

## **A QSAR Study on Antimicrobial Activity of Some New Sulfonylhydrazinotiazoles**

**Brindusa Tiperciuc,<sup>a</sup> V. Zaharia,<sup>a</sup> R. Campean,<sup>a</sup> Manuela Curticapean,<sup>b</sup>  
Adina Costescu<sup>c</sup> and M. V. Diudea<sup>c</sup>**

<sup>a</sup>University of Medicine and Pharmacy "Iuliu Hatieganu" Cluj-Napoca, Romania

<sup>b</sup> University of Medicine and Pharmacy Tg-Mures, Romania

<sup>c</sup> "Babes-Bolyai" University, Faculty of Chemistry and Chemical Engineering,  
Cluj, 400028, Romania

(Received May 12, 2008)

### **Abstracts**

A quantitative structure-activity relationship QSAR study on antimicrobial activity of a series of newly synthesized sulfonylhydrazinotiazoles was performed by using the Free-Wilson and auto-correlating partial charges approaches. Several models were developed using stepwise multiple linear regression analysis and the results were further validated by leave-one-out method. They showed the most important contributors at the antimicrobial activity of sulfonylhydrazinotiazoles are the substituents on 4 and 5 position at the thiazole ring. Also, the main feature describing the activity of these compounds is a partial charge-based descriptor, a measure of the molecular electronic properties, used within an auto-correlation weighting scheme.

## Introduction

Thiazole nucleus is known to be present in various molecules having a biological activity. They display antitumor, anti-inflammatory, antifungal and antibacterial activity against both Gram-positive and Gram-negative bacteria [1-3]. Compounds having sulfonylhydrazine moiety are also known to possess a wide range of biological and pharmacological activity: antimicrobial [4-5], antitumor [5-7], analgesic [8], anti-inflammatory, antipyretic [4]. Also thiosemicarbazides and hydrazinethiazoles with IMAO activity [9] and arylidenehydrazinethiazoles with antimicrobial and anti-inflammatory potential [10], have been reported.

Motivated by these facts and in pursuing our research in the field of synthesis and antimicrobial evaluation of heterocyclic compounds with thiazolic nucleus, we aimed to develop a QSAR model and to explain the antibacterial and antifungal activity of some newly synthesized benzensulfonylhydrazinethiazoles, for which antimicrobial activity has been evaluated experimentally [11].

To obtain additional information on structural requirements necessary for antimicrobial activity, we performed both a Free-Wilson and auto-correlating partial charges approaches on a set of 14 compounds [12, 13]. The data set includes substances with variations on the position 4 and 5 of the thiazolic ring and the hydrazine moiety.

## Materials and Methods

The in vitro minimum inhibitory concentration (MIC) in  $\mu\text{M}/\text{ml}$  required for inhibiting the growth of *Bacillus subtilis*, *Citrobacter*, *Escherichia coli* and *Candida albicans* was the property studied [11]. The  $\log(1/\text{MIC})$  is being used as biological response, for QSAR analysis (Table 2).

The QSAR analysis consists of the following steps: (i) structure optimization by using semiempirical method PM3; (ii) calculation of molecular descriptors; (iii) correlation analysis by step-forward selection of descriptors; (iv) evaluation of the significance level of the model; (v) validation of the model (leave-one-out *loo* cross-validation procedure); (vi) interpretation of the model.

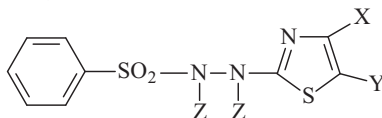
Most of the applications of molecular descriptors have been dedicated to QSAR studies because of the great importance for biology of the structure-activity relationship [15]. The computation of such descriptors is accessible by using available software products. The complete set of molecular descriptors described in this study (some of them will be defined below), was calculated by Dragon program package [16].

The structures were optimized by using the semiempirical PM3 Hamiltonian, available in HyperChem.

In order to build the regression models, we tried to fit a linear function, which quality was estimated by the squared correlation coefficient ( $R^2$ ), the standard error of estimate (s), the Fischer ratio ( $F$ ), chance statistics lower than 0.01 ( $p < 0.01$ ) and the coefficient of variance (CV%). The predicting ability of each model was estimated by the cross-validated squared correlation coefficient ( $R^2_{cv}$ ), calculated by LOO method [14].

### Free-Wilson Analysis

The chemical structure of all compounds studied and the descriptors for the Free-Wilson analysis are given in Figure 1.



X: X<sub>1</sub>-X<sub>3</sub>, Y: Y<sub>1</sub>-X<sub>4</sub> Z: Z<sub>1</sub>-Z<sub>2</sub>

|                | X                                |                | Y                                |                | Z                 |
|----------------|----------------------------------|----------------|----------------------------------|----------------|-------------------|
| X <sub>1</sub> | CH <sub>3</sub>                  | Y <sub>1</sub> | H                                | Z <sub>1</sub> | H                 |
| X <sub>2</sub> | CH <sub>2</sub> Cl               | Y <sub>2</sub> | COCH <sub>3</sub>                | Z <sub>2</sub> | COCH <sub>3</sub> |
| X <sub>3</sub> | C <sub>6</sub> H <sub>5</sub> Cl | Y <sub>3</sub> | COOC <sub>2</sub> H <sub>5</sub> |                |                   |
|                |                                  | Y <sub>4</sub> | Br                               |                |                   |

**Figure 1.** General structure of 2-(2-benzensulfonyl-1,2-di-Z-hydrazino)-4X,5Y-thiazole antimicrobials considered.

Free and Wilson's model is based on the assumption that each substituent makes an additive and constant contribution to the biological activity regardless of substituent variation in the rest of molecule. The values of individual contributions are calculated by regression analysis and the constant term obtained, is a theoretically predicted activity value of the unsubstituted compound (all R = H) [12].

The code for each compound and the matrix used in the Free-Wilson analysis are given in Table 1.

**Table 1.** The codes for S1-S14 compounds and the Free-Wilson matrix

| Comp | Code   | X              |                |                | Y              |                |                |                | Z              |                |
|------|--|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|      |  | X <sub>1</sub> | X <sub>2</sub> | X <sub>3</sub> | Y <sub>1</sub> | Y <sub>2</sub> | Y <sub>3</sub> | Y <sub>4</sub> | Z <sub>1</sub> | Z <sub>2</sub> |
| S1   | X <sub>1</sub> Y <sub>1</sub> Z <sub>1</sub> | 1              |                |                | 1              |                |                |                | 1              |                |
| S2   | X <sub>2</sub> Y <sub>1</sub> Z <sub>1</sub> |                | 1              |                | 1              |                |                |                | 1              |                |
| S3   | X <sub>3</sub> Y <sub>1</sub> Z <sub>1</sub> |                |                | 1              | 1              |                |                |                | 1              |                |
| S4   | X <sub>1</sub> Y <sub>2</sub> Z <sub>1</sub> | 1              |                |                |                | 1              |                |                | 1              |                |
| S5   | X <sub>1</sub> Y <sub>3</sub> Z <sub>1</sub> | 1              |                |                |                |                | 1              |                | 1              |                |
| S6   | X <sub>1</sub> Y <sub>1</sub> Z <sub>2</sub> | 1              |                |                | 1              |                |                |                |                | 1              |
| S7   | X <sub>2</sub> Y <sub>1</sub> Z <sub>2</sub> |                | 1              |                | 1              |                |                |                |                | 1              |
| S8   | X <sub>3</sub> Y <sub>1</sub> Z <sub>2</sub> |                |                | 1              | 1              |                |                |                |                | 1              |
| S9   | X <sub>1</sub> Y <sub>2</sub> Z <sub>2</sub> | 1              |                |                |                | 1              |                |                |                | 1              |
| S10  | X <sub>1</sub> Y <sub>3</sub> Z <sub>2</sub> | 1              |                |                |                |                | 1              |                |                | 1              |
| S11  | X <sub>1</sub> Y <sub>4</sub> Z <sub>1</sub> | 1              |                |                |                |                |                | 1              | 1              |                |
| S12  | X <sub>3</sub> Y <sub>4</sub> Z <sub>1</sub> |                |                | 1              |                |                |                | 1              | 1              |                |
| S13  | X <sub>1</sub> Y <sub>4</sub> Z <sub>2</sub> | 1              |                |                |                |                |                | 1              |                | 1              |
| S14  | X <sub>3</sub> Y <sub>4</sub> Z <sub>2</sub> |                |                | 1              |                |                |                | 1              |                | 1              |

The matrix was solved by multiple linear regression analysis, using the Excel software package. The 95% confidence interval is given for each regression coefficient.

The correlation was sought between inhibitory activity and various substituent at position 4 (X<sub>1,3</sub>) and 5 (Y<sub>1,4</sub>) of the thiazolic ring and for the acetyl group on the hydrazinic moiety (Z<sub>1,2</sub>).

Leave-one-out analysis was performed, in view of testing the predicting ability of the regression equation. The question of outliers was addressed for points which do not fall within a specified error limits (standard residual >2×s). These compounds were not included in the further analysis. With these outliers removed, we observed an improvement in correlation with the same descriptors.

### Auto-correlating Partial Charges Analysis

A QSAR method is based on the comparison of a measured and calculated molecular activity and then relating a few of the most informative structural descriptors to the target bioactivity. The quantitative structure-activity relationships constructed this way provide a means of investigating and predicting antimicrobial activities.

The subset of electronic parameters includes molecular descriptors on partial charges.

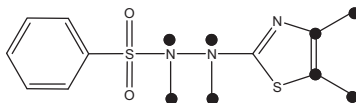
Within TOPOCLUJ program, the partial charges  $Ch_i$  are calculated as follows [17]:

$$Ch_{i,j} = \log(S_j / S_i)^{1/(d_{i,j})^2} \quad (1)$$

$$Ch_i = \sum_j c_j ch_{i,j} \quad (2)$$

In the above relations,  $S_i$ ,  $S_j$  represent the Sanderson group electronegativities, calculated for the hydride groups (*i.e.*, the heavy atoms with their surrounding hydrogen atoms) in the molecule and  $d_{ij}$  is the Euclidean distance separating atoms  $i$  and  $j$  in a minimal energy optimized chemical structure (HyperChem). For other topological partial charge calculations see refs. [18, 19]. Any sulfonylhydrazinothiazole compound can be described by these partial charges which characterize both the substituted/unsubstituted aromatic positions and the heteroatom (nitrogen).

The partial charges are calculated for the positions marked by black points in Figure 2.



**Figure 2.** Partial charges selected for general structure of 2-(2-benzensulfonyl-1,2-di-Z-hydrazino)-4X,5Y-thiazole.

On this ground, a flexible global descriptor ( $CD$ ), can be defined as an additive function of correlation weights of the partial charges corresponding to each atom  $i$ :

$$CD_i = \sum_j c_j \cdot Ch_{i,j} \quad (3)$$

where  $c_j$  represents the regression coefficient (*i.e.*, the correlation weight) as given by the multivariate regression  $\log(A_{i,\text{exp}}) = f(Ch_i)$  (see Table 2). These “ad-hoc” weightings depend on the set of molecules in work as well as on the considered molecular (or local) property.

The *Dragon 5.4* software was used to calculate a total of 1600 molecular descriptors, for each of the studied compounds. The descriptors choose for the analysis are either: 1) *getaway* descriptors; 2) *whim* descriptors and 3) *2D autocorrelation* indices.

Ghetaway descriptors [16] are based on a leverage matrix similar to the defined in statistics and usually used for regression diagnostics. These molecular descriptors try to match 3D-molecular geometry provided by the molecular influence matrix and the atom relatedness

by molecular topology, with chemical information by using various atomic weights (atomic mass, polarizability, van der Waals volume and electronegativity, etc.).

The two descriptors which belong to this group are: R8m (R autocorrelation of lag 8/weighted by atomic masses) and ISH (standardized information content on the leverage equality).

WHIM [16] descriptors are 3-D descriptors based on the calculation of principal component axes calculated from a weighted covariance matrix obtained by the molecular geometric coordinates. Six different weighting schemes are used for the weighted covariance matrix: u (unweighted), m (atomic mass), p (atomic polarizability), v (van der Waals volume), e (atomic electronegativity) and s (atomic electrotopological state). They contain chemical information concerning: size, symmetry, shape and distribution of the molecule atoms.

WHIM descriptors used in our study are: G1e (1<sup>st</sup> component symmetry directional WHIM index/weighted by atomic Sanderson electronegativities), E2s (2<sup>nd</sup> component accessibility directional WHIM index/weighted by atomic electrotopological state) and E1m (1<sup>st</sup> component accessibility directional WHIM index/weighted by atomic masses).

The next group of descriptors is based on 2-D autocorrelation [16] functions applied to a molecular graph, which is a 2-dimensional structural representation of a molecule. This class of descriptors expresses a correlation between numerical values of the graph entries, which can be statistically weighted using various atomic properties, at intervals equal to the given lag value.

From this class of descriptors, the following indices gave good results: AT8e (Broto-Moreau autocorrelation of a topological structure-lag 8/weighted by atomic Sanderson electronegativities) and GATS4m (Geary autocorrelation-lag 4/weighted by atomic masses).

## Results and discussions

### Free-Wilson Analysis

The Free-Wilson equations obtained describes the pharmacological activity of the compounds in a pretty good way ( $R^2=0.626-0.831$ ). Table 2 presents the observed and calculated values of  $\log(1/MIC)$  by the best obtained models.

The best model of antibacterial activity in **Bacillus subtilis**, with the coefficient of correlation  $R=0.886$ , is that given in eq 4 (see also Table 2). The data contain no outliers.

$$\text{Log}(1/\text{MIC}) = 1.35 - 0.091 \times X_1 + 0.11 \times X_2 - 0.375 \times Y_1 - 0.162 \times Y_2 - 0.277 \times Y_3 - 0.358 \times Z_1 \quad (4)$$

n=14, R=0.886, R<sup>2</sup>=0.785; s=0.174, F=4.274, R<sup>2</sup><sub>cv</sub>=0.507

**Table 2.** Observed and calculated log(1/MIC) and the squared correlation coefficient R<sup>2</sup> for *Bacillus subtilis*, *Citrobacter*, *Escherichia coli* and *Candida albicans*.

| No. | <i>Bacillus subtilis</i><br>(log(1/MIC))     |          | <i>Citrobacter</i><br>(log(1/MIC))  |       |            | <i>Escherichia coli</i><br>(log(1/MIC))      |            | <i>Candida albicans</i><br>(log(1/MIC))  |       |            |
|-----|--|----------|---|-------|------------|--|------------|--|-------|------------|
|     | Obs.   | CalcEq 4 | Obs.  | Calc. | Calc. Eq 5 | Obs.   | Calc. Eq 6 | Obs.   | Calc. | Calc. Eq 7 |
| S1  | 0.652  | 0.526    | 0.652   | 0.682 | 0.632      | 1.430  | 1.380      | 1.731  | 1.79  | 1.622      |
| S2  | 0.703  | 0.727    | 0.703   | 0.871 | 0.937      | 1.481  | 1.581      | 1.481  | 1.728 | 1.560      |
| S3  | 0.742  | 0.616    | 0.742   | 0.447 | 0.629      | 1.520  | 1.621      | 1.821  | 1.955 | 1.788      |
| S4  | 0.715  | 0.737    | 1.016   | 1.032 | 1.098      | 1.493  | 1.644      | 2.396  | 1.388 | -          |
| S5  | 0.755  | 0.623    | 1.357   | 1.068 | 1.134      | 1.533  | 1.085      | 2.436  | 2.376 | 2.208      |
| S6  | 0.77   | 0.884    | 0.770   | 1.054 | 0.873      | 1.548  | 1.781      | 1.849  | 2.005 | 2.173      |
| S7  | 1.111  | 1.086    | 1.412   | 1.244 | 1.178      | 1.889  | 1.983      | 2.19   | 1.944 | 2.111      |
| S8  | 0.84   | 0.975    | 0.840   | 0.820 | 0.870      | 1.919  | 2.023      | 2.521  | 2.171 | 2.339      |
| S9  | 1.12   | 1.096    | 1.421   | 1.404 | 1.339      | 2.199  | 2.046      | 0.597  | 1.604 | -          |
| S10 | 0.850  | 0.982    | 1.151   | 1.440 | 1.374      | 1.628  | 1.487      | 2.532  | 2.592 | 2.760      |
| S11 | 0.723  | 0.901    | 1.325   | 1.183 | 1.364      | 1.501  | 1.506      | 1.501  | 1.788 | 1.620      |
| S12 | 0.835  | 0.991    | 0.437   | 0.948 | -          | 1.613  | 1.747      | 1.613  | 1.954 | 1.786      |
| S13 | 1.427  | 1.259    | 1.728   | 1.555 | 1.605      | 1.904  | 1.908      | 2.506  | 2.004 | 2.172      |
| S14 | 1.518  | 1.350    | 1.518   | 1.321 | 1.602      | 2.296  | 2.149      | 2.296  | 2.17  | 2.338      |
|     | R <sup>2</sup> =0.785<br>n=14<br>no outliers |          | R <sup>2</sup> =0.626<br>n=14<br>R <sup>2</sup> =0.825<br>n=13<br>outlier:<br>S12 |       |            | R <sup>2</sup> =0.831<br>n=14<br>no outliers |            | R <sup>2</sup> =0.306<br>n=14<br>R <sup>2</sup> =0.776<br>n=12<br>outliers:<br>S4 and S9 |       |            |

In case of antibacterial activity on **Citrobacter**, for all data (n = 14), correlation explains about 62% of the variance in inhibitory activity (Table 2). S12 is an outlier and with this outlier removed, an improvement in correlation with the same descriptors (82.5%, eq 5) was observed.

$$\text{Log}(1/\text{MIC}) = 1.601 + 0.003 \times X_1 + 0.308 \times X_2 - 0.731 \times Y_1 - 0.266 \times Y_2 - 0.230 \times Y_3 - 0.240 \times Z_1 \quad (5)$$

n=13, R=0.908, R<sup>2</sup>=0.825; s=0.19, F=6.60, R<sup>2</sup><sub>cv</sub>=0.752

For **Escherichia coli**, correlation (R=0.912) is given in eq 6, which explains about 83% of the variance in inhibitory activity (see also Table 2). The data contain no outliers.

$$\text{Log}(1/\text{MIC}) = 2.149 - 0.241250 \times X_1 - 0.039 \times X_2 - 0.224 \times Y_1 + 0.138 \times Y_2 - 0.127 \times Y_3 - 0.401 \times Z_1 \quad (6)$$

n=14, R=0.912, R<sup>2</sup>=0.831, s=0.157, F=5.75, R<sup>2</sup><sub>cv</sub>=0.753

Finally, the statistically low correlation, with R<sup>2</sup> = 0.306, was considered a poor model for antifungal activity in **Candida albicans** (Table 2). After the elimination of two outliers, compounds: S4 and S9, which are not properly predicted by the model, the squared correlation coefficient increased at R<sup>2</sup> = 0.776 (eq 7).

$$\text{Log}(1/\text{MIC}) = 2.337 - 0.166 \times X_1 - 0.228 \times X_2 + 0.001 \times Y_1 + 0.588 \times Y_3 - 0.551 \times Z_1 \quad (7)$$

n=12, R=0.881, R<sup>2</sup>=0.776, s=0.26, F=4.165, R<sup>2</sup><sub>cv</sub>=0.71

The descriptors which take into account the presence or the absence of the substituents in the selected positions are more significant in case of Citrobacter and Escherichia coli (82.5% and 83.1% respectively). Predicting abilities for Bacillus subtilis and Candida albicans are satisfactory (78.5% and 77.6%, respectively - Table 2)

### Auto-correlating Partial Charges Analysis

The regression analysis was used to determine which of the available molecular descriptors were most relevant in modeling of some sulfonylhydrazinotiazoles antimicrobial activities. Several descriptors were selected in this respect. Definitely, the electronic descriptor (CD) is the best predictor in a monovariate regression (Table 3). All-together, this approach is superior to the Free-Wilson analysis, as can be seen from the results given below.

**Table 3.** Observed and calculated log(1/MIC), squared correlation coefficient R<sup>2</sup> and CV% in monovariate regression, for Bacillus subtilis, Citrobacter, Escherichia coli and Candida albicans.

| No. | Bacillus subtilis<br>(log(1/MIC)) |             | Citrobacter<br>(log(1/MIC)) |             | Escherichia coli<br>(log(1/MIC)) |             | Candida albicans<br>(log(1/MIC)) |             |
|-----|-----------------------------------|-------------|-----------------------------|-------------|----------------------------------|-------------|----------------------------------|-------------|
|     | obs.                              | calc.=f(CD) | obs.                        | calc.=f(CD) | obs.                             | calc.=f(CD) | obs.                             | calc.=f(CD) |
| 1   | 0.652                             | 0.671       | 0.652                       | 0.405       | 1.430                            | 1.457       | 1.731                            | 1.714       |
| 2   | 0.703                             | 0.505       | 0.703                       | 0.669       | 1.481                            | 1.370       | 1.481                            | 1.377       |
| 3   | 0.742                             | 0.724       | 0.742                       | 0.760       | 1.520                            | 1.523       | 1.821                            | 1.813       |
| 4   | 0.715                             | 0.820       | 1.016                       | 1.267       | 1.493                            | 1.482       | 2.396                            | 2.331       |
| 5   | 0.755                             | 0.699       | 1.357                       | 1.165       | 1.533                            | 1.567       | 2.436                            | 2.492       |
| 6   | 0.770                             | 1.055       | 0.770                       | 1.195       | 1.548                            | 1.862       | 1.849                            | 2.138       |
| 7   | 1.111                             | 0.966       | 1.412                       | 1.223       | 1.889                            | 1.727       | 2.190                            | 2.287       |
| 8   | 0.840                             | 1.016       | 0.840                       | 0.915       | 1.919                            | 1.961       | 2.521                            | 2.249       |
| 9   | 1.120                             | 1.115       | 1.421                       | 1.408       | 2.199                            | 2.196       | 0.597                            | 0.603       |
| 10  | 0.850                             | 0.868       | 1.151                       | 1.184       | 1.628                            | 1.632       | 2.532                            | 2.522       |
| 11  | 0.723                             | 0.865       | 1.325                       | 1.263       | 1.501                            | 1.535       | 1.501                            | 1.473       |
| 12  | 0.835                             | 0.840       | 0.437                       | 0.700       | 1.613                            | 1.638       | 1.613                            | 1.782       |
| 13  | 1.427                             | 1.344       | 1.728                       | 1.627       | 1.904                            | 1.938       | 2.506                            | 2.360       |
| 14  | 1.518                             | 1.274       | 1.518                       | 1.291       | 2.296                            | 2.067       | 2.296                            | 2.329       |
|     | R <sup>2</sup> = 0.723            |             | R <sup>2</sup> = 0.738      |             | R <sup>2</sup> = 0.81            |             | R <sup>2</sup> = 0.94            |             |
|     | CV%=12.23                         |             | CV%=17.9                    |             | CV%=4.28                         |             | CV%=4.66                         |             |



**Bacillus subtilis**

First molecule (S1) seems to be outlier and it was excluded from the analysis.

**Bivariate regression**

$$\text{Log}(1/\text{MIC}) = -0.240 + 0.675 \times \text{CD} + 1.573 \times \text{R8m} \quad (8)$$

n=13; R<sup>2</sup>=0.908; s=0.10; F=49.51; p=6.49E-06

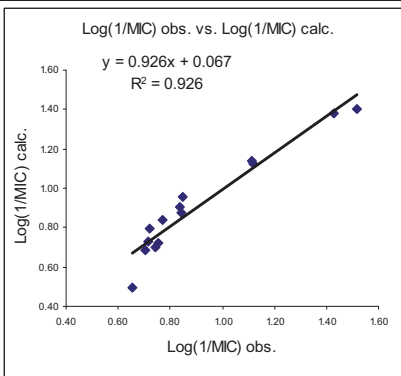
**Trivariate regression**

$$\text{Log}(1/\text{MIC}) = 0.722 + 0.629 \times \text{CD} + 1.615 \times \text{R8m} - 5.471 \times \text{G1e} \quad (9)$$

n=13; R<sup>2</sup>=0.962; s=0.06; F=77.0; p=1E-06; R<sup>2</sup><sub>cv</sub>=0.934

**Table 4.** Observed and calculated log(1/MIC), CV% in trivariate regression for **Bacillus subtilis**

|     | <b>Bacillus subtilis</b> (log(1/MIC)) |       |        | CV%           |
|-----|---------------------------------------|-------|--------|---------------|
|     | Obs                                   | Calc. | Rez.   |               |
| S2  | 0.703                                 | 0.682 | 0.021  | 2.987         |
| S3  | 0.742                                 | 0.701 | 0.041  | 5.530         |
| S4  | 0.715                                 | 0.733 | -0.018 | 2.497         |
| S5  | 0.755                                 | 0.726 | 0.029  | 3.903         |
| S6  | 0.770                                 | 0.842 | -0.072 | 9.298         |
| S7  | 1.111                                 | 1.140 | -0.029 | 2.608         |
| S8  | 0.840                                 | 0.879 | -0.039 | 4.658         |
| S9  | 1.120                                 | 1.127 | -0.007 | 0.646         |
| S10 | 0.850                                 | 0.957 | -0.107 | 12.639        |
| S11 | 0.723                                 | 0.793 | -0.070 | 9.726         |
| S12 | 0.835                                 | 0.906 | -0.071 | 8.474         |
| S13 | 1.427                                 | 1.377 | 0.050  | 3.471         |
| S14 | 1.518                                 | 1.403 | 0.115  | 7.563         |
|     |                                       |       |        | <b>7.010*</b> |



\*CV% average

**Citrobacter**

**Bivariate regression**

$$\text{Log}(1/\text{MIC}) = 0.470 + 0.960 \times \text{CD} - 0.531 \times \text{GATS4m} \quad (10)$$

n=14; R<sup>2</sup>=0.824; s=0.17; F=25.84; p=7E-05;

**Trivariate regression**

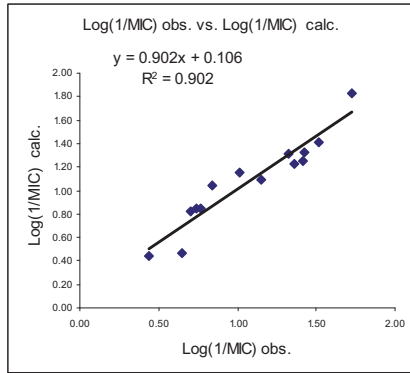
$$\text{Log}(1/\text{MIC}) = -0.007 + 0.832 \times \text{CD} - 0.983 \times \text{GATS4m} + 0.333 \times \text{ATS8e} \quad (11)$$

n=14; R<sup>2</sup>=0.901; s=0.13; F=30.58; p=24E-06; R<sup>2</sup><sub>cv</sub>=0.80

GATS4m and ATS8e belong to the 2D autocorrelation group of descriptors. It seems that partial charges descriptor (CD) associated with this type of indices (eq 10 and eq 11) give good correlations for Citobacter antimicrobial activity.

**Table 5.** Observed and calculated log(1/MIC), CV% in trivariate regression for **Citrobacter**

|     | <b>Citrobacter (log(1/MIC))</b> |       |                | CV%    |
|-----|---------------------------------|-------|----------------|--------|
|     | Obs                             | Calc  | Rez            |        |
| S1  | 0.652                           | 0.471 | 0.181          | 27.698 |
| S2  | 0.703                           | 0.821 | -0.118         | 16.768 |
| S3  | 0.742                           | 0.841 | -0.099         | 13.407 |
| S4  | 1.016                           | 1.150 | -0.134         | 13.238 |
| S5  | 1.357                           | 1.223 | 0.134          | 9.876  |
| S6  | 0.770                           | 0.847 | -0.077         | 9.998  |
| S7  | 1.412                           | 1.253 | 0.159          | 11.230 |
| S8  | 0.840                           | 1.038 | -0.198         | 23.537 |
| S9  | 1.421                           | 1.329 | 0.092          | 6.489  |
| S10 | 1.151                           | 1.094 | 0.057          | 4.987  |
| S11 | 1.325                           | 1.317 | 0.008          | 0.582  |
| S12 | 0.437                           | 0.447 | -0.010         | 2.205  |
| S13 | 1.728                           | 1.828 | -0.100         | 5.771  |
| S14 | 1.518                           | 1.413 | 0.105          | 6.945  |
|     |                                 |       | <b>10.909*</b> |        |



\* CV% average

**Escherichia coli**

**Bivariate regression**

$$\text{Log}(1/\text{MIC}) = 0.034 + 0.785 \times \text{CD} + 1.008 \times \text{R8m} \quad (12)$$

n=14; R<sup>2</sup>=0.92; s=0.08; F=61.86; 1E-06;

**Trivariate regression**

$$\text{Log}(1/\text{MIC}) = -0.011 + 0.894 \times \text{CD} + 0.950 \times \text{R8m} - 0.349 \times \text{E2s} \quad (13)$$

n=14; R<sup>2</sup>=0.951; s=0.07; F=64.81; p=1E-06; R<sup>2</sup><sub>cv</sub>=0.87

Calculated values and residuals together with the coefficient of variance (CV%) for eq 13 are presented in Table 6.

**Candida albicans**

**Bivariate regression**

$$\text{Log}(1/\text{MIC}) = 2.23 + 0.97 \times \text{CD} - 2.46 \times \text{ISH} \quad (14)$$

n = 14; R<sup>2</sup> = 0.975; s = 0.09; F = 216.41; p = 1.4E-09;

**Trivariate regression**

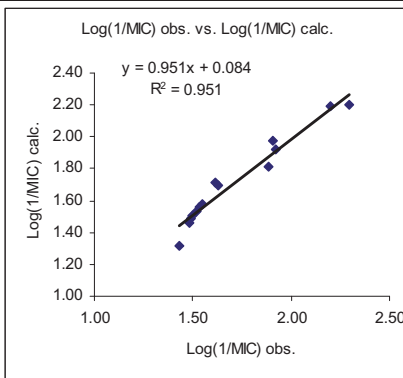
$$\text{Log}(1/\text{MIC}) = 2.061 + 0.989 \times \text{CD} - 2.441 \times \text{ISH} + 0.261 \times \text{E1m} \quad (15)$$

n = 14; R<sup>2</sup> = 0.985; s = 0.07; F = 226.06; p = 1.7E-09; R<sup>2</sup><sub>cv</sub> = 0.972

Calculated values and residuals together with CV% for eq 15 are given in Table 7.

**Table 6.** Observed and calculated log(1/MIC), CV% in trivariate regression for **Escherichia coli**

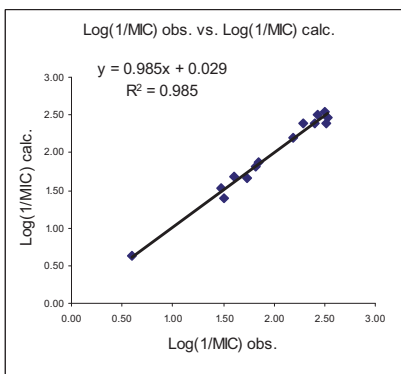
|     | Escherichia coli (log(1/MIC)) |       |        | CV%           |
|-----|-------------------------------|-------|--------|---------------|
|     | Obs                           | Calc  | Rez    |               |
| S1  | 1.430                         | 1.317 | 0.113  | 7.901         |
| S2  | 1.481                         | 1.463 | 0.018  | 1.215         |
| S3  | 1.520                         | 1.534 | -0.014 | 0.944         |
| S4  | 1.493                         | 1.486 | 0.007  | 0.485         |
| S5  | 1.533                         | 1.558 | -0.025 | 1.610         |
| S6  | 1.548                         | 1.579 | -0.031 | 1.994         |
| S7  | 1.889                         | 1.810 | 0.079  | 4.192         |
| S8  | 1.919                         | 1.923 | -0.004 | 0.231         |
| S9  | 2.199                         | 2.196 | 0.003  | 0.148         |
| S10 | 1.628                         | 1.693 | -0.065 | 3.979         |
| S11 | 1.501                         | 1.508 | -0.007 | 0.459         |
| S12 | 1.613                         | 1.717 | -0.104 | 6.478         |
| S13 | 1.904                         | 1.971 | -0.067 | 3.520         |
| S14 | 2.296                         | 2.199 | 0.097  | 4.219         |
|     |                               |       |        | <b>2.670*</b> |



\* CV% average

**Table 7.** Observed and calculated log(1/MIC), CV% in trivariate regression for **Candida albicans**

|     | Candida albicans (log(1/MIC)) |       |        | CV%           |
|-----|-------------------------------|-------|--------|---------------|
|     | Obs                           | Calc  | Rez    |               |
| S1  | 1.731                         | 1.667 | 0.064  | 3.679         |
| S2  | 1.481                         | 1.521 | -0.040 | 2.682         |
| S3  | 1.821                         | 1.815 | 0.006  | 0.306         |
| S4  | 2.396                         | 2.385 | 0.011  | 0.472         |
| S5  | 2.436                         | 2.497 | -0.061 | 2.498         |
| S6  | 1.849                         | 1.881 | -0.032 | 1.719         |
| S7  | 2.190                         | 2.203 | -0.013 | 0.590         |
| S8  | 2.521                         | 2.392 | 0.129  | 5.111         |
| S9  | 0.597                         | 0.629 | -0.032 | 5.421         |
| S10 | 2.532                         | 2.463 | 0.069  | 2.719         |
| S11 | 1.501                         | 1.398 | 0.103  | 6.854         |
| S12 | 1.613                         | 1.682 | -0.069 | 4.305         |
| S13 | 2.506                         | 2.550 | -0.044 | 1.771         |
| S14 | 2.296                         | 2.386 | -0.090 | 3.907         |
|     |                               |       |        | <b>3.002*</b> |



\* CV% average

## Conclusions

The Free-Wilson analysis was performed on a set of 14 sulphonylhydrazino-thiazoles on four microbial strains (Bacillus subtilis, Citrobacter, Escherichia coli and Candida albicans. Using Free-Wilson procedure, QSAR equations with moderate predictability were obtained.

All four antimicrobial activities are favorably influenced by the presence of the following substituents at the thiazole ring: methyl, methyl chloride, acetyl and ethoxycarbonyl. The methyl chloride substituent on position 4 of the thiazole ring increases the activity on *Escherichia* while the ethoxycarbonyl substituent on position 5 of the thiazole ring increases the activity on *Bacillus Subtilis* and decreases the *Escheria* antimicrobial activities.

The auto-correlating partial charges descriptor CD is the best descriptor in monovariate regression. When this electronic index is combined with Ghetaway, WHIM or 2D autocorrelation descriptors, the correlations are significantly improved.

The models obtained show good estimating and predicting abilities, as can be seen from the CV%, lower than 10, up to 2.7. The models are stable, and statistically significant, despite the rather small set of (newly synthesized) molecules under discussion.

## References

- [1]. Fahmy H.T.Y., *Pharmazie*, **1997**, 52, 750-752;
- [2]. Gursog A., Demirayak S., Cesur Z., Reisch J., Otuk G., *Die Pharmazie*, **1990**, 45, 246-249;
- [3]. Oniga O., Grosu I., Mager S., Simiti., *Monats. Chem.*, **1998**, 129, 661-669;
- [4]. Silva L.L., Navakoski de Oliveira K., Nunes R.J., *Arkivoc*, **2006**, 124-129;
- [5]. Kamal A., Khan M. N. A., Reddy S., Rohini K., *Bioorg. Med. Chem*, **2007**, 15, 1004-1013;
- [6]. Finch R.A., Shyam K., Penketh P.G., Sartorelli A.C., *Cancer Res.*, **2001**, 61, 3033-3038;
- [7]. Giles F., Verstovsek S., Thomas D., Gerson S., Cortes J., Fader S., Ferrajoli A., Ravandi F., Kornblau S., Garcia-Manero G., Jabbour E., O'Brien S., Karsten V., Cahill A., Yee K., Albitar M., Sznol M., Kantarjian H., *Clin. Cancer Res.*, **2005**, 11, 7817-7824;
- [8]. Ariesan V., Marie M., Cuparencu B., Safta L., *Therapie*, **1972**, 27, 309;
- [9]. Pignatello R., Mazzone S., Castelli I., Mazzone P., Racite G., Mazzone G., *Pharmazie*, **1994**, 49, 272-276;
- [10]. Holla B. S., Maloni K. V., Rao B. S., Sarojini B. K., Kumari N. S., *Eur. J. Med. Chem.*, **2003**, 38, 313-318;
- [11]. Zaharia V., Silvestru A., Manuela Curticean, Palibroda N., Vlasea M., Heterocycles 21, Synthesis, characterisation and evaluation of antimicrobial potential of some new sulfonylhydrazinotiazoles (*In press*);
- [12]. Kubinyi H., Kehrhahn O.H., *J. Med. Chem.*, 1976, 19, 578-86;
- [13]. Tmej C., Chiba P., Huber M., Richter E., Hitzler M., Schaper K. J., Ecker G., *Arch. Pharm. Pharm. Med. Chem.*, **1998**, 331, 233-240;
- [14]. Khan N., Soni L. K., Gupta A. K., Wakode S. R., Wagh R. D., Kaskhedikar S. G., *Indian J. Pharm. Sci.*, **2006**, 68, 341-346;
- [15]. Basak S. C., In Practical Applications of Quantitative Structure-Activity Relationships QSAR in Environmental Chemistry and Toxicology, (Karcher W, Devillers J. Eds, Kluwer Academic Publishers: Dordrecht, The Netherlands), **1990**, 83;
- [16]. Dragon Professional version 5.4-**2004**, Copyright (c) TALETE srl.;
- [17]. Iudea M. V., Ursu O., *TOPOCLUJ* (Copyright Babes-Bolyai Univ. Cluj), **2002**;
- [18]. Rios-Santamaria I., Garcia-Domenech R., Cortijo J., Santamaria P., Morcillo E. J., Galvez J., *Internet Elect. J. Molec. Des.*, **2002**, 1, 70-79;
- [19]. Galvez J., Garcia-Domenech R., Salabert M. T., Soler R., *J. Chem. Inf. Comput. Sci.*, **1994**, 34, 520-525.