fast Descriptors)

$X74$ : Octupolexxz (VAMP Electrostatics)

$X18$ : Chi (3): cluster (Fast Descriptors)

Fast descriptors are Sub graph counts (path) and Chi (cluster). They are topological indices which based on graph theory concepts [43]. They help to differentiate molecules according to their size, degree of branching, flexibility, and overall shape. topological indices were first used in QSAR studies of triazines as DHFR inhibitors in 2006 [9].

VAMP Electrostatics descriptors are Mulliken charge [44] and Octupolexxz electrostatic moment components [39]. The VAMP module [45] allows to predict geometries, heats of formation, and a host of molecular properties, including ionization potential, multipole moments, molecular and atomic polarizabilities, and potential-derived charges.

Mulliken charges arise from the Mulliken population analysis [44]. They provide a means of estimating partial atomic charges calculated by the methods of computational chemistry, particularly those based on the linear combination of atomic orbitals molecular orbital method.

The developed model prove that the biological activity of these series of compounds controlled mainly by the molecular size, shape and charge which is clearly represented by descriptors in the constructed equations.

A comparatively reported 3D-QSAR study on dihydro-1,3,5-triazines and their spiro derivatives (series A) as DHFR inhibitors by comparative molecular field analysis (CoMFA) revealed that the biological activity is highly dependent on the molecular size, shape and molecular charge of the ligand [10].

The best of these constructed equations was chosen based on its statistical validation parameters. The internal validation parameters calculated for the model represented by equations 1-8 are shown in Table 3 where the R<sup>2</sup> value was found to be acceptable (0.98).

**Table 3. The internal validation parameters calculated for the developed QSAR model represented by equations 1-8**

<b>Internal validation parameters</b>	<b>Equation (1)</b>	<b>Equation (2)</b>	<b>Equation (3)</b>	<b>Equation (4)</b>
Friedman LOF	0.36815100	0.36815100	0.36815400	0.36815400
R-squared	0.98603600	0.98603600	0.98603600	0.98603600
Adjusted R-squared	0.98312700	0.98312700	0.98312700	0.98312700
Cross validated R-squared	0.97859700	0.97859700	0.97859700	0.97859700
Significant Regression	Yes	Yes	Yes	Yes
<b>Internal validation parameters</b>	<b>Equation (5)</b>	<b>Equation (6)</b>	<b>Equation (7)</b>	<b>Equation (8)</b>
Friedman LOF	0.36993300	0.36993300	0.36995400	0.36995400
R-squared	0.98596800	0.98596800	0.98596800	0.98596800
Adjusted R-squared	0.98304500	0.98304500	0.98304400	0.98304400
Cross validated R-squared	0.97843900	0.97843900	0.97843700	0.97843700
Significant Regression	Yes	Yes	Yes	Yes

The internal validation [39] results prove that the developed QSAR model represented by equations 1-8 is accepted in terms of good correlation coefficient and low LOF value.

External validation [39] of the developed model is achieved by calculating the predicted biological activity for test set using equation (1). These calculated values revealed a good prediction ability of our developed model as shown in Figure 2A and 2B. The residual values are calculated from the difference between the actual and predicted biological activity values ( $\log 1/IC_{50}$ ) for test and training sets as shown in Table 4. It worth noting that compound 43(HMPMM) showed unexpected large residual value deviating from experimental value by (-4.16264042). Our result of the unexpected high predicted biological activity of HMPMM than practical activity is consistent with and supported by reported studies which was explained by its inherent chemical instability problem owing to chemical loss of formaldehyde to give pentamethylmelamine[21, 46, 47] as shown in Figure 3. General instability of the hydroxymethyl species in melamine derivatives makes it impossible to prepare stable hydroxymethylmelamines while having an amino proton (N-H) in the molecule. Thus the oxidative metabolite of PMM, N<sup>2</sup>-hydroxymethyl-N<sup>2</sup>,N<sup>4</sup>,N<sup>6</sup>-tetramethylmelamine is not obtained synthetically [46].

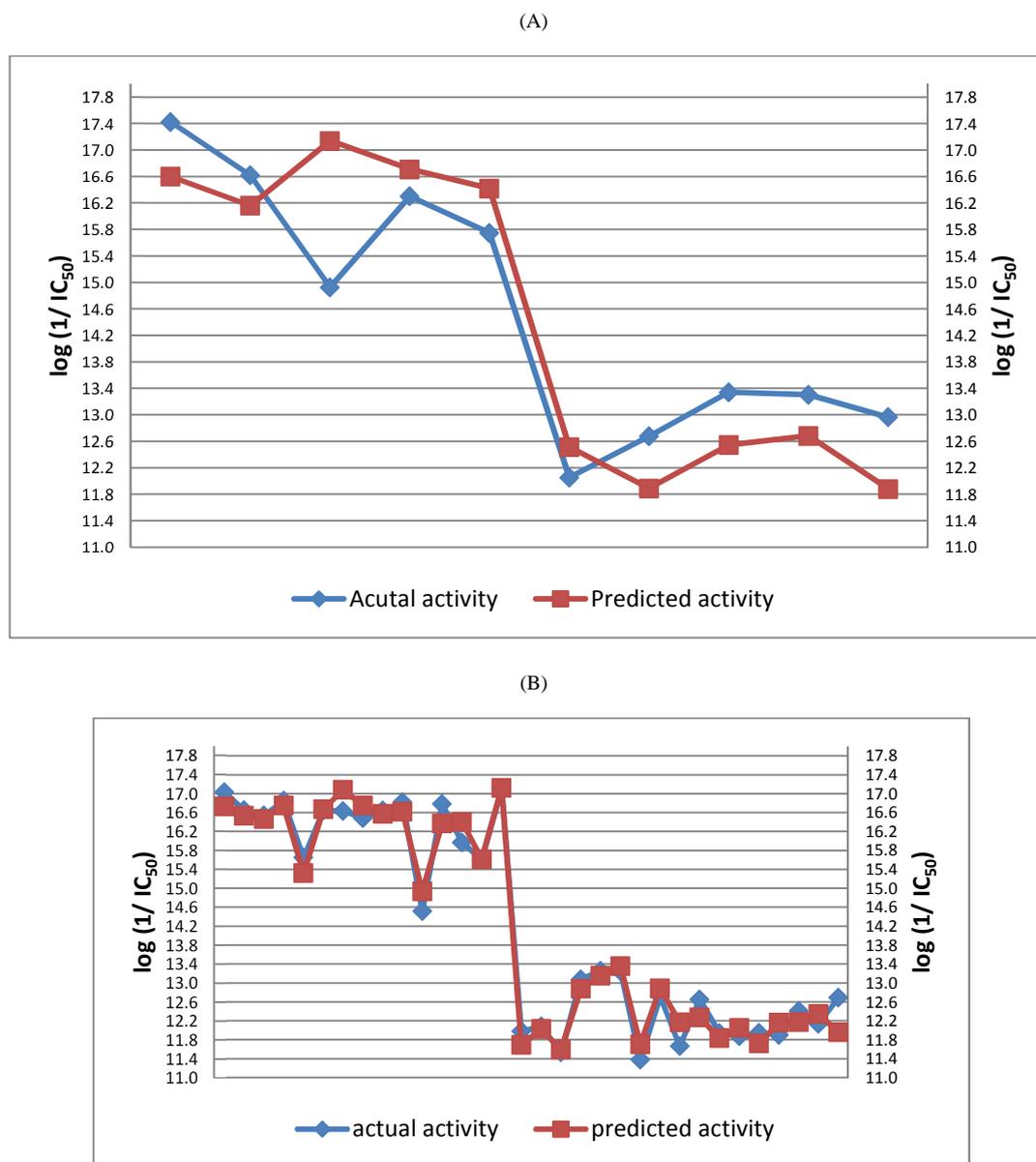


Figure 2. External validation of the developed model, the plot of predicted  $\log(1/IC_{50})$  versus experimental values for test set (A), training set (B). This graph revealed a good prediction ability of our model

Table 4. External validation of the developed model by calculating the residual values for test and training sets using equation (1)

Entry (series)	log (1/IC <sub>50</sub> ) (M) (practical)	log (1/IC <sub>50</sub> ) (M) (predicted)	residual value	Entry (series)	log (1/IC <sub>50</sub> ) (M) (practical)	log (1/IC <sub>50</sub> ) (M) (predicted)	residual value
1. (A)	17.02939884	16.72779417	0.30160467	27. (B)	13.25589477	13.15071598	0.10517879
2. (A)	16.64912398	16.53264790	0.11647608	28. (B)	13.25019675	13.35406075	-0.10386400
3. <sup>a</sup> (A)	17.42373211	16.59879797	0.82493414	29. <sup>a</sup> (B)	13.33927638	12.54397594	0.79530044
4. (A)	16.53664600	16.46134782	0.07529818	30. <sup>a</sup> (B)	13.30268693	12.68359786	0.61908907
5. (A)	16.84998366	16.74708873	0.10289493	31. (B)	11.37664785	11.69883513	-0.32218728
6. (A)	15.64871722	15.31802345	0.33069376	32. (B)	12.71689827	12.88229026	-0.16539199
7. (A)	16.64403491	16.66705088	-0.02301597	33. (B)	11.66142547	12.16429635	-0.50287087
8. <sup>a</sup> (A)	16.61732214	16.16126624	0.45605590	34. (B)	12.64923962	12.26945682	0.37978280
9. (A)	16.53664600	16.46134782	-0.45028696	35. <sup>a</sup> (B)	12.96535963	11.87891508	1.08644455
10. (A)	16.84998366	16.74708873	-0.26687514	36. (B)	11.93604551	11.82890377	0.10714174
11. (A)	15.64871722	15.31802345	0.06902815	37. (B)	-----	-----	-----
12. (A)	16.64403491	16.66705088	0.20345567	38. (B)	-----	-----	-----
13. <sup>a</sup> (A)	14.92629665	17.13777789	-2.21148124	39. (B)	-----	-----	-----
14. <sup>a</sup> (A)	16.30081729	16.70975997	-0.40894268	40. (C)	11.86960041	12.04526966	-0.17566925
15. (A)	14.51749669	14.92976091	-0.41226422	41. (C)	-----	-----	-----
16. (A)	16.77974416	16.37143598	0.40830819	42. (C)	-----	-----	-----
17. (A)	15.96623330	16.40314421	-0.43691091	43. (C)	11.94370838	11.71754993	0.22615845
18. <sup>a</sup> (A)	15.74791236	16.42043906	-0.67252670	44. (C)	-----	-----	-----
19. (A)	15.60887131	15.61294445	-0.00407314	45. (C)	11.89858795	12.15686692	-0.25827897
20. (A)	17.10159513	17.11671607	-0.01512094	46. (C)	12.40452358	12.17359041	0.23093317
21. (B)	11.97496092	11.68691513	0.28804580	47. (C)	-----	-----	-----
22. <sup>a</sup> (B)	12.05249356	12.51538471	-0.46289115	48. (C)	-----	-----	-----
23. (B)	12.08032144	12.02717877	0.05314267	49. (C)	12.12911160	12.33846176	-0.20935016
24. (B)	11.52905485	11.58977800	-0.06072315	50. (C)	12.68410845	11.95702913	0.72707932
25. <sup>a</sup> (B)	12.67447755	11.88911734	0.78536022	51. <sup>a</sup> (C)	9.39666995	13.55931037	-4.16264042
26. (B)	13.07357321	12.87418347	0.19938975				

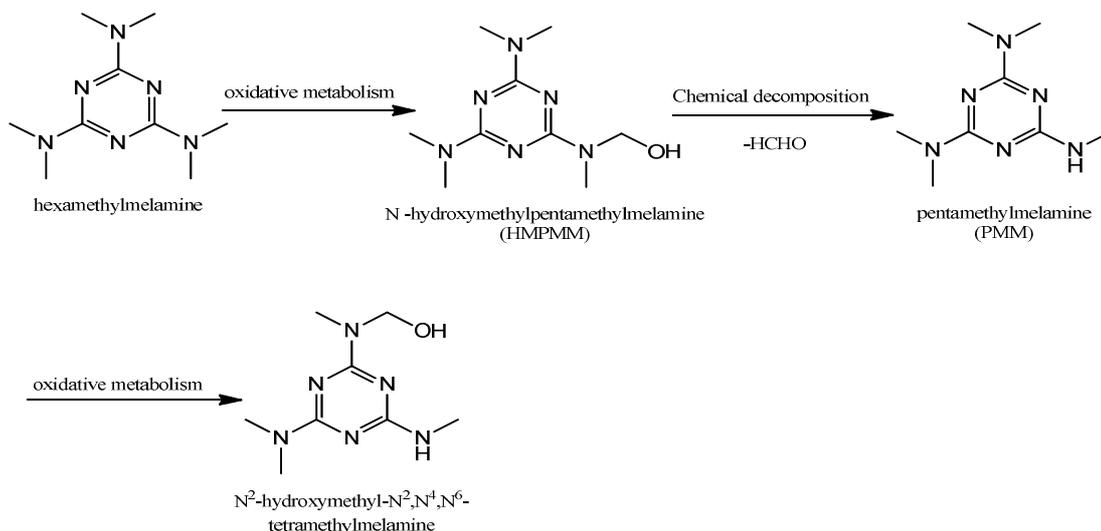
<sup>a</sup> Test compounds

Figure 3. Metabolism of hexamethylmelamine (HMM). HMM is 1,3,5-triazine derivative and is used clinically as antitumor agent. Hydroxymethylpentamethylmelamine (HMPMM) is the hydroxylated metabolite and is the major active form of HMM showing inherent chemical instability problem owing to chemical loss of formaldehyde to give pentamethylmelamine

Further external validation is investigated by calculating the biological activity of four new triazine derivatives synthesized by our group (Figure 4) using the developed model. The developed model revealed very good predictability as shown in Figure 5 and Table 5 and residual values ranged from 0.3574514156 to 0.9234778675.

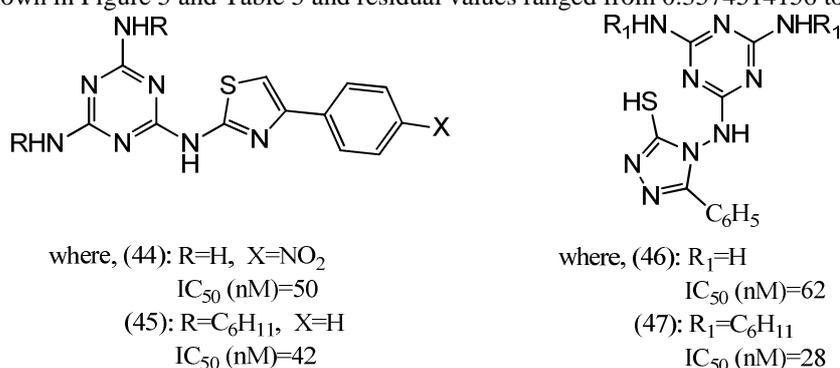
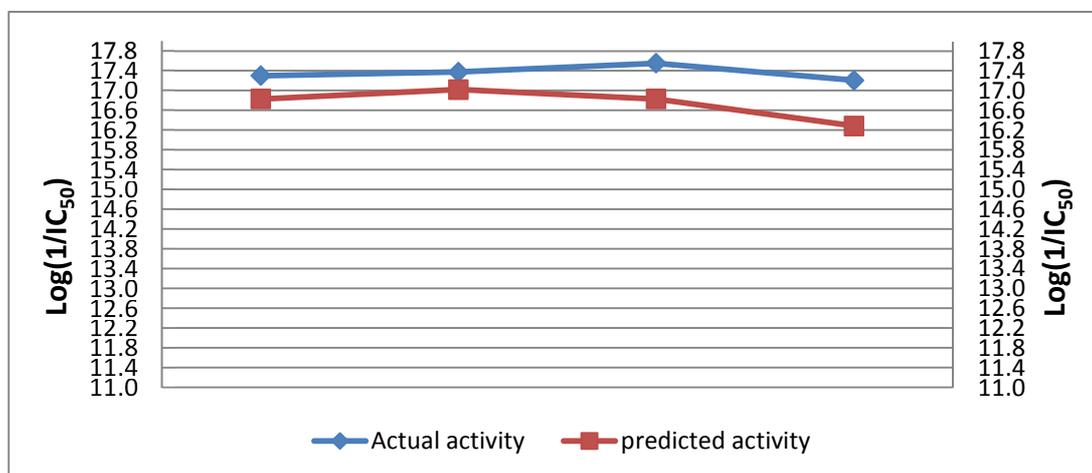


Figure 4. Chemical structures and IC<sub>50</sub> of new 1,3,5-triazine derivatives synthesized in our laboratory

Table 5. External validation of the developed model by calculating predicted  $\log(1/IC_{50})$  and residual values for the newly synthesized four triazine derivatives

Compound number	$\log(1/IC_{50})$ (M) (practical)	$\log(1/IC_{50})$ (M) (predicted)	residual value
44	17.30102999	16.83069918	0.47033081
45	17.37675071	17.0192992944	0.3574514156
46	17.20760831	16.2841304425	0.9234778675
47	17.55284196	16.8303105576	0.7225314024

Figure 5. Plot of predicted  $\log(1/IC_{50})$  versus experimental values for the newly synthesized four triazine derivatives

Quality of our model not limited to predict the scientific fact of HMPMM chemical decomposition or predict the biological activity of new triazine derivatives, but also it can distinguish different stereoisomers as in series B. It can predict different biological activity for compounds (21, 22, 25, 26, 31, and 32), (23 and 24), (27 and 28) and (35 and 37) although they have the same molecular formula and sequence of bonded atoms (constitution), but that differ only in the three-dimensional orientations of their atoms in space. In addition, compounds 49 and 50 in series C were pyrimidine derivatives and the model showed a good prediction.

## CONCLUSION

A new QSAR model with good predictive and diagnostic abilities is developed using GFA approach. It provides a useful guideline in future design of highly promising selective non-small cell lung cancer cytotoxic agents. The developed model can be used to predict the biological activity ( $IC_{50}$ ) of different classes (e.g., inhibitor versus non inhibitors) of compounds before the actual biological testing against A549 cell line. It can also be used in the analysis of physicochemical structural characteristics that can give rise to inhibitors of non-small lung cancer (A549) tumor cell proliferation. Being inhibitor of proliferation of non-small lung tumor cell (A549) require protonated partially charged substituted nitrogen atom on triazine ring to form hydrogen bond and ionic interaction similar to those made by the 2,4-diaminopteridine ring of MTX. Bulky substituent at triazine ring is preferred for hydrophobic interactions similar to benzoyl moiety of MTX.

## REFERENCES

- [1] C. Silipo and C. Hansch, *J. Am. Chem. Soc.*, **1975**, 97, 6849.
- [2] B. R. Baker and G. J. Lourens, *J. Med. Chem.*, **1967**, 10, 1113.
- [3] B. A. Hathaway, Z. R. Guo, C. Hansch, T. J. Delcamp, S. S. Susten and J. H. Freisheim, *J. Med. Chem.*, **1984**, 27, 144.
- [4] R. B. Srivastava, P. P. Singh and S. Singh, *Med. Chem.A.I.J.*, **2005**, 1, 14.
- [5] X. Fradera, L. Amat, E. Besalü and R. Carbó-Dorca, *Quant. Struct.Act.Relat.*, **1997**, 16, 25.
- [6] A. J. Hopfinger, *J. Am. Chem. Soc.*, **1980**, 102, 7196.
- [7] G. Lakshmi, T. Mithilesh and S. K. Singh, *Int J Sci Res*, **2013**, 2, 41.
- [8] M. Tiwari, S. K. Singh and L. Gangwar, *Int J Sci Res*, **2013**, 2, 96.
- [9] T. Abhilash, T. Mamta and T. Suprajnya, *Asian J. Biochem.*, **2006**, 1, 138.
- [10] X. Ma, G. Xiang, C.-W. Yap and W.-K. Chui, *Bioorg. Med. Chem. Lett.*, **2012**, 22, 3194.

- [11] R. V. Patel, P. Kumari, D. P. Rajani, C. Pannecouque, E. De Clercq and K. H. Chikhaliya, *Future med. chem.*, **2012**, 4, 1053.
- [12] X. Ma, S.-T. Tan, C.-L. Khoo, H.-M. Sim, L.-W. Chan and W.-K. Chui, *Bioorg. Med. Chem. Lett.*, **2011**, 21, 5428.
- [13] K. Sarmah, N. Sarmah, K. Kurmi and T. Patel, *Adv. Appl. Sci. Res.*, **2012**, 3, 1459.
- [14] S. Manohar, S. I. Khan and D. S. Rawat, *Chem Biol Drug Des.*, **2013**, 81, 625.
- [15] N. Mibu, K. Yokomizo, A. Koga, M. Honda, K. Mizokami, H. Fujii, N. Ota, A. Yuzuriha, K. Ishimaru and J. Zhou, *Chem. Pharm. Bull.*, **2014**, 62, 1032.
- [16] M. El-Hamamsy and N. El-Mahdy, *Life Sci J.*, **2014**, 11, 382.
- [17] W. Zhu, Y. Liu, Y. Zhao, H. Wang, L. Tan, W. Fan and P. Gong, *Arch. Pharm.*, **2012**, 345, 812.
- [18] G. J. Kumar, H. S. Bomma, E. Srihari, S. Shrivastava, V. Naidu, K. Srinivas and V. J. Rao, *Med. Chem. Res.*, **2013**, 22, 5973.
- [19] X. Ma and W.-K. Chui, *Bioorg. Med. Chem.*, **2010**, 18, 737.
- [20] R. Kumar, L. Gupta, P. Pal, S. Khan, N. Singh, S. B. Katiyar, S. Meena, J. Sarkar, S. Sinha, J. K. Kanaujiya, S. Lochab, A. K. Trivedi and P. M. S. Chauhan, *Eur. J. Med. Chem.*, **2010**, 45, 2265.
- [21] R. Kumar, A. Deep Singh, J. Singh, H. Singh, R. Roy and A. Chaudhary, *Mini Rev Med Chem.*, **2014**, 14, 72.
- [22] V. Llado, S. Teres, M. Higuera, R. Alvarez, M. A. Noguera-Salva, J. E. Halver, P. V. Escriba and X. Busquets, *Proc Natl Acad Sci U S A*, **2009**, 106, 13754.
- [23] J. Neradil, G. Pavlasova, M. Sramek, M. Kyr, R. Veselska and J. Sterba, *Oncol. Rep.*, **2015**, 33, 2169.
- [24] S. S. Abolmaali, A. M. Tamaddon and R. Dinarvand, *Cancer Chemother Pharmacol.*, **2013**, 71, 1115.
- [25] S. Yoon, J. R. Choi, J.-O. Kim, J.-Y. Shin, X. Zhang and J.-H. Kang, *Cancer Treat Res.*, **2010**, 42, 163.
- [26] WHO Fact sheet on cancer. accessed on 30 March 2015, URL: <http://www.who.int/features/factfiles/cancer/en/>.
- [27] M. balaha, M. El-Hamamsy, N. Sharaf El-Din, N. El-Mahdy, *Journal of Applied Pharmaceutical Science*, **2016**, in press.
- [28] T. Matsuno, M. Kato, H. Sasahara, T. Watanabe, M. Inaba, M. Takahashi, S.-i. Yaguchi, K. Yoshioka, M. Sakato and S. Kawashima, *Chem. Pharm. Bull.*, **2000**, 48, 1778.
- [29] R. L. Blakley, *Adv Enzymol Relat Areas Mol Biol*, **1995**, 70, 23.
- [30] T. M. Martin, P. Harten, D. M. Young, E. N. Muratov, A. Golbraikh, H. Zhu and A. Tropsha, *J. Chem. Inf. Model.*, **2012**, 52, 2570.
- [31] Accelrys® Materials Studio. [Accessed: Aug/4<sup>th</sup>/2014] URL: <http://accelrys.com/products/materials-studio/>.
- [32] ChemAxon® MarvinSketch. [Accessed: Aug/4<sup>th</sup>/2014] URL: <http://www.chemaxon.com/products/marvin/>.
- [33] Vienna ab-initio Simulation Package (VASP) Group. [Accessed: Aug/4<sup>th</sup>/2014] URL: <http://www.vasp.at/>.
- [34] M. Chhabria, M. Jani, K. Parmar and M. Singh, *Med. Chem. Res.*, **2012**, 21, 407.
- [35] Accelrys®. Forcite. [Accessed: Aug/4<sup>th</sup>/2014] URL: [accelrys.com/products/datasheets/ms\\_forcite.pdf](http://accelrys.com/products/datasheets/ms_forcite.pdf).
- [36] C. Nantasenamat, C. Isarankura-Na-Ayudhya, T. Naenna and V. Prachayasittikul, *EXCLI Journal*, **2009**, 8, 74.
- [37] P. Gramatica, *Mol. Inf.*, **2014**, 33, 311.
- [38] K. Khaled and N. Abdel-Shafi, *Int. J. Electrochem. Sci*, **2011**, 6, 4077.
- [39] O. J. Reyes, S. J. Patel and M. S. Mannan, *Ind. Eng. Chem. Res.*, **2011**, 50, 2373.
- [40] L. C. Yee and Y. C. Wei, In: M. Dehmer, K. Varmuza and D. Bonchev (Ed.), *Statistical Modelling of Molecular Descriptors in QSAR/QSPR* (Wiley-Blackwell, **2012**), 1.
- [41] S. A. Mirkhani, F. Gharagheizi and M. Sattari, *Chemosphere*, **2012**, 86, 959.
- [42] T. F. El-Moselhy, *Chem Biol.*, **2012**, 2, 172.
- [43] A. Q. Baig, M. Imran and H. Ali, *Can J Chem.*, **2015**, 93, 730.
- [44] T. Yuan and K. Larsson, *J. Phys. Chem. C.*, **2014**, 118, 26061.
- [45] X. Ru, Z. Cheng, L. Song, H. Wang and J. Li, *J Mol Struct.*, **2012**, 1030, 10.
- [46] W. A. Denny, In: D. E. V. Wilman (Ed.), *The Chemistry of Antitumour Agents* (Springer Netherlands, **1990**) 322.
- [47] D. R. Shah, R. P. Modh and K. H. Chikhaliya, *Future med. chem.*, **2014**, 6, 463.