Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazole- and Thiadiazoline-Disulfonamides: Integration of Structure Information

Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ
Technical University, Iuliu Hatieganu University
Cluj-Napoca, Romania

Aim

The paper presents the abilities in estimation and prediction of the inhibition on carbonic anhydrase I of some substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides through the integration of complex structures information by using of an original molecular descriptors family on the structure-activity relationships approach.

Material and Method

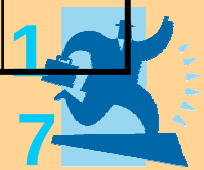
The proposed approach uses the complex information obtained from substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides structure in order to generate and calculate the molecular descriptors family. The structure-activity relationship models were built based on the generated descriptors. The obtained multivariate models (the models with two, respectively four descriptors) were validated by computing the cross-validation leave-one-out score ($r^2_{cv\text{-}loo}$), and analyzed through assessment of the squared correlation coefficients (r^2), and the models stability (r^2 - $r^2_{cv\text{-}loo}$). The prediction ability of the multivariate MDF-SAR model with four descriptors was analyzed in training versus test sets.

Results

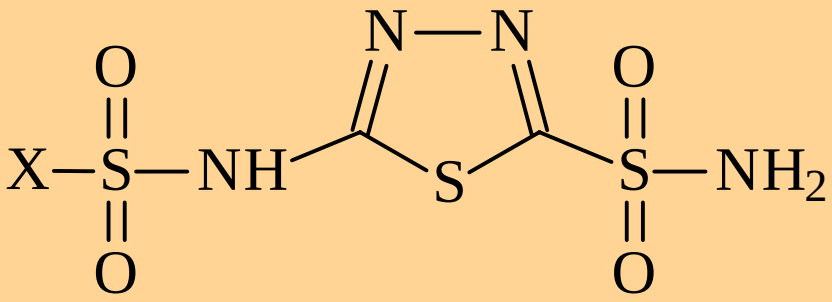
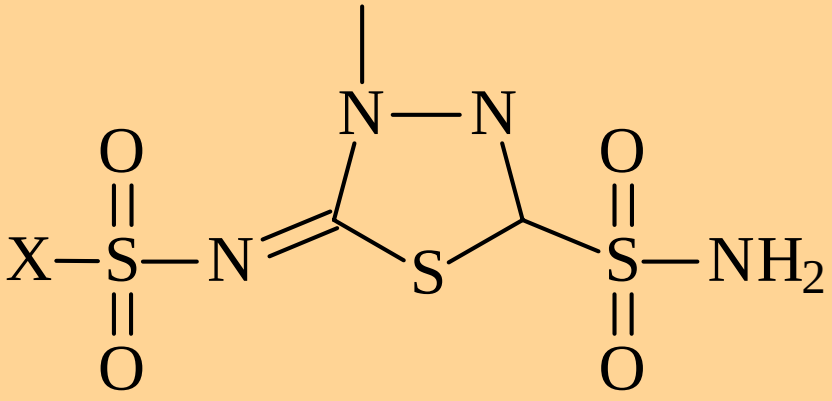
The best performing MDF-SAR model proved to be the model with four descriptors $r^2 = 0.918$. The MDF-SAR model with four descriptors shown that the inhibition on carbonic anhydrase I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides is likely to be of geometry and topology nature, being in relation with the partial charge and relative atomic mass of compounds. The estimation ability of this model is sustained by the multiple correlation coefficient ($r = 0.9579$, $^{95\%}CI = [0.9212, 0.9776]$) and by the significance of the model ($F = 97$, $p < 0.001$). The prediction ability is sustained by the cross validation leave-one-out score ($r^2_{cv\text{-}loo} = 0.8911$), the model stability ($r^2 - r^2_{cv\text{-}loo} = 0.0264$), and by the results on training versus test analysis. The model with four descriptors proved to render higher value of the correlation coefficient comparing with previously reported model ($p < 0.01$).

Previously Reported SARs (Supuran & Clare, 1999)

inhibitors	SAR	r^2	q^2	s	F	n
thi-Adiazoles (A)	$\log IC_{50} = 59.43 \cdot Q_{S1} + 0.1359 \cdot \mu_x - 0.0300 \cdot \mu_z - 0.0204 \cdot \Delta H_S + 98.87 \cdot Q_{O1} + 27.83$	0.909	0.502	0.18	28	20
thi-adiazolines (B)	$\log IC_{50} = 8.47 \cdot 10^{-3} \cdot \Pi_{yy} - 5.871 \cdot Q_{S2} - 1.787 \cdot E_H - 1.575 \cdot E_L + 0.0501 \cdot \Delta H_S - 82.31 \cdot Q_{O1} - 16.36 \cdot Q_{O2} - 182.6$	0.917	0.712	0.21	19	20



Structures

1,3,4-thiadiazole-disulfonamides (A)	1,3,4-thiadiazoline-disulfonamides (B)
	

X – substituent (one of twenty in total)

Substituent's and LogIC_{50} (nM)

Substituent X	A	B	Substituent X	A	B
Me	1.00	1.23	4-O_2N-C_6H_4	0.48	0.90
$PhCH_2$	0.85	0.78	3-O_2N-C_6H_4	0.30	0.85
4-Me-C_6H_4	0.70	0.70	2-O_2N-C_6H_4	0.70	0.70
4-F-C_6H_4	0.60	0.90	Me_2N	1.28	0.95
4-Cl-C_6H_4	0.60	0.90	2-$HO_2CC_6H_4$	0.00	0.00
4-Br-C_6H_4	0.48	0.70	4-(2,4,6-Me_3Py^+)C_6H_4	1.26	1.23
4-MeO-C_6H_4	0.70	0.78	4-(2,4,6-Ph_3Py^+)C_6H_4	2.56	2.66
4-AcNH-C_6H_4	1.00	0.30	2,4-$(O_2N)_2C_6H_3$	1.08	1.00
4-H_2N-C_6H_4	0.78	0.00	4-Cl-3-O_2N-C_6H_3	0.95	0.85
3-H_2N-C_6H_4	0.95	0.00	2,4,6-$Me_3C_6H_4$	1.18	1.11

Me = methyl; Ph = phenyl; Ac = acetyl; Py^+ = pyridine

MDF SARs (This work)

- *n = 40 (both A and B)*
- Two independent variables SAR

- $1.74 + 1.01 \cdot 10^{-1} \cdot inPRlQg + 3.10 \cdot 10^{-3} \cdot lPDMqMg$
  - Four independent variables SAR
- $1.14 + 8.79 \cdot 10^{-2} \cdot inPRlQg + 3.52 \cdot 10^{-3} \cdot lPDMoMg + 2.43 \cdot iAMRqQg + 1.04 \cdot inMRkQt$

Determination analysis

r^2		*2v*	*2v, 4v*	*4v*	*4v*	*4v*
		lPDMqMg	*inPRlQg*	*lPDMoMg*	*iAMRqQg*	*inMRkQt*
2v	lPDMqMg	1				
2v, 4v	inPRlQg	0.021	1			
4v	lPDMoMg	0.999	0.02	1		
4v	iAMRqQg	0.144	0.06	0.142	1	
4v	inMRkQt	0.310	0.02	0.312	0.5	1

Multivariate statistics

Parameter (abbreviation)	2v	4v
Number of experiments (n)	40	40
Correlation coefficient (r)	0.8975	0.9579
95% CI for r ($^{95\%}CI_r$)	0.81-0.94	0.92-0.98
Determination coefficient (r^2)	0.8056	0.9175
Standard error of estimation (s_{est})	0.2426	0.1624
Fisher parameter (F_{est})	77	97
Cross-validation leave-one-out score ($r^2_{cv\text{-}loo}$)	0.7888	0.8911
Standard error of leave-one-out analysis (s_{loo})	0.2532	0.1869
Fisher parameter of loo analysis (F_{pred})	69	71
r^2 - $r^2_{cv\text{-}loo}$	0.0167	0.0264

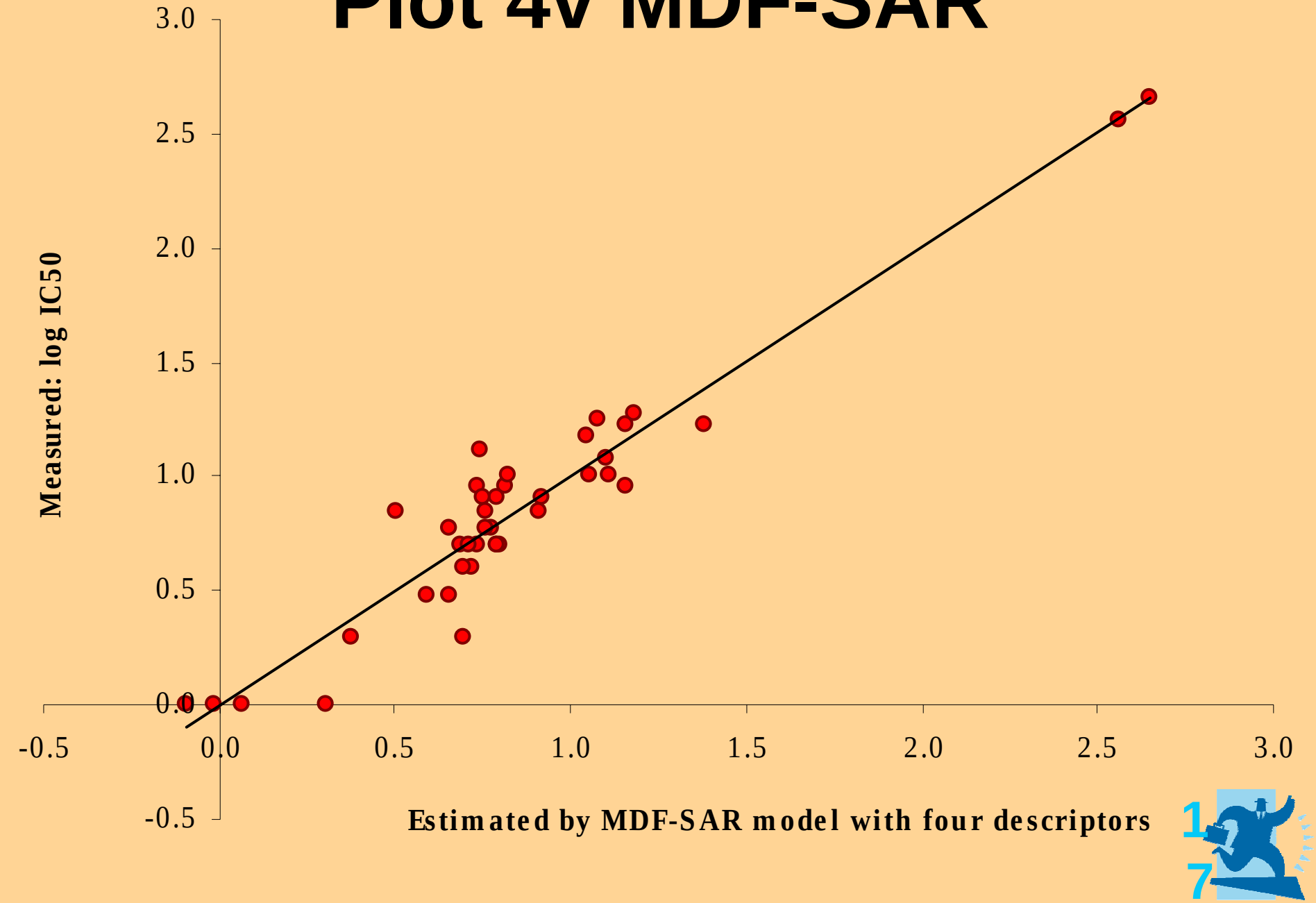
ANOVA

	StdError	$r^2(Y, desc)$	*t*	95%CI
MDF-SAR model with two descriptors				
Intercept	0.0845	n.a.*	20.618†	[1.5715-1.9140]
inPRlQg	0.0174	0.2822	5.8097†	[0.0657-0.1360]
lPDMqMg	0.0003	0.6282	9.9803†	[0.0025-0.0037]
MDF-SAR model with four descriptors				
Intercept	0.1295	n.a.*	8.7986†	[0.8768-1.4028]
inPRlQg	0.0119	0.2822	7.3752†	[0.0637-0.1121]
lPDMoMg	0.0002	0.6274	14.241†	[0.0030-0.0040]
iAMRqQg	0.3812	0.2663	6.3782†	[1.6576-3.2055]
inMRkQt	0.1663	0.1299	6.2486†	[0.7013-1.3764]

StdError = standard error; Y = log IC_{50}, desc = molecular descriptor; t = parameter of the Student test
$95\%CI_{ai}$ = 95% confidence interval associated with a_i; * n.a. = not applicable

† $p < 1‰$

Plot 4v MDF-SAR



Training versus Test Analysis

n_{tr}	a_0	a_1	a_2	a_3	a_4	r_{tr}	95% CI_{rtr}	F_{tr}	n_{ts}	r_{ts}	95% CI_{rts}	F_{ts}
20	1.257	$9.92\cdot10^{-2}$	$3.65\cdot10^{-3}$	2.209	1.1543	0.936	[0.842-0.975]	27†	20	0.972	[0.929-0.989]	60†
21	1.078	$9.01\cdot10^{-2}$	$3.58\cdot10^{-3}$	2.720	1.1522	0.961	[0.905-0.984]	49†	19	0.954	[0.881-0.982]	34†
22	0.899	$8.01\cdot10^{-2}$	$2.91\cdot10^{-3}$	2.527	0.6065	0.984	[0.961-0.993]	129†	18	0.901	[0.750-0.963]	14†
23	1.070	$8.63\cdot10^{-2}$	$3.41\cdot10^{-3}$	2.456	0.8262	0.945	[0.873-0.977]	38†	17	0.965	[0.902-0.987]	38†
24	0.690	$7.51\cdot10^{-2}$	$2.87\cdot10^{-3}$	3.317	1.2557	0.934	[0.851-0.971]	33†	16	0.942	[0.837-0.980]	18†
25	1.455	$9.16\cdot10^{-2}$	$4.46\cdot10^{-3}$	2.095	0.4501	0.892	[0.766-0.951]	19†	15	0.962	[0.881-0.988]	6‡
26	0.909	$1.06\cdot10^{-1}$	$3.09\cdot10^{-3}$	2.841	0.8498	0.945	[0.880-0.975]	44†	14	0.951	[0.848-0.985]	18†
27	1.193	$8.45\cdot10^{-2}$	$3.52\cdot10^{-3}$	2.213	0.9843	0.945	[0.882-0.975]	46†	13	0.972	[0.905-0.992]	32†
28	1.169	$8.88\cdot10^{-2}$	$3.50\cdot10^{-3}$	2.332	1.0154	0.934	[0.860-0.969]	39†	12	0.988	[0.957-0.997]	70†
29	1.088	$9.77\cdot10^{-2}$	$3.61\cdot10^{-3}$	2.694	1.1626	0.958	[0.911-0.982]	67†	11	0.966	[0.872-0.991]	13‡
30	1.094	$9.07\cdot10^{-2}$	$3.05\cdot10^{-3}$	2.167	0.9770	0.916	[0.830-0.960]	33†	10	0.976	[0.897-0.994]	15‡
31	1.205	$8.66\cdot10^{-2}$	$3.59\cdot10^{-3}$	2.221	1.0053	0.945	[0.887-0.973]	55†	9	0.981	[0.908-0.996]	21‡

† $p \leq 0.001$; ‡ $0.001 < p < 0.05$

Steiger's Test

It's a significant difference between previously reported SAR (Supuran & Clare, 1999) and MDF-SAR (This study)?

previous SARs vs. MDF-SARs	Steiger's Z	p value
previous SAR(A,5v) vs. 2v MDF-SAR(A+B)	0.582	0.2803
previous SAR(A,7v) vs. 2v MDF-SAR(A+B)	1.041	0.1489
previous SAR(A,5v) vs. 24 MDF-SAR(A+B)	2.563	0.0052
previous SAR(A,7v) vs. 2v MDF-SAR(A+B)	2.965	0.0015

2v MDF-SAR(A+B) embed 28% of previous SAR(A,5v) and 15% of previous SAR(B,7v)

4v MDF-SAR(A+B) is significantly different from both previously reported models (below 5‰ match with previous SAR(A,5v) and below 1‰ match with previous SAR(B,7v)

Remarks

- Number of predictor variables are dramatically reduced (5v, 7v) → (2v, 4v)
- 4v MDF-SARs *are significantly different* from previously reported SARS (Steiger Z); *are better* than (0.909, 0.917 < 0.9175); thus, *are significantly better*
- MDF-SARs embed more knowledge (A+B)
- MDF-SARs allow structure-activity analysis (next example for **inPRlQg** MDF member):
  - **g** from molecular geometry (vs. topology)
  - **Q** from partial charge (vs. other atomic properties)
  - **l** from elastic type force (vs. other interaction descriptor types)
  - **R** from rare model interactions (vs. other two model interaction types)
  - **P** from path based fragmentation (Diudea & all, 2000)
  - **n** from ‘smallest absolute’ fragment’s superposing method
  - **i** from inversed global structure descriptor

Conclusion

Modelling the inhibition activity on carbonic anhydrase I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides by integration of complex structure information provide stable MDF-SAR models, revealing that there is a relationship between the compounds structure and their inhibition activity on carbonic anhydrase I.

Acknowledgement

- UEFISCSU Romania (Project ET36/2005)
- Thank you for your attention

Open Access

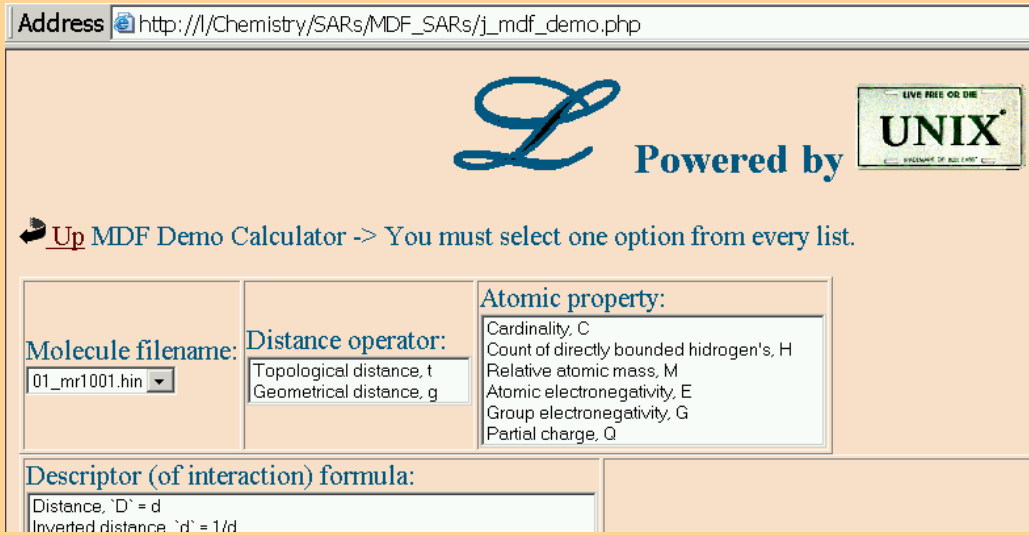
http://l.academicdirect.org

↳ Chemistry

↳ SARs

↳ MDF_SARs

↳ j_mdf_demo.php



http://lori.academicdirect.org

http://sorana.academicdirect.ro