



ISSN: 1697-090X

Inicio
Home

Indice del
volumen
Volume index

Comité Editorial
Editorial Board

Comité Científico
Scientific
Committee

Normas para los
autores
Instruction to
Authors

Derechos de autor
Copyright

Contacto/Contact:



MODELLING THE INHIBITORY ACTIVITY ON CARBONIC ANHYDRASE IV OF SUBSTITUTED THIADIAZOLE - AND THIADIAZOLINE - DISULFONAMIDES: INTEGRATION OF STRUCTURE INFORMATION

Lorentz Jäntschi and Sorana Daniela Bolboaca*

Technical University of Cluj-Napoca
**"Iuliu Hatieganu" University of Medicine and Pharmacy
Cluj-Napoca, Romania

lori@academicdirect.org

Rev Electron Biomed / Electron J Biomed 2006;2:22-33

Comment of the reviewer Jose Luis Hernandez Caceres, PhD. Full Professor. Center for Cybernetics Applications to Medicine (CECAM). Havana, Cuba.

Comment of the reviewer Prof.Dr. Amalio Garrido Escudero Chem Eng PhD. Head Environmental Engineering and Toxicology Dpt. Universidad Católica S. Antonio. Campus Los Jerónimos. Guadalupe (Murcia). SPAIN

ABSTRACT:

Purpose: To analyze the relationships between inhibitory activities on carbonic anhydrase IV and structures of substituted 1,3,4-thiadiazole and 1,3,4-thiadiazoline disulfonamide through integration of compounds complex structure information by the use of Molecular Descriptors Family.

Method: A number of forty compounds were used to generate and compute the molecular descriptors family and to build structure-activity relationships models. The obtained multi-varied models (the models with two, respectively with four descriptors) were validated by computing the cross-validation leave-one-out score (r^2_{cv-100}), and analyzed through assessment of the squared correlation coefficients (r^2), and the models stability ($r^2 - r^2_{cv-100}$). The estimation abilities of the multi-varied MDF-SAR model with four descriptors were analyzed in training and test sets.

Results: Analysis of the obtained models shows that the best results was obtained by the multi-varied model with four molecular descriptors ($r^2 = 0.920$). The prediction abilities of this model is sustained by the cross validation leave-one-out score ($r^2_{cv-100} = 0.903$), the model stability ($r^2 - r^2_{cv-100} = 0.017$), and the results on training versus test analysis (no significant differences between correlation coefficients in training and test sets, $p > 0.05$). The multi-varied model which used four descriptors proved to render higher value of correlation coefficient comparing with previous reported models ($p < 0.01$).

Conclusion: The multi-varied model with four descriptors is a solid and reliable one and indicates that the inhibition activity on carbonic anhydrase IV of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides is like to be of geometry and topology nature, being in relation with compounds partial charges.

KEYWORDS: Molecular Descriptors Family (MDF); Carbonic anhydrase IV inhibitors; Substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides

RESUMEN:

MODELIZACIÓN DE LA ACTIVIDAD INHIBIDORA DE LA CARBOANHIDRASA IV PRODUCIDA POR LAS DISULFONAMIDAS SUSTITUIDAS DEL TIADIAZOL Y DE LA TIADIAZOLINA: INTEGRACIÓN DE LA INFORMACIÓN ESTRUCTURAL.

Objetivo: Analizar las relaciones entre las actividades inhibidoras en la carboanhidrasa IV y las estructuras de las sulfonamidas sustituidas del 1,3,4-tiadiazol y de las disulfonamidas sustituidas de la 1,3,4-tiadiazolina mediante la integración compleja de la información estructural de los compuestos por el uso de Familias de Descriptores Moleculares.

Método: Un número de cuarenta compuestos fue utilizado para generar y para calcular las familias de descriptores moleculares y para construir modelos de las relaciones estructura-actividad. Los modelos multivariantes obtenidos (con dos y cuatro familias de descriptores respectivamente) fueron validados calculando la validación cruzada "leave-one-out" (r^2_{cv-100}), y analizadola mediante la

evaluación de los coeficientes de correlación ajustados (r^2), y la estabilidad de los modelos ($r^2 - r^2_{cv-100}$). Las capacidades de estimación del modelo multivariante de MDF-SAR con cuatro descriptores fueron comprobadas en los conjuntos de datos experimentales y previstos por el modelo.

Resultados: El análisis de los modelos obtenidos demuestra que los mejores resultados fueron obtenidos por el modelo multivariante con cuatro descriptores moleculares ($r^2 = 0.920$). La capacidad de predicción de este modelo son sostenidas por la validación cruzada "leave-one-out" $r^2_{cv-100} = 0.903$, la estabilidad del modelo ($r^2 - r^2_{cv-100} = 0.017$), y los resultados entre valores predichos por el modelo frente a valores experimentales (sin diferencias significativas entre los coeficientes de correlación en los conjuntos de datos, $p > 0.05$). El modelo multivariante que utilizó cuatro descriptores mostró un valor más alto del coeficiente de correlación en comparación con los modelos divulgados anteriormente ($p < 0.01$).

Conclusión: El modelo multivariante con cuatro descriptores es sólido y fiable e indica que la actividad de la inhibición en la carboanhidrasa IV producida por las sulfonamidas sustituidas del 1,3,4-tiadiazol- y de la 1,3,4-tiadiazolina- dependen de la naturaleza de la geometría y de la topología del compuesto, estando relacionadas con las cargas parciales de los mismos.

PALABRAS CLAVES: Familia de Descriptores Moleculares (MDF); inhibidores de la carboanhidrasa IV; Sulfonamidas sustituidas del 1,3,4-tiadiazol- y la 1,3,4-tiadiazolina.

INTRODUCTION

Quantitative structure-activity relationships (QSAR) is widely use in modelling of physico-chemical properties or biological activities of chemically active compounds¹⁻³, in most cases by the used of topological indices as: Szeged index^{4,5}, Wiener index⁶⁻⁸, PI (Padmakar - Ivan) index^{9,10}, Balaban index¹¹, Zagreb index¹², valence connectivity index¹³, Lu index¹⁴, DAI index (distance-based atom-type topological)¹⁵, AI index (atom-type topological index)¹⁶. Carbonic anhydrase IV, a membrane-bound form of carbonic anhydrase enzymes, is known to be, comparing with carbonic anhydrase II, a faster intra-molecular proton transporter¹⁷. Carbonic anhydrase inhibitors were studied by many authors through quantitative structure-activity relationships¹⁸⁻²⁴. More, the studies of carbonic anhydrase IV inhibition with sulfonamide lead to development of drugs used as antiglaucoma agents²⁵, anticonvulsant²⁶ agents, and as novel types of anticancer agents²⁷.

A number of forty substituted disulfonamide, twenty 1,3,4-thiadiazole disulfonamide and twenty 1,3,4-thiadiazoline disulfonamide, with carbonic anhydrase IV inhibition properties were previously studied by the use of quantum chemical quantitative structure-activity relationships²⁸. The models obtained by Supuran&Clare are in Table 1, and associated statistics in Table 2. In table 1, there were used the following abbreviations: Π_{zz} = the polarizability tensor, Q_{Cr1} = the changes of the atoms of the attached ring carbon, Q_{S1} = the changes of the atoms of the primary sulfonamide group, μ_x = the dipole moment, Q_{Nr1} = the charges on the nitrogen and ΔH_{S2} = the solvation energy for the secondary sulfonamide group²⁸.

Model no.	Expression
1	$\log IC_{50} = 7.31 \cdot 10^{-3} \cdot \Pi_{zz} - 5.570 \cdot Q_{Cr1} + 11.46 \cdot Q_{S1} + 0.0602 \cdot \mu_x - 37.16$
2	$\log IC_{50} = -2.798 \cdot Q_{Nr1} + 8.447 \cdot Q_{S1} + 0.1800 \cdot \mu_x + 0.5295 \cdot \text{LogP} + 0.0298 \cdot \Delta H_{S2} - 20.11$

Table 1. QSAR models for carbonic anhydrase IV inhibition activity

In the table 2 were used the following abbreviations: n for the studied sample size, R^2 for the square of the multiple correlation coefficient, Q^2 for the same quantity base on the predicted errors (the leave-one-out techniques), s for the standard errors of estimate of the equation, and F for the Fisher variance ratio.

Model no.	R^2	Q^2	s	F	n
1	0.632	0.444	0.34	15.07	40
2	0.769	0.646	0.20	20.05	36

Table 2. Statistics of models recorded in table 1

Starting from the successful results obtained by an original molecular descriptors family (MDF) on structure-activity relationships (SAR) methodology²⁹⁻³¹, the aim of the research was to investigate and to assess the estimation and prediction abilities of the MDF-SAR methodology on a sample of substituted thiadiazole- and thiadiazoline-disulfonamides.

MATERIALS AND METHODS

Carbonic anhydrase IV inhibitors

A sample of twenty 1,3,4-thiadiazole disulfonamides and twenty 1,3,4-thiadiazoline disulfonamides, with carbonic anhydrase IV inhibition properties was included into the study. The measured inhibitory activity of compounds, expressed as logarithm of concentration of the agent that is required for fifty percent inhibition in vitro ($\log IC_{50}$), was taken from a previous study²⁸.

The generic structure of compounds, the compounds abbreviation, substituents and compounds measured inhibition activity are in Table 3.

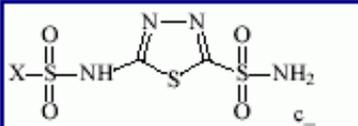
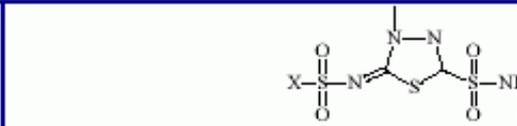
					
Abb.	X	log IC ₅₀ (nM)	Abb.	X	log IC ₅₀ (nM)
c 01	Me	0.6990	e 01	Me	0.9031
c 02	PhCH ₂	0.7782	e 02	PhCH ₂	0.9542
c 03	4-Me-C ₆ H ₄	0.4771	e 03	4-Me-C ₆ H ₄	0.4771
c 04	4-F-C ₆ H ₄	0.8451	e 04	4-F-C ₆ H ₄	0.8451
c 05	4-Cl-C ₆ H ₄	0.6990	e 05	4-Cl-C ₆ H ₄	0.6990
c 06	4-Br-C ₆ H ₄	0.6021	e 06	4-Br-C ₆ H ₄	0.7782
c 07	4-MeO-C ₆ H ₄	0.6021	e 07	4-MeO-C ₆ H ₄	0.6990
c 08	4-AcNH-C ₆ H ₄	0.9031	e 08	4-AcNH-C ₆ H ₄	0.3010
c 09	4-H ₂ N-C ₆ H ₄	0.6990	e 09	4-H ₂ N-C ₆ H	-0.0969
c 10	3-H ₂ N-C ₆ H ₄	0.8451	e 10	3-H ₂ N-C ₆ H ₄	-0.0969
c 11	4-O ₂ N-C ₆ H ₄	0.3010	e 11	4-O ₂ N-C ₆ H ₄	0.7782
c 12	3-O ₂ N-C ₆ H ₄	0.0000	e 12	3-O ₂ N-C ₆ H ₄	0.6990
c 13	2-O ₂ N-C ₆ H ₄	0.6021	e 13	2-O ₂ N-C ₆ H ₄	0.4771
c 14	Me ₂ N	1.1139	e 14	Me ₂ N	0.9031
c 15	2-HO ₂ CC ₆ H ₄	-0.2218	e 15	2-HO ₂ CC ₆ H ₄	-0.3010
c 16	4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄	1.0000	e 16	4-(2,4,6-Me ₃ Py ⁺)C ₆ H ₄	1.0792
c 17	4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄	2.5051	e 17	4-(2,4,6-Ph ₃ Py ⁺)C ₆ H ₄	2.2553
c 18	2,4-(O ₂ N) ₂ C ₆ H ₃	1.4472	e 18	2,4-(O ₂ N) ₂ C ₆ H ₃	0.9031
c 19	4-Cl-3-O ₂ N-C ₆ H ₃	0.8451	e 19	4-Cl-3-O ₂ N-C ₆ H ₃	0.6990
c 20	2,4,6-Me ₃ C ₆ H ₄	1.0792	e 20	2,4,6-Me ₃ C ₆ H ₄	0.9542

Table 3. Structure of compounds and measured IC₅₀**Modelling the inhibition activity of substituted thiadiazole- and thiadiazoline disulfonamide**

The inhibition activity of substituted thiadiazole- and thiadiazoline- disulfonamide were modelled through integration of complex information obtained from compounds structure. The applied methodology used an original molecular descriptors family approach in order to estimate and predict the inhibition activity of studied substituted thiadiazole- and thiadiazoline-disulfonamide.

The steps applied in MDF-SAR modelling²⁹ were:

- **Step 1 (3D representation of compounds)**
The three-dimensional representations of twenty 1,3,4-thiadiazole disulfonamide and twenty 1,3,4-thiadiazoline disulfonamide compounds were building up by the use of *HyperChem* software³².
- **Step 2 (creation of measured properties file)**
The measured inhibition concentration of each compound transformed to logarithmic scale was stored in *property.txt* file.
- **Step 3 (generation of molecular descriptors family)**
All forty compounds were used in construction and generation of the molecular descriptors family. The algorithm of molecular descriptors family list generation for studied substituted thiadiazole- and thiadiazoline-disulfonamide was strictly based on compounds structure. In order to discard redundant information a bias algorithm with a significance level of 10⁻⁹ was applied after molecular descriptors family generation. Each descriptor from family has an individual seven-letters name which expresses the modality of construction. The 7th letter indicates the compound characteristic relative to its geometry (*g*) or topology (*t*); the 6th letter, the atomic property (cardinality - *C*, number of directly bonded hydrogen's - *H*, atomic relative mass - *M*, atomic electronegativity - *E*, group electronegativity - *G*, partial charges - *Q*); the 5th letter indicates the atomic interaction descriptor; the 4th letter indicates the overlapping interaction model; the 3rd letter denotes the fragmentation criterion (the minimal fragments - *m*, the maximal fragments - *M*, the Szeged fragments criterion - *D*, and the Cluj fragments criterion - *P*^{33,34}); the 2nd letter refer the cumulative method of fragmentation properties (one out of nine selectors²⁹), and the 1st letter indicates the linearization procedure applied in generation of global molecular descriptor (identity - *I*, inverse - *i*, absolute - *A*, a inverse of absolute - *a*, natural logarithm of absolute value - *L*, simple natural logarithm - *l*).
- **Step 4 (search and identification of MDF- SAR models)**
The criterions imposed in search and identification of the MDF-SAR models were the highest values for the correlation coefficient and for the squared correlation coefficient (close to +1 or -1).
- **Step 5 (MDF-SAR models validation)**
Analysis of predictive abilities of the chosen MDF-SAR multi-varied models was performed by computing the following parameters: cross-validation leave-one-out (loo) scores (*r*²_{cv-loo}), Fisher parameter (*F*_{pred}) and associated significance (*p*_{pred}), and standard error (*s*_{loo}). In leave-one-out analysis the property of each compound was predicted by the regression equation calculated based on the all other compounds. The internal predictive analysis of MDF-SAR models was performed with Leave-

*one-out Analysis application*³⁵.

o *Step 6 (MDF-SAR models analysis)*

The chosen MDF-SAR models were analyzed through computing and interpreting the following parameters: the correlation coefficient (r), the squared correlation coefficient (r^2), the adjusted squared correlation coefficient (r^2_{adj}), the standard error of estimation (s_{est}), the Fisher parameter (F_{est}) and its significance (p_{est}), the 95% confidence intervals (95%CI_c) of multi-varied regression coefficients, the Student parameter ($t_{descriptor}$) and associated significance, the parameters of co-linearity (analysis of the squared correlation coefficients between descriptors, $r^2(descriptor, descriptor)$, and between each descriptor and measured IC₅₀ activity $r^2(log IC_{50}, descriptor)$), and the model stability (defines as the difference between squared correlation coefficient and leave-one-out correlation coefficient score ($r^2 - r^2_{cv-100}$); the model was consider stable if the difference had a lower value).

The comparison between multi-varied model with two and respectively four descriptors was performed through a correlated correlation analysis by the use of Steiger test³⁶.

The comparison between the correlation coefficients of previous reported models and the correlation coefficients obtained by MDF-SAR multi-varied models was analyzed by the use of Fisher's Z test. The estimation abilities of the model with the highest squared correlation coefficient was analyzed in training and test sets with *Training vs. Test* application³⁷. There were analyzed twelve situations, starting with sample sizes in training set between twenty to thirty-one and corresponding sample sizes in test set from twenty to nine. For each training and test set, the application rebuilds the regression model and based on the obtained model the activity of compound from test set was predicted.

RESULTS

Two MDF-SAR multi-varied models, one with two and other with four molecular descriptors, proved to have abilities in estimation and prediction of 1,3,4-thiadiazole disulfonamide and 1,3,4-thiadiazoline disulfonamides inhibitory property. The equations of the MDF-SAR models are:

$$\bullet_{2d} = 0.802 + 0.111 \cdot \ln PRIQg + 9.980 \cdot 10^{-9} \cdot iHMMTQt \quad (1)$$

$$\bullet_{4d} = 0.625 + 0.105 \cdot \ln PRIQg + 9.919 \cdot 10^{-9} \cdot iHMMTQt - 9.248 \cdot IHMDTQg + 1.727 \cdot \ln PdJQg \quad (2)$$

The molecular descriptors used by the models, their calculated values and the values of the inhibition concentration 50% (IC₅₀) estimated by each models (\bullet_{2d} respectively \bullet_{4d}) are in Table 4.

Abb.	inPRIQg	iHMMTQf	IHMDTQg	InPdJQg	\hat{Y}_{2d}	\hat{Y}_{4d}
c 01	-0.7353	$1.18 \cdot 10^3$	0.0007	0.0956	0.7222	0.7076
c 02	-2.2423	$3.95 \cdot 10^6$	0.0003	0.2080	0.5927	0.7850
c 03	-2.8218	$1.43 \cdot 10^6$	0.0012	0.1283	0.5029	0.5535
c 04	-0.1521	$1.11 \cdot 10^3$	0.0218	0.2435	0.7870	0.8293
c 05	-0.0950	$2.06 \cdot 10^3$	0.0105	0.1693	0.7943	0.8126
c 06	-0.0102	$1.47 \cdot 10^6$	0.0009	0.0309	0.8164	0.6836
c 07	-0.2617	$1.91 \cdot 10^3$	0.0163	0.1145	0.7756	0.6462
c 08	-2.0460	$3.17 \cdot 10^3$	0.0296	0.3202	0.5782	0.6923
c 09	-3.5436	$1.63 \cdot 10^3$	0.0130	0.3811	0.4100	0.7923
c 10	-3.5385	$1.51 \cdot 10^3$	0.0386	0.3287	0.4104	0.4655
c 11	-0.5461	$1.64 \cdot 10^3$	0.0405	0.1993	0.7437	0.5392
c 12	-1.7861	$1.42 \cdot 10^3$	0.1154	0.3446	0.6054	-0.0333
c 13	-1.6233	$1.22 \cdot 10^3$	0.0522	0.2935	0.6233	0.4799
c 14	-0.8071	$5.36 \cdot 10^4$	0.0045	0.2888	0.7135	0.9980
c 15	-3.6509	$1.34 \cdot 10^3$	0.0399	0.0129	0.3977	-0.1043
c 16	-1.0801	$4.10 \cdot 10^7$	0.0009	0.0611	1.0918	1.0155
c 17	-0.0312	$1.67 \cdot 10^8$	0.0016	0.1169	2.4673	2.4664
c 18	-0.4637	$4.29 \cdot 10^3$	0.0313	0.7259	0.7555	1.5449
c 19	-0.4667	$3.24 \cdot 10^3$	0.0067	0.1848	0.7541	0.8363
c 20	-0.0543	$1.54 \cdot 10^6$	0.0011	0.1266	0.8122	0.8433
e 01	-0.8572	$5.09 \cdot 10^3$	0.0003	0.0546	0.7125	0.6317
e 02	-0.9993	$4.97 \cdot 10^6$	0.0004	0.1940	0.7412	0.9006
e 03	-0.0549	$1.78 \cdot 10^6$	0.0007	0.0203	0.8145	0.6655
e 04	-0.6339	$3.50 \cdot 10^3$	0.0098	0.0531	0.7358	0.5631
e 05	-0.0465	$4.11 \cdot 10^3$	0.0065	0.1281	0.8018	0.7856
e 06	-0.3026	$4.00 \cdot 10^6$	0.0002	0.0613	0.8091	0.7366
e 07	-0.8485	$5.46 \cdot 10^3$	0.0076	0.1111	0.7138	0.6627
e 08	-0.6516	$1.01 \cdot 10^6$	0.0121	0.1075	0.7404	0.6402
e 09	-9.0604	$4.27 \cdot 10^3$	0.0056	0.1805	-0.2016	-0.0627
e 10	-8.7988	$4.10 \cdot 10^3$	0.0087	0.1296	-0.1726	-0.1519
e 11	-0.3044	$5.87 \cdot 10^3$	0.0139	0.1104	0.7748	0.6609
e 12	-0.5566	$5.19 \cdot 10^3$	0.0200	0.2415	0.7461	0.8040
e 13	-0.7842	$4.47 \cdot 10^3$	0.0211	0.2449	0.7200	0.7744
e 14	-0.2030	$1.27 \cdot 10^3$	0.0043	0.2208	0.7815	0.9468
e 15	-7.5361	$4.67 \cdot 10^3$	0.0195	0.1050	-0.0315	-0.1610
e 16	-0.2658	$2.75 \cdot 10^7$	0.0009	0.0621	1.0475	0.9687
e 17	-0.0708	$1.50 \cdot 10^8$	0.0010	0.1308	2.2873	2.3170
e 18	-0.0811	$9.32 \cdot 10^3$	0.0202	0.2532	0.8031	0.8759
e 19	-0.8712	$6.99 \cdot 10^3$	0.0128	0.2206	0.7128	0.8033
e 20	-0.0068	$3.85 \cdot 10^6$	0.0009	0.0937	0.8405	0.8158

Table 4. Descriptors used in MDF-SAR models, their values and estimated IC50

The statistics of the MDF-SAR models are in Table 5 and Table 6.

Parameter	Value	
	Model with two descriptors	Model with four descriptors
N	40	40
V	2	4
R	0.8673	0.9593
r^2	0.7521	0.9202
r^2_{adj}	0.7387	0.9111
S_{est}	0.2741	0.1599
F_{est}	56	101
p_{est}	< 0.001	< 0.001
r^2_{cv-100}	0.7279	0.9034
S_{100}	0.2873	0.1762
F_{pred}	49	82
p_{pred}	< 0.001	< 0.001
$r^2 - r^2_{cv-100}$	0.0242	0.0168
$r^2(\text{inPRIQg, iHMMTQt})$	0.0271	0.0271
$r^2(\text{inPRIQg, IHMDTQg})$	n.a.	0.0122
$r^2(\text{inPRIQg, InPdJQg})$	n.a.	0.0010
$r^2(\text{iHMMTQt, IHMDTQg})$	n.a.	0.0365
$r^2(\text{iHMMTQt, InPdJQg})$	n.a.	0.0212
$r^2(\text{IHMDTQg, InPdJQg})$	n.a.	0.2085

**n = number of compounds, v = number of descriptors,
n.a. = not applicable**

Table 5. The statistical parameter of MDF-SAR models

	StdError	$r^2(\text{Y, desc})$	t	95% CI _c
Model with two descriptors				
Intercept	0.0549		14.616*	[0.6915, 0.9141]
InPRIQg	0.0197	0.3335	5.6589*	[0.0715, 0.1512]
iHMMTQt	$1.26 \cdot 10^{-9}$	0.5376	7.9051*	$[7.4 \cdot 10^{-9}, 1.2 \cdot 10^{-8}]$
Model with four descriptors				
Intercept	0.0485		12.891*	[0.5265, 0.7233]
inPRIQg	0.0115	0.3335	9.1175*	[0.0817, 0.1284]
iHMMTQt	0.0000	0.5376	13.223*	$[8.4 \cdot 10^{-9}, 1.1 \cdot 10^{-8}]$
IHMDTQg	1.3826	0.1198	-6.6894*	[-12.055, -6.4418]
InPdJQg	0.2218	0.0208	7.7875*	[1.2771, 2.1778]

StdError = standard error, Y = log IC₅₀, desc = molecular descriptor, t = parameter of the Student test; * p < 0.001

Table 6. The qualities of MDF-SAR models

The plots of residuals (defined as the difference between the measured IC₅₀ and estimated by the regression equations (1) and (2)) obtained with the MDF-SAR models with two and respectively four molecular descriptors are in Figure 1.

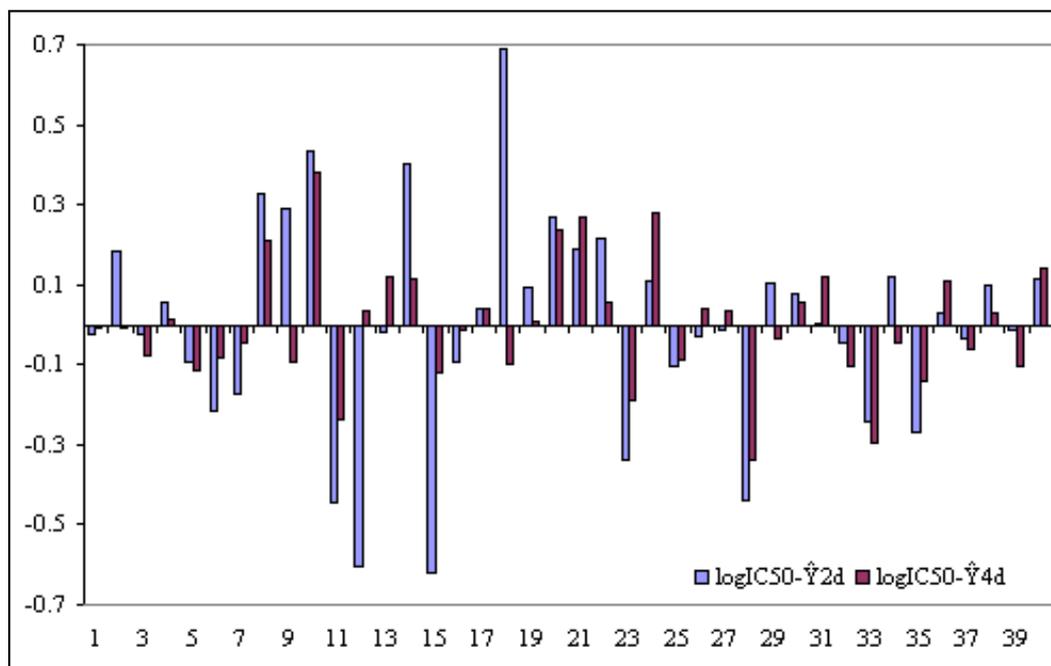


Figure 1. The plots of residuals obtained with Eq. (1) and (2)

The graphical representation of the measured IC_{50} versus estimated with MDF-SAR model with four descriptors (Eq.(2)) is in Figure 2.

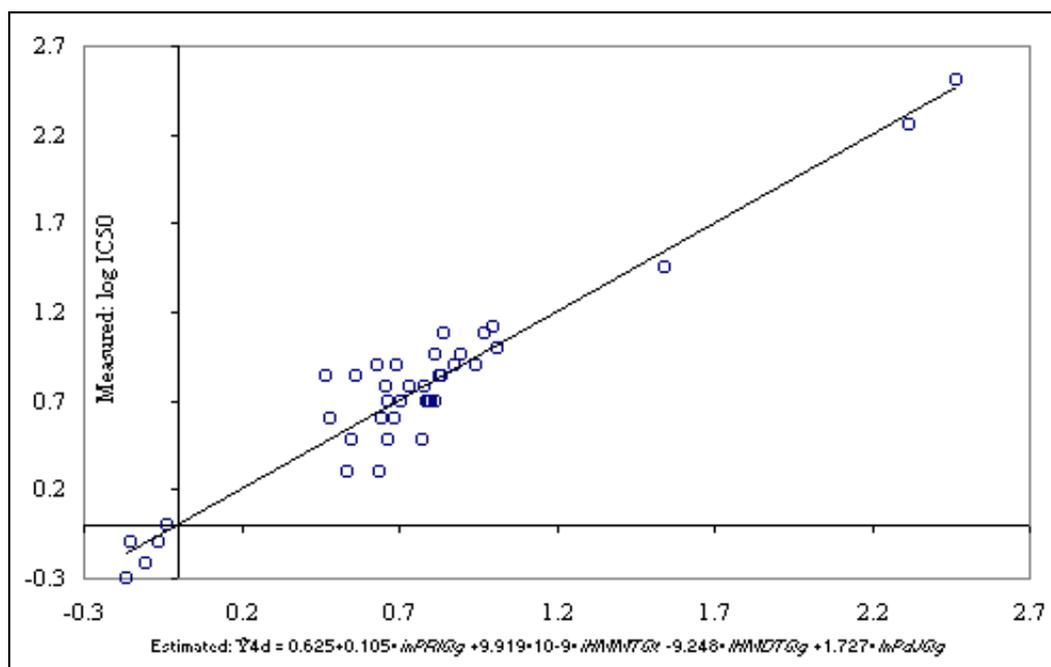


Figure 2. Estimated with the Eq. (2) versus measured IC_{50} of substituted thiazole- and thiazoline-disulfonamide

A correlated correlation analysis was used in order to test the hypothesis that the correlation coefficient obtained by the MDF-SAR model with four descriptors is not statistical different comparing with the correlation coefficient obtained with the MDF-SAR model with two descriptors. The results expressed as correlation coefficients, parameter of the Steiger's Z test and associated significance are in Table 7.

Parameter	Value
$r_{(\log IC50 - \hat{Y}4d)}$	0.9593
$r_{(\log IC50 - \hat{Y}2d)}$	0.8670
$r_{(\hat{Y}4d - \hat{Y}2d)}$	0.9041
Steiger's Z	3.9769
$p_{\text{Steiger's Z}}$	$3.49 \cdot 10^{-5}$

Table 7. The results of the correlated correlation analysis

Comparing the performances of the MDF-SAR model with two descriptors (Eq. (1)), respectively with four descriptors (Eq. (2)) with previous reported models (model 1 and 2, table 2), there was observed that the MDF-SAR model with four variables obtained a correlation coefficient statistically grater (Eq.(2)-model 1, $p_{\text{Steiger's Z}} = 1.38 \cdot 10^{-4}$; Eq.(2)-model 2, $p_{\text{Steiger's Z}} = 8.40 \cdot 10^{-3}$).

The results obtained in training versus test analysis, applied for the MDF-SAR model with four descriptors are in Table 8, and the plot of squared correlation coefficients in Figure 3. The abbreviations used in Table 8 had the following significances: the number of training versus test analysis (No.), the coefficients ($a_0, a_1, a_2, a_3,$ and a_4) of the generic MDF-SAR model with four descriptors ($\bullet_{4d} = a_0 + a_1 \cdot \text{inPRIQg} + a_2 \cdot \text{iHMMTQt} + a_3 \cdot \text{iHMDTQg} + a_4 \cdot \text{InPdJQg}$), the number of compounds in training (No_{tr}) respectively in test (No_{ts}) sets, the correlation coefficients obtained in training (r_{tr}) and test (r_{ts}) sets and associated 95% confidence intervals ($95\%CI_{tr}$ - for training sets, respectively $95\% CI_{ts}$ - for test sets); the Fisher Z test for training (F_{tr}) respectively for test (F_{ts}) sets, and the Fisher Z- test (F_{Z-test}) which test the null hypothesis that there were not significant differences between correlation coefficient obtained in training set and the correlation coefficient obtained in the associated test set.

N ^o	Training					Test					r _{tr} vs. r _{ts}			
	a ₀	a ₁	a ₂	a ₃	a ₄	No _{tr}	r _{tr}	95%CI _{tr}	F _{tr}	No _{ts}	r _{ts}	95% CI _{ts}	F _{ts}	F _{Z-test}
1	0.307	0.086	$1.01 \cdot 10^{-8}$	-8.543	1.603	20	0.958	[0.894, 0.983]	42 ^{***}	20	0.970	[0.924, 0.988]	30 ^{***}	0.51†
2	0.411	0.102	$9.79 \cdot 10^{-9}$	-10.04	2.263	21	0.948	[0.874, 0.979]	36 ^{***}	19	0.956	[0.886, 0.983]	35 ^{***}	0.23†
3	0.153	0.094	$1.02 \cdot 10^{-8}$	-8.355	1.726	22	0.969	[0.925, 0.987]	65 ^{***}	18	0.938	[0.838, 0.977]	20 ^{***}	1.02†
4	0.145	0.070	$1.04 \cdot 10^{-8}$	-8.587	1.563	23	0.926	[0.831, 0.968]	27 ^{***}	17	0.964	[0.901, 0.987]	32 ^{***}	1.06†
5	0.232	0.115	$9.94 \cdot 10^{-9}$	-9.229	1.756	24	0.964	[0.917, 0.984]	62 ^{***}	16	0.940	[0.832, 0.979]	19 ^{***}	0.73†
6	0.499	0.122	$1.02 \cdot 10^{-8}$	-9.243	1.563	25	0.959	[0.908, 0.982]	58 ^{***}	15	0.959	[0.879, 0.987]	22 ^{***}	0.00†
7	0.406	0.057	$1.03 \cdot 10^{-8}$	-9.432	1.572	26	0.941	[0.871, 0.973]	4 ^{***}	14	0.950	[0.846, 0.984]	13 ^{***}	0.24†
8	0.406	0.104	$9.59 \cdot 10^{-9}$	-9.076	1.414	27	0.951	[0.894, 0.978]	52 ^{***}	13	0.964	[0.879, 0.989]	19 ^{***}	0.24†
9	0.154	0.101	$1.01 \cdot 10^{-8}$	-9.838	1.823	28	0.965	[0.926, 0.984]	78 ^{***}	12	0.925	[0.747, 0.979]	5 [*]	1.02†
10	0.320	0.111	$9.93 \cdot 10^{-9}$	-9.160	1.642	29	0.927	[0.849, 0.965]	76 ^{***}	11	0.949	[0.812, 0.987]	10 ^{**}	0.47†
11	0.073	0.104	$9.44 \cdot 10^{-9}$	-8.660	1.622	30	0.949	[0.893, 0.975]	56 ^{***}	10	0.985	[0.934, 0.996]	20 ^{**}	1.45†
12	0.320	0.102	$1.03 \cdot 10^{-8}$	-9.795	1.838	31	0.963	[0.924, 0.982]	84 ^{***}	9	0.944	[0.752, 0.988]	7 [*]	0.47†

* 0.01 < p < 0.05, ** 0.001 < p < 0.01, *** p < 0.001, † p > 0.05

Table 8. The quality of MDF-SAR multi-varied model with four descriptors in training versus test analysis

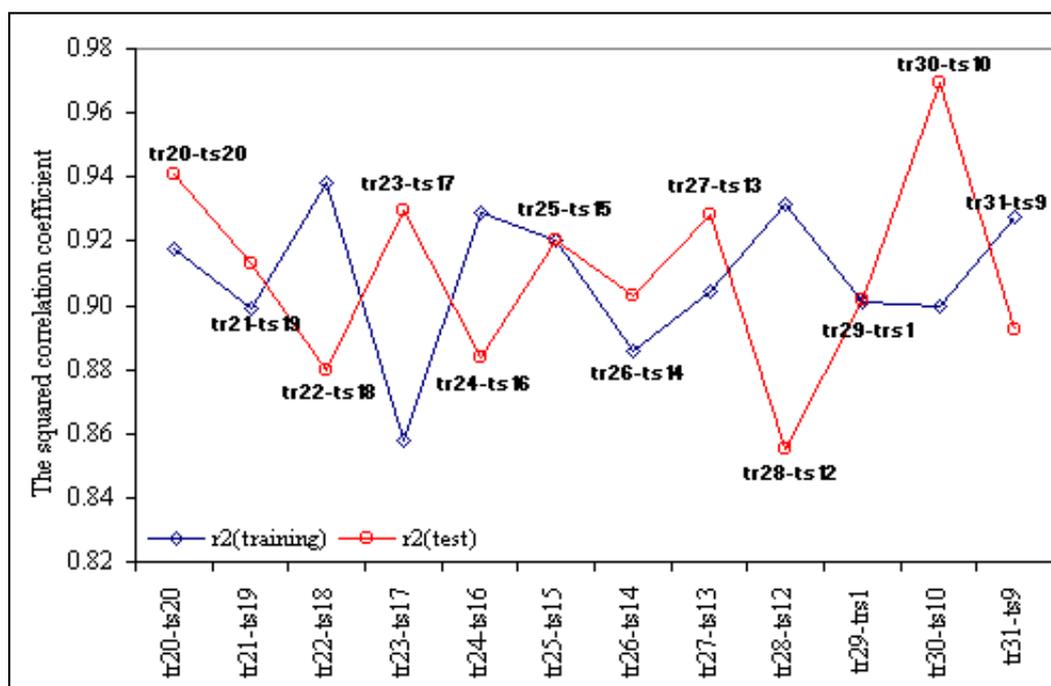


Figure 3. Training and test squared correlation coefficients obtained with MDF-SAR model with 4 descriptors

DISCUSSIONS

The inhibition activity on carbonic anhydrase IV of substituted 1,3,4-thiadiazole and 1,3,4-thiadiazoline disulfonamides proved to be in relationship with the complex information obtained from compounds structure. In both multi-varied MDF-SAR equations (Eq.(1) and Eq.(2)) the inhibition activity on carbonic anhydrase IV was in relationship with the geometry (*inPRIQg*, *IHMDTQg*, *InPdJQg*) as well as the topology (*iHMMTQg*) of the compounds and was strongly depend on the compounds partial charges (*inPRIQg*, *iHMMTQg*, *IHMDTQg*, *InPdJQg*)

Two out of two (in MDF-SAR model with two descriptors, Eq.(1)), respectively three out of four (in MDF-SAR model with four descriptors, Eq.(2)) molecular descriptors had positive regression coefficients. Just one out of four descriptors (in MDF-SAR model with four descriptors, Eq.(2)) had a negative regression coefficient.

Analyzing the performances of MDF-SAR models it can be observed that both are significant in estimation as well as in prediction (see statistical parameters from table 5). Almost seventy-five percent in inhibition activity on carbonic anhydrase IV of studied disulfonamides can be explained by its linear relationship with the variation of two molecular descriptors (Eq.(1)). The MDF-SAR model with two descriptors proved to be valid and stable ($p < 0.001$; $r^2_{cv-loo} = 0.7279$; $r^2 - r^2_{cv-loo} = 0.0242$).

Both descriptors used in Eq.(1) could be found again in MDF-SAR model with four variables, demonstrating that these descriptors had an important contribution in modelling of the inhibition activity on carbonic anhydrase IV of substituted 1,3,4-thiadiazole disulfonamide, 1,3,4-thiadiazoline disulfonamides. The abilities in estimation for the MDF-SAR model with four descriptors are sustained by the value of correlation coefficient ($r^2 = 0.9593$, table 5), confidence boundaries associated with the regression coefficients and probability associated with Student tests (for all coefficients less than 0.001, see table 6). Almost ninety-six percent from variation of inhibition activity on carbonic anhydrase IV can be explained by its linear relationship with the variation of four molecular descriptors used in Eq. (2). The Fisher parameter computed in leave-one-out analysis and its significance ($F_{pred} = 82$, $p < 0.001$) sustains the estimation ability of the model. The stability of the MDF-SAR model (Eq.(2)) is sustained by the values of difference between correlation coefficient ($r^2 - r^2_{cv(lo0)} = 0.0168$) and by the cross validation leave-one-out correlation score ($r^2_{cv-loo} = 0.9034$). The power of the model with four descriptors in prediction of the inhibition on carbonic anhydrase IV of studied disulfonamides is sustained by the absence of co-linearity between descriptors (see the squared correlation coefficients between pairs of descriptors, which always is less than 0.20, table 5).

The MDF-SAR model with four descriptors proved to had a significantly greater correlation coefficient comparing with the MDF-SAR model with two descriptors ($p < 0.001$, table 7) and statistically significant greater correlation coefficients comparing with previous reported models ($p < 0.01$). The residuals of the MDF-SAR model with two descriptors varied from -0.6196 to 0.6917 while the residual of the MDF-SAR model with four descriptors varied from -0.3392 to 0.3796. The qualities of the MDF-SAR model with four descriptors in training versus test analysis showed that excepting the intercept, in more than 83%, the values of equation coefficients (a_1 , a_2 , a_3 , and a_4) are included into the 95% confidence intervals of the model (see table 6 and table 8). More, with a single exception (for training sample size equal with thirty, table 8) the values of the correlation coefficients obtained in test sets are included into the 95% confidence intervals of the correlation coefficients obtained in training sets. There were not identify statistical significances between the correlation coefficients in training and test sets (p -value > 0.05 , table 8).

Starting from the above described MDF-SAR model with four descriptors, and by the use of an original software³⁸, the inhibition activity on carbonic anhydrase IV of new substituted 1,3,4-thiadiazole-, and 1,3,4-thiadiazoline-disulfonamides can be calculated in a short time, without any experiments. Future studies are necessary in order to assess the performances in computing the inhibition activity on carbonic anhydrase IV of new substituted 1,3,4-thiadiazole-disulfonamide or 1,3,4-thiadiazoline-disulfonamide.

CONCLUSIONS

The multi-varied MDF-SAR model with four descriptors is a solid and reliable one and indicates that the inhibition activity on carbonic anhydrase IV of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides is like to be of geometry and topology nature, being in relation with compounds partial charges.

The MDF-SAR methodology opens a new pathway in computing and characterization of inhibitory activity on carbonic anhydrase IV of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides.

REFERENCES

- 1.- Khadikar PV, Karmarkar S, Gour K, Agrawal VK, Singh S. QSAR study using distance-based topological indices. *Oxidation Communications* 2004;27(1):1-11.
- 2.- Allen DD, Geldenhuys WJ. Molecular modeling of blood-brain barrier nutrient transporters: in silico basis for evaluation of potential drug delivery to the central nervous system. *Life Sci.* 2006;78(10):1029-33.
- 3.- Verma RP, Mekapati SB, Kurup A, Hansch C. A QSAR review on melanoma toxicity. *Bioorg Med Chem* 2005;13(19):5508-26.
- 4.- Khadikar PV, Karmarkar S, Agrawal VK, Singh J, Shrivastava A, Lukovits I, Diudea MV. Szeged index - Applications for drug modeling. *Letters in Drug Design and Discovery* 2005;2 (8):606-624.
- 5.- Gutman I, Dobrynin AA. The Szeged index - A success story. *Graph Theory Notes New York* 1998;34:37-44.
- 6.- Nikolic S, Trinajstić N, Mihalić Z. The Wiener index: Development and applications. *Croat Chem Acta* 1995;68: 105-29.
- 7.- Yang F, Wang ZD, Huang YP. Modification of the Wiener index 4. *J Comput Chem.* 2004;25(6):881-7.
- 8.- Vukicević D, Žerovnik J. Altered Wiener indices. *Acta Chimica Slovenica* 2005;52(3):272-81.
- 9.- Khadikar PV, Karmarkar S, Agrawal VK. A Novel PI Index and its Applications to QSPR/QSAR Studies. *J Chem Inf Comput Sci* 2001; 41(4):934-949.
- 10.- Jaiswal M, Khadikar P. QSAR study on tadpole narcosis using PI index: a case of heterogeneous set of compounds. *Bioorg Med Chem* 2004;12(7):1731-6.
- 11.- Thakur A, Thakur M, Khadikar PV, Supuran CT, Sudele P. QSAR study on benzenesulphonamide carbonic anhydrase inhibitors: topological approach using Balaban index. *Bioorg Med Chem* 2004;15;12(4):789-93.
- 12.- Nikolic S, Kovacević G, Milicević A, Trinajstić N. The Zagreb indices 30 years after. *Croatica Chemica Acta* 2003;76(2):113-24.
- 13.- Li XH, Jalbout AF, Solimannejad M. Definition and application of a novel valence molecular connectivity index. *Journal of Molecular Structure: THEOCHEM* 2003;663:81-5.
- 14.- Lu C, Guo W, Hu X, Wang Y, Yin C. A Lu index for QSAR/QSPR studies. *Chem Phys Lett* 2006;417(1-3):11-5.
- 15.- Lu C, Guo W, Wang Y, Yin C. Novel distance-based atom-type topological indices DAI for QSPR/QSAR studies of alcohols. *J Mol Model (Online)*. 2006;1-8.
- 16.- Ren B. Application of novel atom-type AI topological indices to QSPR studies of alkanes. *Comput Chem* 2002;26(4):357-69.
- 17.- Baird TT Jr, Waheed A, Okuyama T, Sly WS, Fierke CA. Catalysis and inhibition of human carbonic anhydrase IV. *Biochemistry* 1997;36(9):2669-78.
- 18.- Clare BW, Supuran CT. A physically interpretable quantum-theoretic QSAR for some carbonic anhydrase inhibitors with diverse aromatic rings, obtained by a new QSAR procedure. *Bioorg Med Chem.* 2005;13(6):2197-211.
- 19.- Jaiswal M, Khadikar PV, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: the first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and heterocyclic sulfonamides. *Bioorg Med Chem Lett.* 2004;14(12):3283-90.
- 20.- Vullo D, Steffansen B, Brodin B, Supuran C, Scozzafava A, Nielsen C. Carbonic anhydrase inhibitors: Transepithelial transport of thioureido sulfonamide inhibitors of the cancer-associated isozyme IX is dependent on efflux transporters. *Bioorg Med Chem* 2006;14(7):2418-2427.
- 21.- 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;14:1108-14.

- 22.- Supuran CT, Clare BW. Quantum theoretic QSAR of benzene derivatives: some enzyme inhibitors. *J Enzyme Inhib Med Chem*. 2004;19(3):237-48.
- 23.- 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(4):1108-14.
- 24.- Gupta SP, Kumaran S. A quantitative structure-activity relationship study on some aromatic/heterocyclic sulfonamides and their charged derivatives acting as carbonic anhydrase inhibitors. *J Enzyme Inhib Med Chem*. 2005;20(3):251-9.
- 25.- Pastorekova S, Parkkila S, Pastorek J, Supuran CT. Carbonic anhydrases: current state of the art, therapeutic applications and future prospects. *J Enzyme Inhib Med Chem* 2004;19(3):199-229.
- 26.- Masereel B, Rolin S, Abbate F, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: anticonvulsant sulfonamides incorporating valproyl and other lipophilic moieties. *J Med Chem* 2002;45(2):312-20.
- 27.- Svichar N, Esquenazi S, Waheed A, Sly WS, Chesler M. Functional demonstration of surface carbonic anhydrase IV activity on rat astrocytes. *Glia* 2006;53(3):241-7.
- 28.- Supuran TC, Clare WB. Carbonic anhydrase inhibitors - Part 57: Quantum chemical QSAR of a group of 1,3,4-thiadiazole- and 1,3,4-thiadiazoline disulfonamide with carbonic anhydrase inhibitory properties. *Eur J Med Chem* 1999;34:41-50.
- 29.- Jäntschi L. Molecular Descriptors Family on Structure Activity Relationships 1. The review of Methodology. *Leonardo Electronic Journal of Practices and Technologies* 2005;6:76-98.
- 30.- Jäntschi L. Delphi Client - Server Implementation of Multiple Linear Regression Findings: a QSAR/QSPR Application. *Applied Medical Informatics* 2004;15:48-55.
- 31.- Bolboaca S, Jäntschi L. Molecular Descriptors Family on Structure Activity Relationships 3. Antituberculous Activity of some Polyhydroxyxanthenes, *Leonardo Journal of Sciences* 2005;7:58-64.
- 32.- HyperChem , Molecular Modelling System [Internet page]; ©2003, Hypercube [about three screens]; [cited 2005 Nov]. Available from URL: <http://hyper.com/products/>
- 33.- Jäntschi L, Katona G, Diudea VM. Modeling Molecular Properties by Cluj Indices. *Commun. Math. Comput. Chem. (MATCH)* 2000;41:151-188.
- 34.- Diudea M, Gutman I, Jäntschi L. *Molecular Topology*, 2nd Edition, Nova Science, Huntington, New York, 2002.
- 35.- Leave-one-out Analysis. ©2005, Virtual Library of Free Software [cited 2006 March]. Available from URL: http://vl.academicdirect.org/molecular_topology/mdf_findings/loo/.
- 36.- Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull* 1980;87:245-51.
- 37.- Training vs. Test Experiment. ©2005, Virtual Library of Free Software [cited 2006 March]. Available from URL: http://vl.academicdirect.org/molecular_topology/qsar_qspr_sl/.
- 38.- MDF SAR Predictor, © 2005, Virtual Library of Free Software [cited 2006 March]. Available from URL: http://vl.academicdirect.org/molecular_topology/mdf_findings/sar/.

ACKNOWLEDGEMENT: Research was in partly supported by UEFISCSU Romania through project ET36/2005

Comment of the reviewer Jose Luis Hernandez Caceres, PhD. Full Professor. Center for Cybernetics Applications to Medicine (CECAM). Havana, Cuba.

One of the benefits from the wide availability of computers at research laboratories is the possibility to make experiments "in silico", especially for predicting compounds with a desired biological effect. This paper is a nice example on this direction.

Authors showed that it is possible to predict the anhydrase IV inhibitory properties of certain compounds on the basis of a set of four descriptors obtained from compounds' chemical structure. A result like this is encouraging not only theoretically, but also from the practical viewpoint. Few years ago pharmacological predictions were made basically by trial and error, and experiments were long lasting and costly.

Comment of the reviewer Prof.Dr. Amalio Garrido Escudero Chem Eng PhD. Head Environmental Engineering and Toxicology Dpt. Universidad Católica S. Antonio. Campus Los Jerónimos. Guadalupe (Murcia). SPAIN

The aim of this paper is to build a Structure Activity Relationships model using a Molecular Description Family methodology to predict the inhibitory activity of anhydrase IV for a group of compounds with a common root structure and different substituents.

Developing a structure activity relationship model is difficult and required a good set of physico-chemical properties and activity data, statistical tools and methodology. Molecules are typically flexible and it is possible to compute many possibly useful properties that might relate to activity. The way to manage these data is a table in which compounds are defined by individual rows and molecular properties (or descriptors) are defined by the columns. A SAR model attempts to find consistent relationships between the variations in the values of molecular properties and the biological activity for a series of compounds so that these rules can be used to predict the activity of new chemical entities with similar structure. A SAR model generally takes the form of a linear equation.

The authors research an original methodology to group the descriptors into different sort of families. This strategy results in a better fit for the predicted inhibitory activity of forty compounds and better model stability than the previous quantum chemical quantitative structure-activity relationship model built by Supuran et al. in 1999. The model shows a very good correlation between real and predicted results for this kind of models.

This strategy can be used not only to predict properties for new sulphonamide root structure compounds to be used as antiglaucoma, anticonvulsants or anticancer agents but to predict carcinogenic, mutagenic, theratogenic or endocrine disruptor activity for other group of substances than reduce the amount of in vivo and in vitro testing needs specially presents in the environment as a result of degradation of synthetic chemicals or byproducts.

The importance to research in this field is relevant to reduce the time needed to identify and set up preventive rules and correction treatments. The contribution of the authors improve the methodology to built such kind of useful models.

Correspondence:

Lorentz JÄNTSCHI	Sorana Daniela BOLBOACA
Technical University of Cluj-Napoca,	"Iuliu Hatieganu" University of Medicine and Pharmacy,
15 Constantin Daicoviciu Street,	13 Emil Isac Street,
400020 Cluj-Napoca, Romania	400023 Cluj-Napoca, Romania,
http://lori.academicdirect.org	http://sorana.academicdirect.ro

Received, June 16, 2006. Received reviewed July 26, 2006
Publicado, 31 de Julio de 2006.