informa
healthcare

# Quantitative structure–activity relationship (QSAR) studies on a series of 1,3,4-thiadiazole-2-thione derivatives as tumor-associated carbonic anhydrase IX inhibitors

MOHAMMED K. ABDEL-HAMID, ATEF A. ABDEL-HAFEZ, NAWAL A. EL-KOUSSI, & NADIA M. MAHFOUZ

*Department of Medicinal Chemistry, Faculty of Pharmacy, Assiut University, Assiut 71526, Egypt*

## Abstract

A linear quantitative structure–activity relationship (QSAR) study that encodes various aspects of physicochemical, topological and electronic descriptors has been developed for a series of 1,3,4-thiadiazole-2-thione derivatives (**1a-r** and **2a-c**). The carbonic anhydrase IX inhibitory activity of the candidates under study (**1a-r** and **2a-c**) were correlated to the selected parameters using stepwise linear regression analyses to achieve the best QSAR model. Promising results were obtained with the employed tetra-parametric model indicating that the information approach used in the present investigation is quite useful for modeling carbonic anhydrase IX inhibitors.

**Keywords:** *1,3,4-Thiadiazole-2-thione, QSAR analysis, topological descriptors, carbonic anhydrase inhibitors*

## Introduction

QSAR has been considered a major tool in drug discovery since the work of Hansch almost 40 years ago [1]. At the present time the QSAR study is one of the major fields of research in medicinal chemistry and drug design. QSAR is an effective way of optimizing or correlating certain structural features or molecular descriptors, such as lipophilicity, polarizability, or electronic and steric properties within a congeneric series of molecules with their biological activities. A better understanding of the modes of actions, prediction of newer analogs with better activity, classification of active/inactive compounds, and optimization of the lead compound to reduce toxicity and increase selectivity are the main goals of QSAR studies [2]. In addition, the methodology of QSAR is very helpful in screening a large library of possible drug candidates for selectivity and potency [3–8].

The carbonic anhydrases (CAs, EC 4.2.1.1) [9–12] constitute interesting targets for the design of pharmacological agents useful in the treatment or prevention of a variety of disorders [13,14]. A quite new and unexpected application of the CA inhibitors (CAIs) regard their potential use in the management (imaging and treatment) of hypoxic tumors [15–22], since at least two CA isozymes of the 15 presently known in humans [9–13], CA IX and XII, are predominantly found in tumor cells and lack (or are present in very limited amount) in normal tissues [14,23–26].

Most of the potent CAIs investigated up to now belong to the aromatic/heterocyclic sulfonamide or sulfamate classes [9–13,27,28], although compounds incorporating other zinc-binding groups have also been investigated [9–13,29,30]. Due to the biological importance of such agents as potent CAIs, many QSAR models have been proposed for the prediction of CA inhibitory activity of different aromatic and heterocyclic sulfonamides using different molecular descriptors.

Recently, we have reported the first thiadiazole-thione non-sulfonamide CAIs (**1a-r** and **2a-c**) showing inhibitory activity against the tumor-associated CA IX

Correspondence: Atef A. Abdel-Hafez, Department of Pharmaceutical Medicinal Chemistry, Faculty of Pharmacy, Assiut University, Assiut 71526, Egypt. Tel: 2088303006. Fax: 20882332776. E-mail: atef@aun.edu.eg

RIGHTSLINK

in low micromolar range [31]. In the present study, a QSAR model is investigated in order to predict the CA IX inhibitory activity of certain thiadiazole-thione derivatives (**1a-r** and **2a-c**) using different groups of molecular descriptors. Our objective is to develop a rapid and reliable method to predict the CA inhibition activity of this new group of candidates. In addition, we aim to gain more information that could be used in further development and optimization of such CA IX inhibitors.

## Materials and methods

### Data set

In this investigation, twenty one non-sulfonamide CAIs (**1a-r** and **2a-c**), that were recently synthesized and reported by our group, are used and their

structures are depicted in Figure 1. *In vitro* inhibitory activity of these structures (**1a-r** and **2a-c**) against human tumor-associated carbonic anhydrase isozyme IX (hCA IX) are expressed as $\log IC_{50}$ in Table II.

### Descriptors calculations

Electronic descriptors calculations were carried out using "Chem. Draw Ultra" version 8.0 and MOPAC package included in "Chem. 3D Ultra" version 8.0, Cambridge Software Corporation, USA.

The structures of the compounds (**1a-r** and **2a-c**) were drawn in Chem. Draw and saved as a template structure. For every compound, the template structure was suitably changed considering its structural features, copied to Chem. 3D to create the 3D model and finally the model was cleaned up. The energy
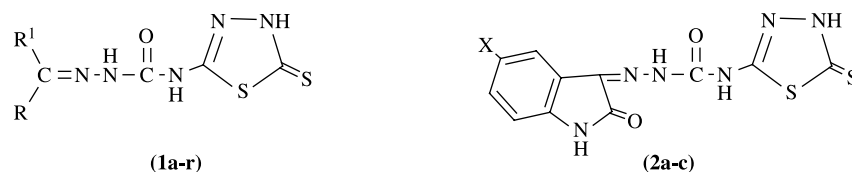


| Compound No. | R | $R^1$ | X |
|:---:|:---:|:---:|:---:|
| **1a** | H | $C_6H_5$ | |
| **1b** | H | $4\text{-}(OH)C_6H_4$ | |
| **1c** | H | $4\text{-}(OCH_3)\,C_6H_4$ | |
| **1d** | H | $3\text{-}(OCH_3)\text{-}4\text{-}(OH)C_6H_3$ | |
| **1e** | H | $2\text{-}(Cl)C_6H_4$ | |
| **1f** | H | $3\text{-}(Br)C_6H_4$ | |
| **1g** | H | $4\text{-}(Br)C_6H_4$ | |
| **1h** | H | $2\text{-}(NO_2)C_6H_4$ | |
| **1i** | H | $4\text{-}(NO_2)C_6H_4$ | |
| **1j** | H | $2\text{-}[N(CH_3)_2]C_6H_4$ | |
| **1k** | H | 3-pyridyl | |
| **1l** | H | 2-furyl | |
| **1m** | H | $CH_2CH(CH_3)_2$ | |
| **1n** | $CH_3$ | $C_6H_5$ | |
| **1o** | $CH_3$ | $4\text{-}(Cl)C_6H_4$ | |
| **1p** | $CH_3$ | $4\text{-}(Br)C_6H_4$ | |
| **1q** | $CH_3$ | $4\text{-}(NO_2)C_6H_4$ | |
| **1r** | $C_6H_5$ | $C_6H_5$ | |
| **2a** | | | H |
| **2b** | | | Br |
| **2c** | | | $NO_2$ |

Figure 1. The structures of 1,3,4-thiadiazole-2-thione derivatives (**1a-r** and **2a-c**) under study.

minimization was carried out under a Molecular Orbital Package (MOPAC) module (CS MOPAC 2000) of Chem. 3D Ultra version 8.0. The energy minimized geometry was used for the calculation of polarizability ($\Pi$), dipole moment ($\mu$), Connolly Accessible Area (CAA) and ionization potential (IP) of the molecules using PM3 Hamiltonian.

Physicochemical and topological descriptors calculations were carried out using "Molecular Operating Environment (MOE) version 2006.02", Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, H3A 2R7, Canada.

The fourth-order valence connectivity index was calculated according to the dissection rule applied by Kier and Hall [34].

Statistical analyses were carried out using "Microsoft Excel XP" software in "Microsoft Office XP" version 2002, Microsoft Corporation, USA and SPSS software, version 10.0, SPSS Inc., USA.

## Results and discussion

In an effort to determine the role of structural features of the investigated derivatives (**1a-r** and **2a-c**), QSAR studies were undertaken using the linear regression analysis model. Carbonic anhydrase inhibitory activity expressed as $\log IC_{50}$ was used as a dependent variable in this study.

### Molecular descriptors

Three groups of molecular descriptors of the synthesized compounds (**1a-r** and **2a-c**) were selected and calculated to cover a wide range of electronic, hydrophobic and topological characters of the molecules. Table I shows the descriptors used in this study together with the tools used for their calculation.

In addition to the calculated descriptors an indicator variable (I), was used to indicate the presence (1.0) or absence (0.0) of hydrogen atom at the benzylic carbon in compounds **1a-r**. The calculated descriptors, indicator variable (I) and $\log IC_{50}$ values against hCA IX are listed in Table II.

### Correlation matrix of the calculated descriptors

Multicolinearity occurs when two independent variables are correlated with each other and therefore contribute redundant information to the model. However, the multicolinearity problem was recently discussed by Randic; his recommendations were used in discussing our results [36].

The simplest method of investigating occurrence of multicolinearity is to obtain the correlation matrix which indicates that most of the descriptors used are not highly correlated (Table III). We decide to avoid the combinations between the highly intercorrelated descriptors with $|r| \geq 0.90$, where $r$ is the simple linear coefficient.

For the moderately intercorrelated descriptors we have to seriously consider Randic recommendations [36]. Randic stated that selection of the descriptors to be used in QSAR studies should not be delegated solely to the computers, although the statistical criteria will continue to be useful for preliminary screening of the descriptors taken from a large pool. Often in an automated selection of parameters, a descriptor will be discarded because it is highly correlated with another descriptor already selected. But what is important is not a descriptor parallel one another, but whether the two descriptors differ in those parts that are important for the activity under concern or not. Accordingly, if the two descriptors differ in the domain, which is important for the property/activity considered then **both** descriptors will be retained in the same model. On the other hand, if they differ in parts that are not relevant for the correlation of considered property/activity then **one** of them can be discarded and only the other could be used for further calculation.

These recommendations of Randic will be applied later in the generation of the multi-parametric models.

### Stepwise linear regression analyses

For the current data set of 21 compounds (**1a-r** and **2a-c**), the QSAR model development was restricted to a maximum of four variables in accordance to the

Table I.　Calculated molecular descriptors used in this study.

| Symbol | Description | Type | Server |
|---|---|---|---|
| Log P | Log of the octanol/water partition coefficient | Physicochemical | MOE [32] |
| W | Wiener index | Topological | Chem3D [33] |
| $W_p$ | Wiener polarity number | Topological | MOE |
| $^4\chi_{pc}^v$ | Fourth-order valence connectivity index | Topological | Manual [34] |
| Z | Zagreb index | Topological | MOE |
| PJ | Value of (diameter - radius)/diameter | Topological | MOE |
| $A_w$ | Van der Waals surface area | Topological | MOE |
| $V_w$ | Van der Waals volume | Topological | MOE |
| $\Pi$ | Mean polarizability | Electronic | MOPAC [35] |
| CAA | Connolly Accessible Area | Topological | Chem3D |
| $\mu$ | Magnitude of dipole moment | Electronic | MOPAC |
| IP | Ionization Potential | Electronic | MOPAC |

Table II. The inhibitory activity ($\log IC_{50}$) and the structural descriptors of the selected compounds (**1a-r** and **2a-c**).

| No. | $\log IC_{50}$ | Log P | W | $W_p$ | $^4\chi^v_{pc}$ | Z | PJ | $A_w$ | $V_w$ | $\Pi$ | CAA | $\mu$ | IP | I |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1a** | 1.362 | 4.73 | 745 | 19 | 0.738 | 86 | 0.50 | 275.17 | 238.13 | 177.15 | 484.17 | 19.66 | 9.21 | 1 |
| **1b** | 0.857 | 4.43 | 873 | 21 | 0.824 | 92 | 0.46 | 285.88 | 243.25 | 187.59 | 489.46 | 18.89 | 9.07 | 1 |
| **1c** | 1.380 | 4.69 | 1020 | 23 | 0.885 | 96 | 0.50 | 305.99 | 265.38 | 189.91 | 526.74 | 20.32 | 8.94 | 1 |
| **1d** | 1.415 | 4.42 | 1131 | 26 | 0.969 | 102 | 0.50 | 312.76 | 266.25 | 183.60 | 534.51 | 17.90 | 9.10 | 1 |
| **1e** | 1.462 | 5.33 | 849 | 22 | 1.000 | 92 | 0.50 | 291.58 | 251.25 | 174.86 | 483.15 | 25.46 | 6.89 | 1 |
| **1f** | 1.006 | 5.57 | 861 | 21 | 1.071 | 92 | 0.50 | 303.02 | 262.75 | 190.72 | 513.18 | 20.07 | 9.20 | 1 |
| **1g** | 1.342 | 5.54 | 873 | 21 | 1.115 | 92 | 0.46 | 303.27 | 264.13 | 184.46 | 513.99 | 18.78 | 9.20 | 1 |
| **1h** | 1.380 | 4.27 | 1097 | 26 | 0.959 | 102 | 0.50 | 298.04 | 254.13 | 187.12 | 516.41 | 22.04 | 9.25 | 1 |
| **1i** | 1.002 | 4.27 | 1169 | 25 | 0.927 | 102 | 0.50 | 299.06 | 256.50 | 189.84 | 522.14 | 16.49 | 9.25 | 1 |
| **1j** | 0.531 | 4.65 | 1169 | 25 | 1.231 | 102 | 0.50 | 329.65 | 284.63 | 217.69 | 561.79 | 20.10 | 9.02 | 1 |
| **1k** | 0.874 | 4.67 | 831 | 22 | 0.989 | 92 | 0.50 | 295.15 | 252.25 | 185.22 | 505.94 | 19.79 | 9.22 | 0 |
| **1l** | 0.934 | 5.26 | 965 | 24 | 1.205 | 98 | 0.46 | 311.56 | 267.00 | 192.65 | 528.08 | 19.21 | 9.20 | 0 |
| **1m** | 0.839 | 5.46 | 965 | 24 | 1.361 | 98 | 0.46 | 322.05 | 280.50 | 200.84 | 537.4 | 19.02 | 9.23 | 0 |
| **1n** | 0.880 | 4.20 | 1275 | 28 | 1.176 | 108 | 0.50 | 319.01 | 272.50 | 202.12 | 540.15 | 17.71 | 9.20 | 0 |
| **1o** | 0.928 | 6.33 | 1450 | 31 | 1.165 | 120 | 0.50 | 355.85 | 314.88 | 203.49 | 592.27 | 19.32 | 9.27 | 0 |
| **1p** | 0.707 | 3.43 | 745 | 22 | 0.726 | 92 | 0.50 | 289.92 | 248.13 | 197.65 | 473.04 | 17.45 | 9.16 | 0 |
| **1q** | 0.763 | 3.45 | 632 | 16 | 0.692 | 82 | 0.45 | 257.29 | 213.75 | 169.50 | 455.83 | 19.88 | 9.05 | 0 |
| **1r** | 0.948 | 4.28 | 627 | 17 | 0.958 | 78 | 0.45 | 293.39 | 246.00 | 165.49 | 499.38 | 21.54 | 9.12 | 0 |
| **2a** | 0.975 | 3.35 | 1025 | 28 | 1.139 | 110 | 0.50 | 294.51 | 256.63 | 206.35 | 503.31 | 20.17 | 8.94 | 0 |
| **2b** | 0.859 | 4.19 | 1157 | 30 | 1.486 | 116 | 0.50 | 319.46 | 283.38 | 217.97 | 534.89 | 20.14 | 8.98 | 0 |
| **2c** | 0.113 | 3.32 | 1467 | 34 | 1.316 | 126 | 0.46 | 317.05 | 273.38 | 253.03 | 539.98 | 21.94 | 9.01 | 0 |

generaly accepted rule of thumb. Based on this rule for a QSAR model development, one should use five compounds for each parameter selected (5:1 for compounds:descriptor) [37,38].

*Mono-parametric models*

The simple linear regression analysis yields two statistically significant correlations, one contains the polarizability parameter ($\Pi$) and the other includes the indicator variable (I). However, both models showed weak correlation with the inhibitory activity, where the former model shows better statistical results (**Model 1**).

$$\text{Log IC}_{50} = 3.2734 - 0.0118(\pm0.0028)\Pi \quad (\textbf{Model 1})$$

$n = 21$, $S_e = 0.2435$, $r = 0.6931$, $R^2_A = 0.4530$, $F = 17.5660$, $Q = 2.8459$.

Here and thereafter:

*n*: number of compounds
$S_e$: standard error of estimation
*r* (*R*):simple (multiple) correlation coefficient
$R^2_A$: adjustable $R^2$
*F*: Fisher's statistics
*Q*: quality factor

The negative coefficient of polarizability indicates that increasing the molecular polarizability is not a favored factor for the CA IX inhibitory activity of the molecule.

*Bi-parametric models*

Stepwise regression analyses using different combinations of two structural descriptors resulted into seven bi-parametric models (Table IV), which having

Table III. Intercorrelation matrix of the molecular descriptors.

| | Log P | W | $W_p$ | $^4\chi^v_{pc}$ | Z | PJ | $A_w$ | $V_w$ | $\Pi$ | CAA | $\mu$ | IP | I |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Log P | 1.00 | | | | | | | | | | | | |
| W | 0.07 | 1.00 | | | | | | | | | | | |
| $W_p$ | 0.07 | 0.94 | 1.00 | | | | | | | | | | |
| $^4\chi^v_{pc}$ | 0.24 | 0.61 | 0.69 | 1.00 | | | | | | | | | |
| Z | −0.06 | 0.94 | 0.99 | 0.68 | 1.00 | | | | | | | | |
| PJ | 0.06 | 0.29 | 0.31 | −0.04 | 0.27 | 1.00 | | | | | | | |
| $A_w$ | 0.47 | 0.79 | 0.73 | 0.74 | 0.72 | 0.22 | 1.00 | | | | | | |
| $V_w$ | 0.5 | 0.78 | 0.74 | 0.75 | 0.74 | 0.26 | 0.99 | 1.00 | | | | | |
| $\Pi$ | −0.22 | 0.75 | 0.83 | 0.68 | 0.85 | 0.07 | 0.59 | 0.6 | 1.00 | | | | |
| CAA | 0.44 | 0.84 | 0.73 | 0.7 | 0.72 | 0.22 | 0.96 | 0.95 | 0.57 | 1.00 | | | |
| $\mu$ | 0.06 | −0.1 | −0.02 | 0.11 | −0.04 | −0.05 | −0.11 | −0.1 | −0.02 | −0.15 | 1.00 | | |
| IP | −0.1 | 0.13 | 0.07 | 0.02 | 0.09 | −0.14 | 0.16 | 0.13 | 0.15 | 0.26 | −0.72 | 1.00 | |
| I | 0.27 | −0.07 | −0.25 | −0.33 | −0.26 | 0.29 | −0.16 | −0.14 | −0.3 | −0.08 | 0.09 | −0.21 | 1.00 |

Table IV.    Regression parameters for the bi-parametric models.

| Descriptors | $S_e$ | $R$ | $R_A^2$ | $F$ |
|---|---|---|---|---|
| $\Pi$, $W_p$ | 0.2010 | 0.8154 | 0.6277 | 17.8609 |
| $\Pi$, Z | 0.2098 | 0.7968 | 0.5943 | 15.6536 |
| $\Pi$, PJ | 0.2102 | 0.7961 | 0.5931 | 15.5790 |
| $\Pi$, I | 0.2106 | 0.7952 | 0.5914 | 15.4748 |
| $\Pi$, W | 0.2192 | 0.7757 | 0.5574 | 13.5970 |
| $\Pi$, $V_w$ | 0.2289 | 0.7520 | 0.5172 | 11.7121 |
| $\Pi$, log P | 0.2353 | 0.7355 | 0.4899 | 10.6055 |

significantly better statistics than the mono-parametric model discussed above. It was noticed that the polarizability is a common descriptor in all of the produced bi-parametric models.

Out of these combinations, the bi-parametric model that contains the polarizability and the Wiener polarity number was the best one (**Model 2**).

$$\text{Log IC}_{50} = 4.0744 - 0.0231(\pm 0.0042)\Pi$$
$$+ 0.0577(\pm 0.0183)W_p \quad (\textbf{Model 2})$$

$n = 21$,  $S_e = 0.2010$,  $R = 0.8154$,  $R_A^2 = 0.6277$, $F = 17.8609$, $Q = 4.0567$.

This model showed a good correlation between the inhibitory activity (log IC$_{50}$) and the selected ($\Pi$, $W_p$) descriptors with a significant improvement in its statistical parameters compared with (**Model 1**).

*Tri-parametric models*

Trials were made to correlate three combined descriptors with the biological activity (log IC$_{50}$) aiming to improve the statistical parameters of the obtained models. Five tri-parametric models were obtained from stepwise regression analyses (Table V).

Two out of these tri-parametric models showed excellent statistics (**Models 3** and **4**), where the former is slightly better.

$$\text{Log IC}_{50} = 3.5490 - 0.0210(\pm 0.0033)\Pi$$
$$+ 0.0578(\pm 0.0139)W_p$$
$$+ 0.2631(\pm 0.0699)I \quad (\textbf{Model 3})$$

Table V.    Regression parameters for the tri-parametric models.

| Descriptors | $S_e$ | $R$ | $R_A^2$ | $F$ |
|---|---|---|---|---|
| $\Pi$, $W_p$, I | 0.1527 | 0.9040 | 0.7850 | 25.3341 |
| $\Pi$, Z, I | 0.1642 | 0.8882 | 0.7516 | 21.1671 |
| $\Pi$, $W_p$, PJ | 0.1909 | 0.8453 | 0.6642 | 14.1857 |
| $\Pi$, PJ, I | 0.1929 | 0.8417 | 0.6570 | 13.7708 |
| $\Pi$, $V_w$, I | 0.1932 | 0.8412 | 0.6560 | 13.7108 |

$n = 21$,  $S_e = 0.1527$,  $R = 0.9040$,  $R_A^2 = 0.7850$, $F = 25.3341$, $Q = 5.9201$.

$$\text{Log IC}_{50} = 2.8303 - 0.0207(\pm 0.0037)\Pi$$
$$+ 0.0206(\pm 0.0058)Z$$
$$+ 0.2645(\pm 0.0751)I \quad (\textbf{Model 4})$$

$n = 21$,  $S_e = 0.1642$,  $R = 0.8882$,  $R_A^2 = 0.7516$, $F = 21.1671$, $Q = 5.4080$.

The results indicated that the inclusion of the indicator variable I with certain topological descriptors produces better correlations.

In the above tri-parametric models **3** and **4**, the correlation parameter I has a positive coefficient indicating that derivatives contain a hydrogen atom on the benzylic carbon (**1a-r**) generally have better inhibitory activity. On the other hand, increasing molecular branching and complexity enhances the biological activity as indicated from the positive coefficient of Zagreb index (**Model 4**).

*Tetra-parametric models*

Finally, successive stepwise regression yielded one tetra-parametric model (**model 5**) of slight improvement in statistical parameters.

$$\text{Log IC}_{50} = 3.5578 - 0.0219(\pm 0.0033)\Pi$$
$$+ 0.0527(\pm 0.0144)W_p$$
$$+ 0.2687(\pm 0.2306)^4\chi_{pc}^v$$
$$+ 0.2791(\pm 0.0705)I \quad (\textbf{Model 5})$$

$n = 21$,  $S_e = 0.1512$,  $R = 0.9120$,  $R_A^2 = 0.7894$, $F = 19.7397$, $Q = 6.0304$.

It is noteworthy that, in this model the fourth-order valence connectivity index ($^4\chi_{pc}^v$) appears for the first time in the multi-parametric models. This index encodes the orientation, the length and the presence of heteroatom in ring substituent. The positive coefficient of this index indicates that increasing the number and length of ring substituents as well as the presence of heteroatom containing substituents is favored for the inhibitory activity.

*Validation and cross-validation of model (5)*

Cross-validation statistical technique has been applied to estimate the quality with regard to predictive ability of the generated model (**model 5**) using MOE 2005.06 software [32]. The simplest and most general cross-validation procedure is the leave-one-out technique (LOO technique), where each object of the data set is taken away, one at a time. In this case, given n objects, n reduced models are developed [39].

Table VI. Validation and cross-validation parameters for model (5).

| No. | Obs. $\log IC_{50}$ | MLR validation | | | Cross validation (LOO) | | |
|---|---|---|---|---|---|---|---|
| | | Pred. $\log IC_{50}$ | Residual | Z-score | Pred. $\log IC_{50}$ | Residual | Z-score |
| 1a | 1.362 | 1.158 | 0.204 | 1.554 | 1.114 | 0.248 | 1.987 |
| 1b | 0.857 | 1.043 | $-0.185$ | 1.411 | 1.077 | $-0.219$ | 1.732 |
| 1c | 1.380 | 1.120 | 0.259 | 1.978 | 1.084 | 0.295 | 2.476 |
| 1d | 1.415 | 1.435 | $-0.020$ | 0.153 | 1.441 | $-0.026$ | 0.197 |
| 1e | 1.462 | 1.431 | 0.031 | 0.241 | 1.424 | 0.038 | 0.284 |
| 1f | 1.006 | 1.042 | $-0.036$ | 0.273 | 1.051 | $-0.044$ | 0.329 |
| 1g | 1.342 | 1.187 | 0.155 | 1.186 | 1.148 | 0.194 | 1.509 |
| 1h | 1.380 | 1.366 | 0.014 | 0.109 | 1.362 | 0.018 | 0.134 |
| 1i | 1.002 | 1.238 | $-0.236$ | 1.795 | 1.278 | $-0.275$ | 2.263 |
| 1j | 0.531 | 0.720 | $-0.189$ | 1.436 | 0.827 | $-0.295$ | 2.383 |
| 1k | 0.874 | 0.931 | $-0.056$ | 0.427 | 0.938 | $-0.063$ | 0.478 |
| 1l | 0.934 | 0.917 | 0.017 | 0.135 | 0.914 | 0.0206 | 0.154 |
| 1m | 0.839 | 0.781 | 0.058 | 0.442 | 0.759 | 0.079 | 0.595 |
| 1n | 0.880 | 0.922 | $-0.041$ | 0.315 | 0.928 | $-0.048$ | 0.359 |
| 1o | 0.928 | 1.034 | $-0.105$ | 0.800 | 1.082 | $-0.154$ | 1.171 |
| 1p | 0.707 | 0.572 | 0.135 | 1.031 | 0.484 | 0.223 | 1.737 |
| 1q | 0.763 | 0.866 | $-0.103$ | 0.782 | 0.924 | $-0.160$ | 1.221 |
| 1r | 0.948 | 1.079 | $-0.130$ | 0.991 | 1.136 | $-0.187$ | 1.446 |
| 2a | 0.975 | 0.823 | 0.152 | 1.160 | 0.801 | 0.174 | 1.344 |
| 2b | 0.859 | 0.755 | 0.103 | 0.791 | 0.717 | 0.142 | 1.078 |
| 2c | 0.113 | 0.146 | $-0.032$ | 0.245 | 0.191 | $-0.077$ | 0.577 |

The observed activity (Obs. $\log IC_{50}$) together with the predicted activity (Pred. $\log IC_{50}$) for the tested compounds calculated using multi-linear regression (MLR) and LOO techniques are listed in Table VI. All compounds showed very good results with Z-scores not exceed the value of 2.5 indicating high predictive ability of the model. The observed $\log IC_{50}$ is plotted against their predicted values (calculated by MLR method) with a value of $R^2$ found to be 0.8331 (Figure 2).

*Statistical features*

It is now important to discuss the significant of some statistical parameters to judge the quality of our work and support the obtained results. These parameters include probable error of correlation (PE), adjustable $R^2$ ($R_A^2$) and the quality parameter (Q).
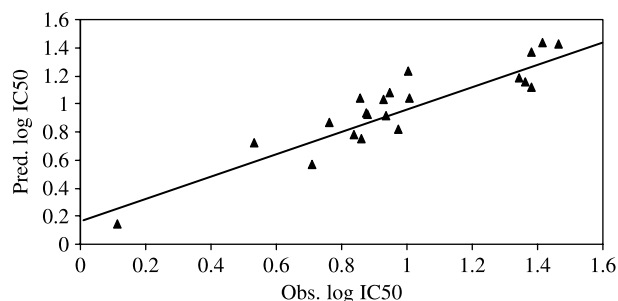


Figure 2. Correlation of observed and predicted $\log IC_{50}$ using model 5 ($R^2 = 0.8331$, MLR method).

Probable error of correlation (PE); (Table VII) is calculated by the following expression:

$$PE = 2(1 - r^2)/3\sqrt{n}$$

It is recommended that:

If, r < PE, it is not statistically significant
r > PE several times, at least three times greater, the correlation is statistically significant and
r > 6 PE, correlation is definitely significant.

For all the proposed models r values were found to be highly greater than 6PE. Thus all the correlations attempted are definitely acceptable.

The statistical parameter adjustable $R^2$ ($R_A^2$) takes into account of adjustment of $R^2$ and the statistical significance of stepwise regression (Table VII). Therefore, if a variable is added that does not contribute its fair share, the $R_A^2$ will actually decline. If $R_A^2$ values goes on increasing as we pass from bi- to higher parametric regressions, then the added parameters

Table VII. Some statistical parameters of the generated models.

| Model No. | $R$ | $R_A^2$ | $Q$ | PE |
|---|---|---|---|---|
| 1 | 0.6931 | 0.4530 | 2.8459 | 0.0756 |
| 2 | 0.8154 | 0.6277 | 4.0567 | 0.0487 |
| 3 | 0.9040 | 0.7850 | 5.9201 | 0.0266 |
| 4 | 0.8882 | 0.7516 | 5.4080 | 0.0307 |
| 5 | 0.9120 | 0.7894 | 6.0304 | 0.0245 |

have favorable contribution in the exhibition of $\log IC_{50}$.

It is necessary that, the proposed model should have both statistical quality as well as better predictive power. The simplest parameter to decide the predictive power of the model is the quality parameter $Q$, the highest value of which indicates best predictive power. In all the models discussed above $Q$ goes on increasing as we pass from mono- to higher parametric regressions (Table VII). Consequently, all the models generated here are quality models having quality predictive power.

## Conclusion

The QSAR studies performed on a small series of thiadiazole CA inhibitors (**1a-r** and **2a-c**) produced five statistically acceptable models. The produced models presented some information about the importance of some structural descriptors for the biological activity of the compound under study. It was found that molecular polarizability is not a favorable contributor for activity. On the other hand, branching, increasing the length and complexity of ring substituent as well as the presence of hydrogen atom on the benzylic carbon could increase the CA IX inhibitory activity of the synthesized compounds (**1a-r** and **2a-c**).

***Declaration of interest***: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## References

[1] Hansch C, Fujita T. A method for the correlation of biological activity and chemical structure. J Am Chem Soc 1964;86: 1616–1626.

[2] Prosenjit B, Kounal R. QSAR of adenosine $A_3$ receptor antagonist 1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives using chemometric tools. Bioorg Med Chem Lett 2005;15: 3737–3743.

[3] Diudea MV. QSAR/QSPR studies by molecular descriptors. Hungtington: Nova Science Publishers; 2000.

[4] Kubinyi H, Folkers G, Martin VC. 3D QSAR in drug design ligand–protein interaction and molecularity. Amsterdam, The Netherlands: Klawer/ESCOM; 1998.

[5] Kubinyi H, Folkers G, Martin YC. 3D QSAR in drug design: Recent advances. Amsterdam, The Netherlands: Klawer/ESCOM; 1998.

[6] Devilliers J. Comparative QSAR. Washington, DC: Taylor & Francis; 1998.

[7] Devillers J, Balaban AT. Topological indices and related descriptors in QSAR and QSPR. Amsterdam, The Netherlands: Gordon and Breach Science Publication; 1999.

[8] Karelson M. Molecular descriptors in QSAR/QSPR. New York: John Wiley & Sons; 2000.

[9] Supuran CT. Carbonic anhydrases: Novel therapeutic applications for inhibitors and activators. Nat Rev Drug Discov 2008;7:168–181.

[10] Scozzafava A, Mastrolorenzo A, Supuran CT. Modulation of carbonic anhydrase activity and its applications in therapy. Expert Opin Ther Patents 2004;14:667–702.

[11] Supuran CT. Carbonic anhydrases: Catalytic and inhibition mechanisms, distribution and physiological roles. In: Supuran CT, Scozzafava A, Conway J, editors. Carbonic anhydrase—its inhibitors and activators. Boca Raton: CRC; 2004. p 1–23.

[12] Supuran CT, Vullo D, Manole G, Casini A, Scozzafava A. Designing of novel carbonic anhydrase inhibitors and activators. Curr Med Chem Cardiovasc Hematol Agents 2004;2:49–68.

[13] Pastorekova S, Parkkila S, Pastorek J, Supuran CT. Proceedings of the 6th International Conference on Carbonic Anhydrases. June, 2003. Bratislava, Slovak Republic. J Enz Inhib Med Chem 2004;19:199–291.

[14] Supuran CT. Carbonic anhydrase inhibitors in the treatment and prophylaxis of obesity. Expert Opin Ther Patents 2003;13: 1545–1550.

[15] Svastova E, Hulikova A, Rafajova M, Zat'ovicova M, Gibadulinova A, Casini A, Cecchi A, Scozzafava A, Supuran CT, Pastorek J, Pastorekova S. Hypoxia activates the capacity of tumor-associated carbonic anhydrase IX to acidify extracellular pH. FEBS Lett 2004;577:439–445.

[16] Supuran CT. Indisulam: An anticancer sulfonamide in clinical development. Expert Opin Invest Drugs 2003;12:283–287.

[17] Supuran CT, Scozzafava A, Casini A. Carbonic anhydrase inhibitors. Med Res Rev 2003;23:146–189.

[18] Scozzafava A, Owa T, Mastrolorenzo A, Supuran CT. Anticancer and antiviral sulfonamides. Curr Med Chem 2003;10:925–953.

[19] Vullo D, Franchi M, Gallori E, Pastorek J, Scozzafava A, Pastorekova S, Supuran CT. Carbonic anhydrase inhibitors. Inhibition of cytosolic isozymes I and II and transmembrane, cancer-associated isozyme IX with anions. J Enz Inhib Med Chem 2003;18:403–406.

[20] Vullo D, Franchi M, Gallori E, Pastorek J, Scozzafava A, Pastorekova S, Supuran CT. Carbonic anhydrase inhibitors: Inhibition of the tumor-associated isozyme IX with aromatic and heterocyclic sulfonamides. Bioorg Med Chem Lett 2003; 13:1005–1009.

[21] Winum JY, Vullo D, Casini A, Montero JL, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: Inhibition of transmembrane, tumor-associated isozyme IX, and cytosolic isozymes I and II with aliphatic sulfamates. J Med Chem 2003; 46:5471–5477.

[22] Winum JY, Vullo D, Casini A, Montero JL, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors. Inhibition of cytosolic isozymes I and II and transmembrane, tumor-associated isozyme IX with sulfamates including EMATE also acting as steroid sulfatase inhibitors. J Med Chem 2003;46: 2197–2204.

[23] Potter C, Harris AL. Hypoxia inducible carbonic anhydrase IX, marker of tumour hypoxia, survival pathway and therapy target. Cell Cycle 2004;3:164–167.

[24] Pastorekova S, Casini A, Scozzafava A, Vullo D, Pastorek J, Supuran CT. Carbonic anhydrase inhibitors: The first selective, membrane-impermeant inhibitors targeting the tumor-associated isozyme IX. Bioorg Med Chem Lett 2004; 14:869–873.

[25] Rafajova M, Zatovicova M, Kettmann R, Pastorek J, Pastorekova S. Induction by hypoxia combined with low glucose or low bicarbonate and high posttranslational stability upon reoxygenation contribute to carbonic anhydrase IX expression in cancer cells. Int J Oncol 2004;24:995–1004.

[26] Robertson N, Potter C, Harris AL. Role of carbonic anhydrase IX in human tumor cell growth, survival, and invasion. Cancer Res 2004;64:6160–6165.

[27] Supuran CT, Scozzafava A, Casini A. Development of sulfonamide carbonic anhydrase inhibitors. In: Supuran CT, Scozzafava A, Conway J, editors. Carbonic anhydrase—its inhibitors and activators. Boca Raton: CRC; 2004. p 67–147.

[28] Casini A, Antel J, Abbate F, Scozzafava A, David S, Waldeck H, Schafer S, Supuran CT. Carbonic anhydrase inhibitors: SAR and X-ray crystallographic study for the interaction of sugar sulfamates/sulfamides with isozymes I, II and IV. Bioorg Med Chem Lett 2003;13:841–845.

[29] Bonnac L, Innocenti A, Winum JY, Casini A, Montero JL, Scozzafava A, Barragan V, Supuran CT. Carbonic anhydrase inhibitors: Aliphatic N-phosphorylated sulfamates—a novel zinc-anchoring group leading to nanomolar inhibitors. J Enz Inhib Med Chem 2004;19:275–278.

[30] Scozzafava A, Supuran CT. Hydroxyurea is a carbonic anhydrase inhibitor. Bioorg Med Chem 2003;11:2241–2246.

[31] Abdel-Hamid M, Abdel-Hafez A, El-Koussi N, Mahfouz N, Innocenti A, Supuran CT. Design, synthesis, and docking studies of new 1,3,4-thiadiazole-2-thione derivatives with carbonic anhydrase inhibitory activity. Bioorg Med Chem 2007;15:6975–6984.

[32] Molecular Operating Environment (MOE 2005.06) version 2006.02, Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, H3A 2R7, Canada.

[33] Chem. 3D Ultra version 8.0, Cambridge Software Corporation, USA.

[34] Kier LB, Hall LH. Molecular connectivity in structure activity analysis. USA: John Wiley & Sons Inc. 1984.

[35] Molecular Orbital Package (MOPAC) module (CS MOPAC 2000) of Chem. 3D Ultra version 8.0, Cambridge Software Corporation, USA.

[36] Jaiswal M, Khadikar V, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: The first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and hetero-cyclic sulfonamides. Bioorg Med Chem Lett 2004;14:3283–3290, and references therein.

[37] Manivannan E, Prasanna S. First QSAR report on FSH receptor antagonistic activity: Quantitative investigations on physico-chemical and structural features among 6-amino-4-phenyltetrahydroquinoline derivatives. Bioorg Med Chem Lett 2005;15:4496–4501.

[38] Narasimhan B, Kumari M, Jain N, Dhake A, Sundaravelan C. Correlation of antibacterial activity of some N-[5-(2-furanyl)-2-methyl-4-oxo-4H-thieno[2,3-d]pyrimidin-3-yl]-carboxamide and 3-substituted-5-(2-furanyl)-2-methyl-3H-thieno[2,3-d] pyrimidin-4-ones with topological indices using Hansch analysis. Bioorg Med Chem Lett 2006;16:4951–4958, and references therein.

[39] Melagraki G, Afantitis A, Sarimveis H, Igglessi-Markopoulou O, Supuran CT. QSAR study on para-substituted aromatic sulfonamides as carbonic anhydrase II inhibitors using topological information indices. Bioorg Med Chem 2006;14:1108–1114.