



## QSAR Study of the Biologically Active Organosulfurs in Natural Products

FAKHR M. ABU-AWWAD

Chemistry Department  
Islamic University of Gaza, Box 108 – Gaza, Palestine  
*dr-fakhr@live.com*

Received 9 March 2010; Accepted 4 May 2010

**Abstract:** The biological activities of two sets of a total of 30 different polysulfides were investigated using QSAR. The semiempirical AM1 in Gaussian 2003 for windows was used to optimize the structures whereas a subsequent calculation of hundreds of various types of descriptors at the density functional (B3LYP/6-31G\*) using CODESSA package was employed. The known 15-lipoxygenase inhibitory activity data ( $IC_{50}$ ) of 19 polysulfides out of the whole data set were correlated in a multiple linear regression procedures with the computed descriptors. Statistically, the most significant overall correlations were five- and four- parameter equations with good statistical parameters;  $R^2=0.9981$ ,  $R^2_{CV}=0.9970$  and  $R^2=0.9967$ ,  $R^2_{CV}=0.9933$  respectively. The models concluded that biological activity of polysulfides is mainly attributed to quantum-chemical, geometrical and topological descriptors with neither electrostatic contribution nor chief role of sulfur atoms. Also sulfur related descriptors were not the most significant contributors in the concluded models. The obtained models were efficiently employed to estimate the biological activities of the other 11 polysulfides available in natural products such as garlic and mushroom.

**Keywords:** Benzodiazepine, Toxicity, AM1, CODESSA, QSAR

### Introduction

Sulfur plays a major role in biology where organosulfur compounds are identified for their unique pharmacological tasks associated with redox processes, metal binding and catalytic activities. These tasks are the basis of the antibiotic and anticancer activities. Medicinal properties of *Allium* species, such as garlic, onion, or shallot, including their role as natural cures for various human viral, bacterial and fungal infections have been mainly attributed to their content of organosulfur compounds<sup>1</sup>.

Polysulfides are formed in *Allium* species by decomposition of its content of allicin, which is seen as the “mother antibiotic”. Diallyltrisulfide and tetrasulfide altogether with

higher polysulfides, such as the diallylpenta-, hexa- and heptasulfide have been reported as constituents of low concentrations<sup>2</sup>. The variation of antibiotic activity with increasing number of sulfur atoms has revealed a vague trend in many biological tests, where the relationship does not seem to be linear<sup>3</sup>. While monosulfides are often inactive, the disulfide has a limited activity. However, trisulfides and tetrasulfides have shown significant antibiotic efficiency with an ability to cause the death of certain cancer cells. Tapiero *et al* has recently reviewed the research activities in this field and concluded that diallyltrisulfide and diallyltetrasulfide of garlic exhibit a wide spectrum of antibacterial, antifungal, antimicrobial and anticancer activity<sup>4</sup>.

The chemistry underlying the biological activity of these organopolysulfides, mainly allicin itself, has been dragging a big deal of research at both experimental and computational levels seeking a better understanding of their reactions within pharmaceutical and biochemical contexts<sup>2-7</sup>. Eventually, this understanding may provide the basis for developing naturally sourced medications with dramatically reduced side effects. Along this line, garlic was investigated extensively for health benefits with hundreds of publications in the past few years<sup>8</sup>. Based on its potential and varied effects Allium family has been considered as one of the best disease-preventive foods with a wide range of therapeutic effects. Eating garlic, for example is one of the best methods to decrease high blood pressure and defend against cardiovascular disease due to the interaction between garlic's polysulfides and red blood cells which would release H<sub>2</sub>S to relax blood vessels<sup>9-10</sup>. Additionally, garlic prevents cold and flu symptoms through immune enhancement and exhibits anticancer and chemo-preventive activities<sup>11</sup>.

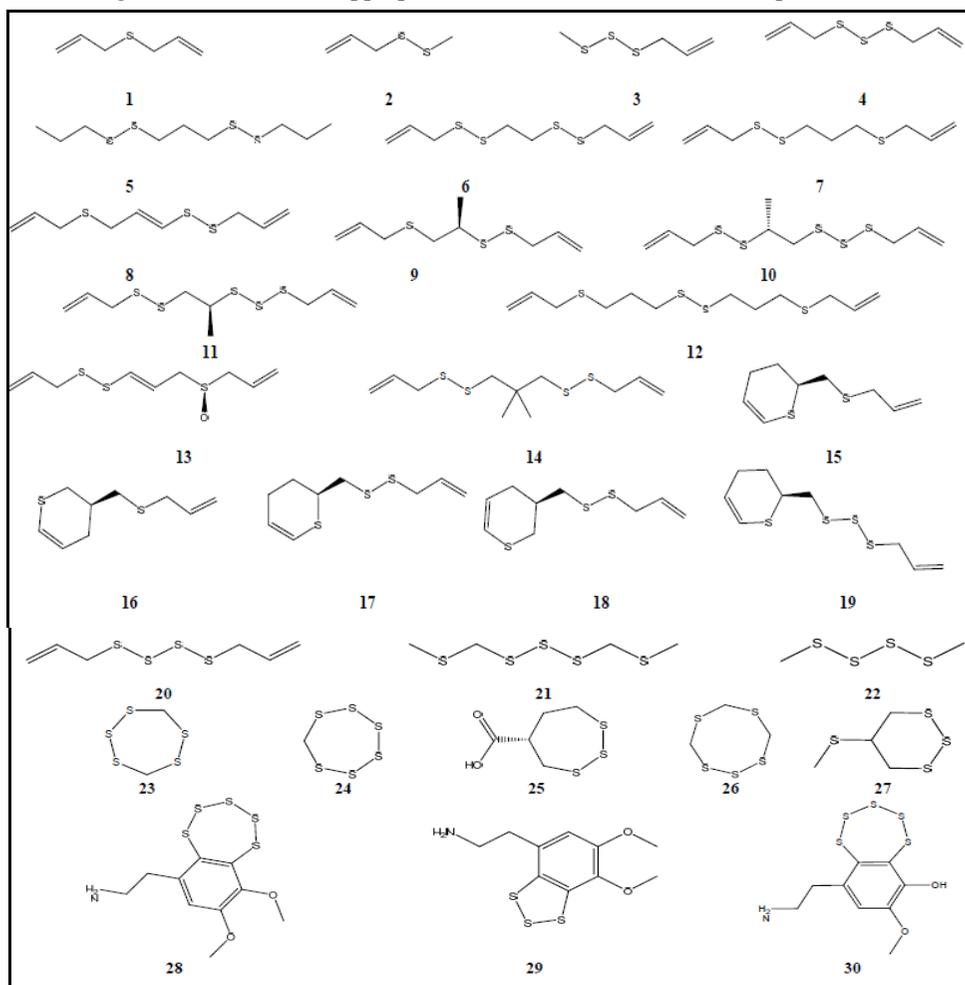
Quantitative structure activity relationships (QSAR) have been widely used to correlate biological activities of various biochemical compounds with various computed observables<sup>12-13</sup>. However, very limited number of computational studies including QSAR type of the organosulfurs in natural products was found in literature. Luco *et al* has a recent QSAR study where multiple linear regression (MLR) and partial least-squares (PLS) techniques were used for modeling the soybean 15-lipoxygenase inhibitory activity of a varied group of mono-, di- and trisulfides derived from the essential oils of garlic<sup>14</sup>. In their study, the model obtained was dependent on two nonempirical descriptors; the solvent-accessible surface area and the average distance/distance degree descriptor index. The energy of the lowest unoccupied molecular orbital ( $\epsilon_{LUMO}$ ) showed high contribution in their PLS models derived for the di- and trisulfides. However, the geometric related indices indicated a low positive contribution of molecular linearity in the enzyme-inhibitor binding.

Lipoxygenases (LOs) are natural non-heme iron-containing oxygenases; with their products are essential intermediates in the biosynthesis of several vital compounds in humans. The LO inhibitors have been classified into four distinct classes of which is the inhibitors of the 5-lipoxygenase activating protein. Block *et al* studied the inhibition of soybean 15-lipoxygenase (15-sLO) by using several garlic essential oil preparations to determine which organosulfurs are responsible for the observed biological activities<sup>15</sup>.

In this study, QSAR analysis is employed to correlate the reported inhibitory activity data of several organosulfur compounds by means of MLR with various types of descriptors in CODESSA software<sup>16</sup>. The data on 15-Lipoxygenase Inhibitory activity was taken from Luco previously published research<sup>14</sup>. The IC<sub>50</sub> values were used as the dependent variable in the correlation, where IC<sub>50</sub> represents the molar concentration of compound required to achieve 50% of inhibition of 15-Lipoxygenase.

### Computational work

Initial geometries of the structures 1-30 in scheme 1 were generated from standard geometrical parameters and were minimized without any constraint using the semiempirical AM1 method in GAUSSIAN 2003 for Windows (G03W)<sup>17</sup>. The optimized structures were properly attributed to their local minima at the same level of theory with no imaginary frequencies. For QSAR calculations, a single point run was carried again for each of the previously optimized structures at the density functional theory (DFT) with the three-parameter hybrid functional (B3)<sup>18</sup> for the exchange part and the Lee-Yang-Parr (LYP) correlation function<sup>19</sup> and the basis set 6-31G\* in G03W using the route command appropriate for CODESSA version 2.10 input file.



**Scheme 1.** Biologically active organosulfurs (1-30) used in the study

In CODESSA procedures, about 300 of various quantum chemical, electrostatic, geometrical, constitutional and topological descriptors were computed for the optimized organosulfurs (1-30). Heuristic method (HM) was applied to the first dataset of the structures (1-19) where pre-selection of descriptors was initiated utilizing the available, nicely correlated and variant descriptors only, which then were sorted according to their

subsequent correlation coefficients. Four criteria were employed to verify the validity of the reached models; the square coefficient of fitting model ( $R^2$ ), the square coefficient of cross validated ( $R^2_{cv}$ ), Fisher  $F$ -criterion value which reflects the ratio of the variance explained by the model and the variance due to the error in it, and the standard deviation of the regression ( $s^2$ ). Commonly, the closer the values of  $R^2$  and  $R^2_{cv}$ , the better the model is.

## Results and Discussion

Employing HM of CODESSA for the dataset (1-19), QSAR models of the best one-, two-, three-, four- and five- parameters are displayed in Table 1, where the maximum number of descriptors utilized was set to 5 as commonly suggested in similar MLR studies<sup>12-13</sup>. In the first step, only models for which  $R^2 > 0.6$  were considered for further selection, followed by a subsequent validation through discarding several models from the initial ones. Unexpectedly, the most significant descriptors in each of the models are quantum chemicals which relate mostly to C atoms but not to the structurally dominant S atoms in the compounds. This may insist that the relation between the number of sulfur atoms in organosulfur compounds and the subsequent biological activities may not necessarily be linear.

Three types of descriptors have been concluded in the correlated models in eqs 1-5; quantum chemical, electrostatic and geometrical. Obviously, neither constitutional nor topological descriptors have been concluded in any of the models. The concluded five descriptors are average electrophilic reaction index for a C atom ( $D_1$ ), minimum nucleophilic reaction index for a S atom ( $D_2$ ) and maximum SIGMA-SIGMA bond order ( $D_5$ ) as quantum chemicals; the geometrical Moment of inertia C ( $D_3$ ) and the electrostatic WNSA-1 Weighted PNSA ( $PNSA1^*TMSA/1000$ ) [Zefirov's PC] ( $D_4$ ).

The numerical values of the concluded descriptors which were related to the biological activity of the whole dataset (1-30) are listed in Table 2. Also, a comparison between the experimental and computed  $IC_{50}$  for 19 polysulfides (1-19) along with the predicted  $IC_{50}$  for the other 10 structures (20-30) are listed in Table 4. The HM based one-parameter regression expression is as in eq. (1), where  $D_1$  is the minimum electrophilic reaction index for a C atom.  $N$  is the number of compounds;  $R^2$  is the correlation coefficient,  $R^2_{cv}$  is the 'leave one out' (LOO) cross-validated coefficient,  $F$  is the Fisher-statistic

$$IC_{50} = -(375.23 \pm 81.619) + (1.0559 * 10^5 \pm 1.1344 * 10^4) * D_1 \quad (1)$$

$$N = 19; R^2 = 0.8360, R^2_{cv} = 0.7017, F = 86.63, s^2 = 36257$$

In the above equation, the quantum chemical descriptor  $D_1$  measures the average electrophilic reaction index for a C atom. It has a positive-sign coefficient, implying that increasing its magnitude would enhance the value of  $IC_{50}$  of the biological activity of the organopolysulfides. Obviously, being the only selected descriptor out of another about 300 descriptors implies its utmost significance of the quantum chemical description of C atom in the biological activity of the organosulfur natural compounds.

Among the obtained two-parameter models, statistically the best one is as follows:

$$IC_{50} = (-329.98 \pm 44.220) + (7.9331 * 10^4 \pm 7264.1) * D_1 + (7713.6 \pm 1171.6) * D_2 \quad (2)$$

$$N = 19; R^2 = 0.9558, R^2_{cv} = 0.9431, F = 172.89, s^2 = 10385.5$$

In this model,  $D_2$  is a quantum mechanical descriptor measures the minimum nucleophilic reaction index for S atom in each molecule. Again both  $D_1$  and  $D_2$  have coefficients with positive signs and thus are directly increasing the value of  $IC_{50}$ . According to t test of the model,  $D_1$  is almost twice as significant as  $D_2$  which obviously give the quantum chemical contribution a much higher part in this correlation.

**Table 1.** QSAR models and statistical parameters of the correlations in the present study

Eq.	X	DX	t-test	Descriptors	Symbol	Statistical parameters
1	-375.23	81.619	-4.597	<i>Intercept</i>		$R^2 = 0.836, R_{cv}^2 = 0.7017$ $F = 86.63, s^2 = 36256.9$
	$1.0559 \times 10^5$	$1.1344 \times 10^4$	9.3075	<i>Avg electroph. react. index for a C atom</i>	D <sub>1</sub>	
2	-329.98	44.220	-7.462	<i>Intercept</i>		$R^2 = 0.9558, R_{cv}^2 =$ $0.9431, F = 172.89$ $s^2 = 10385.5$
	$7.9331 \times 10^4$	$7.2641 \times 10^3$	10.921	<i>Avg electroph. react. index for a C atom</i>	D <sub>1</sub>	
3	$7.71136 \times 10^3$	$1.1716 \times 10^3$	6.584	<i>Min nucleoph. react. index for a S atom</i>	D <sub>2</sub>	$R^2 = 0.9905, R_{cv}^2 =$ $0.9850, F = 522.27$ $s^2 = 2375.3$
	-335.61	23.508	-14.276	<i>Intercept</i>		
4	$1.2004 \times 10^5$	$5.9370 \times 10^3$	20.219	<i>Avg electroph. react. index for a C atom</i>	D <sub>1</sub>	$R^2 = 0.9967, R_{cv}^2 =$ $0.9933, F = 1048.4$ $s^2 = 892.95$
	-1.4034 $\times 10^4$	$1.6696 \times 10^3$	-8.406	<i>Max electroph. react. index for a C atom</i>	D <sub>3</sub> D <sub>4</sub>	
5	$1.4138 \times 10^4$	$1.7460 \times 10^3$	8.098	<i>Moment of inertia B</i>		$R^2 = 0.9981, R_{cv}^2 =$ $0.9970, F = 1388.9$ $s^2 = 540.03$
	-796.37	61.719	-12.903	<i>Intercept</i>		
6	$8.1698 \times 10^4$	$2.8404 \times 10^3$	28.762	<i>Avg electroph. react. index for a C atom</i>	D <sub>1</sub>	$R^2 = 0.9981, R_{cv}^2 =$ $0.9970, F = 1388.9$ $s^2 = 540.03$
	$4.2566 \times 10^3$	$4.6408 \times 10^2$	9.806	<i>Min nucleoph. react. index for a S atom</i>	D <sub>2</sub>	
7	$2.2814 \times 10^4$	$1.7587 \times 10^3$	12.972	<i>Moment of inertia C</i>	D <sub>5</sub>	$R^2 = 0.9981, R_{cv}^2 =$ $0.9970, F = 1388.9$ $s^2 = 540.03$
	3.3589	0.4686	7.1679	<i>WNSA-1 Weighted PNSA</i> <i>(PNSA1*TMSA/1000) [Zefirov's PC]</i>	D <sub>6</sub>	
8	$3.3623 \times 10^4$	$1.0804 \times 10^4$	3.112	<i>Intercept</i>		$R^2 = 0.9981, R_{cv}^2 =$ $0.9970, F = 1388.9$ $s^2 = 540.03$
	$7.8727 \times 10^4$	$2.3977 \times 10^3$	32.835	<i>Avg electroph. react. index for a C atom</i>	D <sub>1</sub>	
9	$3.5867 \times 10^3$	$3.9771 \times 10^2$	8.710	<i>Min nucleoph. react. index for a S atom</i>	D <sub>2</sub>	$R^2 = 0.9981, R_{cv}^2 =$ $0.9970, F = 1388.9$ $s^2 = 540.03$
	$2.8352 \times 10^4$	$2.2118 \times 10^3$	12.818	<i>Moment of inertia C</i>	D <sub>5</sub>	
10	3.5649	0.3701	9.632	<i>WNSA-1 Weighted PNSA</i> <i>(PNSA1*TMSA/1000) [Zefirov's PC]</i>	D <sub>6</sub>	$R^2 = 0.9981, R_{cv}^2 =$ $0.9970, F = 1388.9$ $s^2 = 540.03$
	$-3.4620 \times 10^4$	$1.0867 \times 10^4$	-3.1858	<i>Max SIGMA-SIGMA bond order</i>	D <sub>7</sub>	

**Table 2.** The 15-Lipoxygenase inhibitory activity (IC<sub>50</sub>) (μM)<sup>2</sup> and computed descriptors for the organosulfurs in scheme 1 according to the QSAR models of eqs. 4 and 5

Sys	IC <sub>50</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>	D <sub>7</sub>
1	1800	1.760E-02	1.008E-01	2.620E-02	2.690E-02	2.650E-02	34.6989	0.995
2	860	8.254E-03	2.750E-02	2.020E-02	3.230E-02	3.160E-02	41.8827	0.9982
3	425	6.153E-03	3.450E-02	1.340E-02	1.870E-02	1.870E-02	59.8165	0.998
4	289	3.624E-03	4.150E-02	9.725E-03	2.350E-02	1.560E-02	64.8574	0.9948
5	8	4.329E-03	2.290E-02	1.670E-02	3.475E-03	3.239E-03	85.9291	0.995
6	25	4.266E-03	4.060E-03	9.267E-03	4.289E-03	4.217E-03	112.9561	0.9952
7	29	5.054E-03	2.555E-04	2.060E-02	4.506E-03	4.433E-03	95.7687	0.9951
8	37	4.102E-03	3.889E-03	1.650E-02	5.652E-03	5.181E-03	97.1850	0.9952
9	31	5.023E-03	1.170E-03	2.040E-02	6.378E-03	6.075E-03	82.5075	0.9951
10	20	2.928E-03	3.202E-03	6.396E-03	3.999E-03	3.622E-03	128.7944	0.9951
11	20	2.829E-03	2.190E-02	9.404E-03	4.907E-03	4.627E-03	124.0813	0.9951
12	2	3.304E-03	8.970E-04	9.626E-03	1.995E-03	1.912E-03	138.0794	0.995
13	90	4.462E-03	3.299E-03	1.670E-02	4.702E-03	4.582E-03	123.9404	0.9952
14	10	3.946E-03	9.809E-03	8.837E-03	3.378E-03	3.348E-03	118.7012	0.9951
15	650	1.250E-02	1.896E-03	4.840E-02	1.090E-02	9.970E-03	52.2646	0.995
16	650	1.280E-02	9.280E-05	4.970E-02	1.060E-02	9.743E-03	56.8963	0.995
17	29	4.735E-03	6.952E-04	1.950E-02	7.261E-03	6.783E-03	78.9253	0.9951
18	37	5.108E-03	1.879E-04	2.060E-02	7.853E-03	7.556E-03	76.9630	0.9951
19	51	4.511E-03	3.034E-03	1.880E-02	5.652E-03	5.226E-03	98.1743	0.9949
20	-	1.174E-03	8.330E-02	2.483E-03	9.098E-03	8.951E-03	102.1879	0.995
21	-	6.528E-03	2.310E-02	9.714E-03	7.826E-03	7.575E-03	82.5828	0.9978
22	-	3.213E-03	1.822E-03	4.459E-03	7.518E-03	7.480E-03	84.0023	0.998
23	-	3.502E-03	8.290E-03	3.529E-03	3.910E-02	2.600E-02	59.7223	0.9946
24	-	2.815E-03	4.855E-03	2.815E-03	3.380E-02	2.020E-02	80.6155	0.9962
25	-	1.896E-03	1.600E-02	4.048E-03	2.160E-02	1.730E-02	70.6952	0.9982
26	-	2.508E-03	8.263E-03	2.994E-03	3.410E-02	2.030E-02	60.4810	0.996
27	-	3.273E-03	6.591E-03	6.521E-03	2.240E-02	1.780E-02	64.9810	0.9979
28	-	4.346E-03	1.354E-03	8.342E-03	9.078E-03	5.816E-03	118.6309	0.9977
29	-	4.088E-03	1.040E-02	1.040E-02	1.070E-02	6.488E-03	58.7726	0.9986
30	-	3.685E-03	9.776E-03	9.962E-03	1.050E-02	6.380E-03	58.8086	0.9986

*D*<sub>1</sub>: Avg electroph. react. index for a C atom, *D*<sub>2</sub>: Min nucleoph. react. index for a S atom, *D*<sub>3</sub>: Max electroph. react. index for a C atom, *D*<sub>4</sub>: Moment of inertia B, *D*<sub>5</sub>: Moment of inertia C, *D*<sub>6</sub>: WNSA-1 Weighted PNSA (PNSA1\*TMSA/1000) [Zefirov's PC], *D*<sub>7</sub>: Max SIGMA-SIGMA bond order

Among the obtained three-parameter models, statistically the best one is as follows:

$$\begin{aligned}
 \text{IC}_{50} = & (-335.61 \pm 23.508) + (1.2004 \cdot 10^5 \pm 5937.0) \cdot D_1 - (14034 \pm 1669.6) \cdot D_3 + \\
 & (1.4138 \cdot 10^4 \pm 1.7460 \cdot 10^3) \cdot D_4 \quad (3) \\
 N = & 19; R^2 = 0.9905; R^2_{cv} = 0.9850; F = 522.27; s^2 = 2375.3
 \end{aligned}$$

In this model, *D*<sub>3</sub> is the quantum chemical descriptor measures the maximum electrophilic reaction index for a C atom which replaces *D*<sub>2</sub> in the previous model in eq. 2. Similar to *D*<sub>1</sub> and different from *D*<sub>3</sub>, the geometrical descriptor of moment of inertia B (*D*<sub>4</sub>) has a direct proportional with the numerical values of IC<sub>50</sub>. Very close to the significance of *D*<sub>4</sub>, *D*<sub>3</sub> is the second least significant in this model compared with *D*<sub>1</sub>.

Among the obtained four-parameter models, the best one is as shown in eq. 4 and plotted in Figure 1, where *D*<sub>5</sub> is moment of inertia C, which replaces the B component of the moment in eq. 3 and *D*<sub>6</sub> is the electrostatic descriptor WNSA-1 Weighted PNSA (PNSA1\*TMSA/1000) [Zefirov's PC], which refers to partial negative surface area.

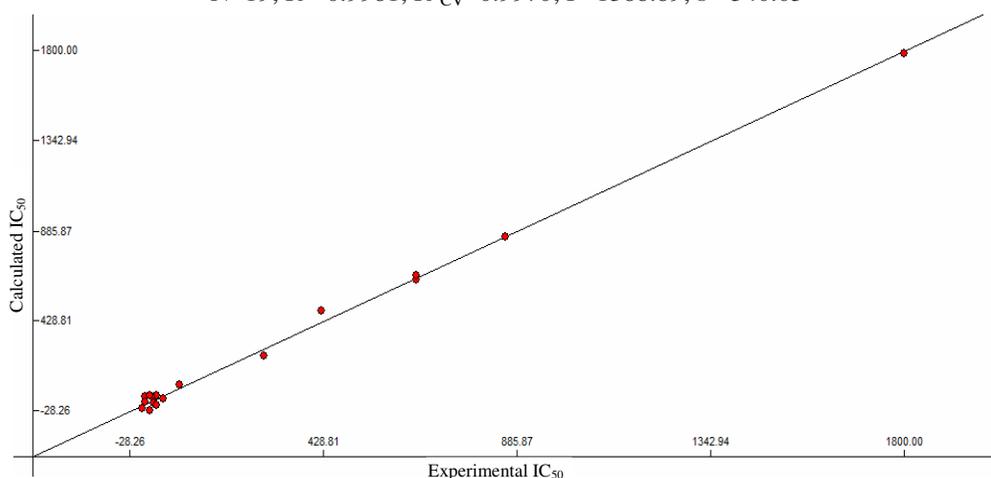
$$\begin{aligned}
 IC_{50} = & (-796.37 \pm 61.719) + (8.1698 * 10^4 \pm 2840.4) * D_1 - (4256.6 \pm 434.08) * D_2 + \\
 & (2.2814 * 10^4 \pm 1.7587 * 10^3) * D_5 + (3.3589 \pm 46.860) * D_6 \\
 N = 19; R^2 = & 0.9967; R^2_{cv} = 0.9933; F=1048.42; s^2 = 893.0
 \end{aligned}
 \tag{4}$$

In this model, all the descriptors have their positive-signed coefficients which highlight direct proportional relations with the magnitude of  $IC_{50}$ . Implying a little impact of the S atom in the activity, the significance of the descriptor in this model is arranged as  $D_1 \gg D_5 > D_2 > D_6$ .

The best five-parameter model was obtained in eq. (5) and plotted in Figure 2, where  $D_7$  is the quantum-chemical descriptor of Max SIGMA-SIGMA bond order, which has a negative impact on the value of  $IC_{50}$ . Going with the same previous trend, the significance of the descriptors here is arranged according to  $D_1 \gg D_5 > D_6 > D_2 > D_7$

$$\begin{aligned}
 IC_{50} = & (3.3623 * 10^4 \pm 1.0804 * 10^4) + (7.8727 * 10^4 \pm 2.3977 * 10^3) * D_1 + (3.5867 * 10^3 \pm \\
 & 397.71) * D_2 + (2.8352 * 10^4 \pm 2.2118 * 10^3) * D_5 + (3.5649 \pm 3.7011) * D_6 - (3.4620 * 10^4 \pm \\
 & 1.0867 * 10^4) * D_7
 \end{aligned}
 \tag{5}$$

$$N=19; R^2=0.9981; R^2_{cv}=0.9970; F=1388.89; s^2=540.03$$

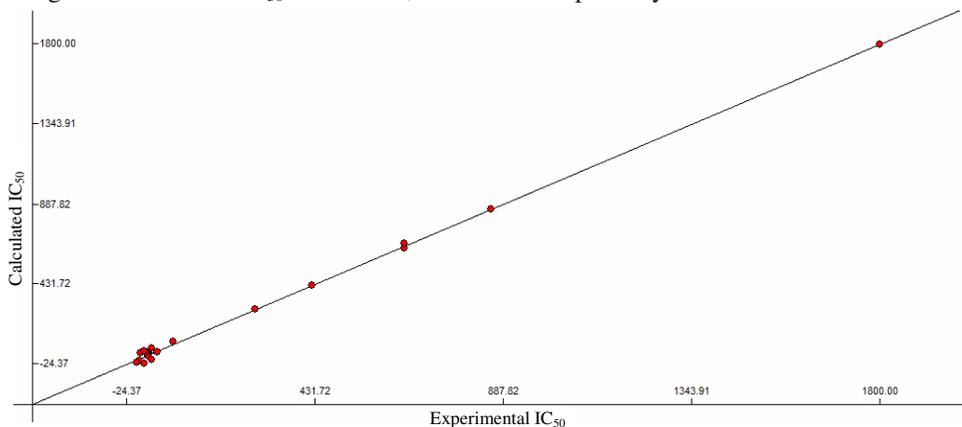


**Figure 1.** Comparison of experimental and calculated biological activity from the regression analysis in eq. (4).  $R^2 = 0.9967$ ;  $R^2_{cv} = 0.9933$ ;  $F=1048.4$ ;  $s^2 = 892.95$  for 19 organosulfurs.

Obviously, the significantly high value of the standard deviation of the regression in each of the five models may be attributed to the large gaps between the experimental data of the relatively small set of molecules (1-19), which ranges between 1800 and 2. However, the excellent squared cross-validity coefficients mostly in eqs. 4 and 5 reflect the predictive validity of the concluded models, where they provide more realistic assessment of the usefulness of the equations than the other parameters. In Table 3, the inter-correlation matrix of descriptors involved in eq. 4 is depicted.

Lastly, the four- and five-descriptor models in eqs. 4 and 5 are used to predict the biological activities of the last ten organosulfurs in scheme 1; (20-30). The experimental and computed values of  $IC_{50}$ , along with the predicted ones for the systems in this study are listed in Table 4. Obviously, according to the estimated  $IC_{50}$ , both diallyltrisulfide (4) and diallyltetrasulfide (20) from garlic have relatively high predicted values of  $IC_{50}$  (290-250, and 186-202, respectively), which support the long-term claim of garlic's wide range antimicrobial, antibacterial and antifungal activities. The compounds 1,3-bis (methylthiomethyl)trislpane (21) and 1,4-dimethyltetrasulfane (22) can be isolated

from *Allium* species including garlic and the bark of specific African trees which has a garlic distinct smell<sup>2</sup>. According to the models in eqs. 5 and 4, both compounds 21 and 22 have relatively high biological activities with  $IC_{50}$  are 186-285, and 207-236 respectively.



**Figure 2.** Comparison of experimental and calculated biological activity from the regression analysis in eq. (5).  $R^2 = 0.9981$ ;  $R^2_{cv} = 0.9970$ ;  $F=1388.89$ ;  $s^2 = 540.03$  for 19 organosulfurs

**Table 3.** Inter-correlation matrix of descriptors involved in the obtained QSAR model in eq. 4

	<i>D1</i>	<i>D2</i>	<i>D5</i>	<i>D6</i>
<i>D1</i>	1.0000	0.1550	0.2445	-0.4592
<i>D2</i>		1.0000	0.4360	-0.3309
<i>D5</i>			1.0000	-0.7302
<i>D6</i>				1.0000

**Table 4.** Experimental  $IC_{50}$  (1-20) ( $\mu M$ )<sup>2</sup> and estimated  $IC_{50}$  (1-30) for the organosulfurs in scheme 1 using QSAR models in eqs. 4 and 5

Sys	Exp $IC_{50}$	<sup>5</sup> cal $IC_{50}$	<sup>4</sup> cal $IC_{50}$	Sys	Exp $IC_{50}$	<sup>5</sup> cal $IC_{50}$	<sup>4</sup> cal $IC_{50}$
1	1800	1796.1	1789.0	16	650	661.9	661.0
2	860	859.9	856.1	17	29	20.4	13.3
3	425	424	479.7	18	37	0.6	-0.05
4	289	289.6	250.3	19	51	43.4	34.1
5	8	-3.7	17.2	20	-	186.2	201.7
6	25	43.9	45.0	21	-	186.5	285.4
7	29	37.8	40.4	22	-	207.4	235.9
8	37	63.0	52.6	23	-	445.2	319.5
9	31	37.8	34.7	24	-	234.9	185.5
10	20	-24.4	-28.3	25	-	16.7	59.6
11	20	48.5	50.4	26	-	162.8	111.0
12	2	-13.9	-15.2	27	-	95.6	124.4
13	90	105.3	103.0	28	-	-105.9	8.3
14	10	35.8	42.8	29	-	-52.7	82.1
15	650	636.8	636.6	30	-	17.6	95.6

<sup>5</sup>cal  $IC_{50}$ : calculated biological activities using model in eq. 5, <sup>4</sup>cal  $IC_{50}$ : calculated biological activities using model in eq. 4

The cyclic polysulfides (23) known as lenthionine, and hexathiepane (24) were isolated from Shiitake mushrooms, have even higher predicted  $IC_{50}$  than those of garlic ingredients; 445-320, 235-186 respectively in models 5 and 4. These findings may support the belief that (24) might generate a blend of sulfurs including  $S_2$ ,  $S_3$ ,  $S_4$  and  $S_5$ <sup>2</sup>.

1,2,3-Trithiane-5-carboxylic acid in (25), which is found in asparagus, showed a reasonable estimated biological activity of about 17 and 60 according to our last two models. 1,2,3,5,7-pentathiocane (26) from *Parkia speciosa* has an approved activity against specific leukemia cells in vitro ( $IC_{50}$   $13 \mu\text{g ml}^{-1}$ ) and could inhibit specific Gram-positive bacterium and fungus<sup>20</sup>. In our models, the data predicted for the compound's activity confirms its experimental efficiency with  $IC_{50}$  ranges between 163 and 111, which is very close to that of 5-(methylthio)-1,2,3-trithiane (27) known as alga *Chara globularis*; ( $IC_{50}$  = 96 -124).

Naturally isolated cyclic polysulfides with significant therapeutic potential are lissoclinotoxin A (28) and Polycitor sp. (varacins B–C (29 and 30), which have common activity against *C. albicans* and *B. subtilis*<sup>2</sup>. According to the data in Table 4 and the QSAR model summarized in eq. 4, the three compounds are moderately biologically active with their  $IC_{50}$  equal 8, 82, and 96 respectively.

## Conclusion

In the present QSAR investigations, the descriptors of two sets of 30 biologically active organosulfurs have been correlated with their activity, ( $ID_{50}$ ) using the AM1/B3LYP/6-31G\* level of theory in G09W and CODESSA package. The concluded QSAR models reflected high efficiency in predicting the biological activity data for other analogous organosulfur ingredients of widely consumed natural products such as garlic and mushroom. According to these models, the motivations behind the biological activity of organosulfurs are mainly quantum chemical while S atoms themselves have a relatively little significance in this regard. In relative basis, the predicted biological activity of garlic's organosulfur ingredients is high and thus assures vital medical role of garlic as a traditional medications.

## References

1. Rabinowitch H D and Currah L, Eds., Health and alliums. In *Allium Crop Science: Recent Advances*; *CAB International*: Wallingford, U.K., 2002, 515.
2. Jacob C, *Nat Prod Rep.*, 2006, **23**, 851.
3. Edris A E, *Phytother Res.*, 2007; **21(4)**, 308.
4. Tapiero H, Townsend D M and Tew K D, *Biomed Pharmacother*, 2004, **58(3)**, 183-193.
5. Arzanlou M and Bohlooli S, *Food Chemistry*, 2010, **120(1)**, 179-183.
6. Butt M S A, Sultan M T and *et al*, *Critical Rev Food Sci and Nut*, 2009, **49(6)**, 538.
7. Iciek M, Kwiecien I and Włodek L., *Environ Molec Mutagenesis*, 2009, **50(3)**, 247-265.
8. Amagase, H, *J Nutrition*, 2006, **136**, 716S-725S.
9. Kraus D W and *et al*, *Proc Natl Acad Sci.*, 2007, **104(46)**, 17977-17982.
10. Rivlin R S, *J Nutrition*, 2001, **131(3s)**, 951S-954S.
11. Kandil O M, Abdellah T H and Elkadi A, *Fed Proc.*, 1987, **46**, 441.
12. Abu-Awwad F M, *Int J Chem Tech Res.*, 2009, **1(3)**, 742.
13. Abu-Awwad F M, *Der Pharma Chemica.*, 2010, **2(1)**, 1-13.
14. Camargo A B, Marchevsky E and Luco J M, *J Agric Food Chem.*, 2007, **55**, 3096-3103
15. Brash A R, *J Biol Chem.*, 1999, **274**, 23679-23682.
16. CODESSA<sup>TM</sup>, V.2.7.10 (PC), Semichem, 7204 Mullen, Shawnee, KS, USA, Copyright Semichem and the University of Florida, 2002.
17. Frisch M J and *et al*, Gaussian 03 W, (Rev 6.1), Gaussian Inc., Pittsburgh, PA, USA. 2003.
18. Becke A D, *J Chem Phys.*, 1993, **98**, 5648.
19. Lee C, Yang W and Parr R G, *Phys Rev.*, 1988, **B37**, 785.
20. Copp B R, Blunt J W, Munro M H G and Pannell L K, *Tetrahedron Lett.*, 1989, **30**, 3703.

