MODELING HERBICIDAL ACTIVITY OF A SUBSTITUTED TRIAZINES CLASS BY INTEGRATION OF COMPOUNDS COMPLEX STRUCTURAL INFORMATION

S. Tigan[†1], S. Bolboaca[1], L. Jäntschi[2]

[1] *Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania;*
[2] *Technical University of Cluj-Napoca, Romania*

[†] E-mail: *stigan@umfcluj.ro*

Relationships between herbicidal activity and structure of a set of thirty 1,3,5-substituted-triazines was studied by applying an original methodology, through integration of compounds complex structural information by the use of Molecular Descriptors Family. The obtained models (mono-, bi-, tri-, and tetra-varied models) were assess and validate through the study of correlation coefficients, the cross validation leave-one-out scores, models stability defined as the differences between the squared correlation coefficient and the cross validation leave-one-out score, and in training versus test analysis. Comparison of the models with the previous reported model was performed by applying of the correlated correlation analysis. Analysis of the obtained models shows that the four-variated model, which has a squared correlation coefficient equal with 0.9885, obtains the best results. The prediction abilities of the four-varied model is justified by the cross validation leave-one-out score (0.9849), by the model stability, and by the results of training vs. test analysis ($p < 0.05$). Comparison of the correlation coefficient obtained by the four-varied model with previous reported model shows that the four varied model obtained significantly greater value ($p = 0.002$). It can be concluded that herbicidal activity of 1,3,5- substituted-triazines is of geometrical and topological nature, and is strongly dependent on the partial change and on the number of directly bonded hydrogen′s.

MODELING HERBICIDAL ACTIVITY OF A SUBSTITUTED TRIAZINES CLASS BY INTEGRATION OF COMPOUNDS COMPLEX STRUCTURAL INFORMATION

Stefan ŢIGAN*, Lorentz JÄNTSCHI, Sorana BOLBOACĂ***

**"Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania, E-mails: stigan@umfcluj.ro, sbolboaca@umfcluj.ro

*Technical University of Cluj-Napoca, Romania, E-mail: lori@academicdirect.org

Abstract

The herbicidal activity of a set of thirty 1,3,5-substituted-triazines were studied using an original structure-activity relationships methodology. The cross-validation leave-one-out correlation score, the training vs. test procedure, and the model stability sustain the prediction abilities of the optimal multi-varied molecular descriptors family on structure-activity relationships model. The comparison with the previous reported model was performed by the used of a correlated correlation analysis. The optimal MDF-SAR model shows that the herbicidal activities of studied 1,3,5-substituted-triazines are of geometrical and topological nature.

Introduction

The Quantitative Structure-Activity Relationships methods are use today for finding the link between the activity of and structure of chemical compounds in order to obtain new compounds with better properties, lowest expenses, and without time-consuming experiments [1].

Diudea & all [2] studied a sample of thirty 1,3,5-substituted-triazines by the used of topological substituent descriptors and the optimal model reported had the following characteristics:

- Squared correlation coefficient equal with 0.9694;
- Fisher parameter equal with 274.3;
- Leave-one-out score equal with 0.9778.

An original molecular descriptors family on structure-activity relationships methodology (MDF-SAR) [3] was applied on the sample of 1,3,5-substituted-triazines and the estimation and prediction abilities of the optimal multi-varied model are presented.

Materials and Methods

The inhibition activity of thirty 1,3,5-substituted-triazines on Chorella, expressed as pI50 (the negative logarithm of concentration required for 50% inhibition of Hill reaction), was taken from a previous study [2].

The steps applied in MDF-SAR modeling [3] were:

- Sketch of the thirty 1,3,5-substituted-triazines compounds
- Create the 1,3,5-substituted-triazines herbicidal activities file
- Generate the MDF members for the 1,3,5-substituted-triazines compounds
- Find the herbicidal MDF-SAR models
- Validate the MDF-SAR models
- Analyze the selected MDF-SAR model.

Results

In MDF-SAR modeling of herbicidal activities of 1,3,5-substituted-triazines, from the total possible number of molecular descriptors (787968), 298462 proved to had real and distinct values. A number of 74467 MDF members were significantly different molecular descriptors after the application of a 10^{-9} significance selector to bias the values. The MDF-SAR models which obtained values of correlation coefficients closest to one were:

- Mono-varied model:

$$Y_{1d} = 7.47 - 4284.7 \cdot iSDRFHg \quad (1)$$

- MDF-SAR model with two descriptors:

$$Y_{2d} = 5.52 - 8112.2 \cdot iSMMWHg + 194.35 \cdot iSMmEQt \quad (2)$$

- MDF-SAR model with three descriptors:

$$Y_{3d} = 1.74 - 9261 \cdot iSMMWHg + 10.34 \cdot iAMdEHg + 3.89 \cdot INDRLQg \quad (3)$$

- MDF-SAR model with four descriptors:

$$Y_{4d} = 5.75 + 199 \cdot iSMmEQt - 9010 \cdot iSMMWHg - 0.071 \cdot LADmkQt + 2.86 \cdot INPRJQg \quad (4)$$

where $Y_{1d\text{-}4d}$ are the estimators of the herbicidal activities and iSDRFHg, iSMmEQt, iSMMWHg, iAMdEHg, LADmkQt, and INPRJQg are MDF members.

The statistical characteristics of the MDF-SAR models are in table 1. The graphical representation of measured pI50 versus estimated by the MDF-SAR model with four descriptors is in figure 1.

The MFD-SAR model with four descriptors proved to have the greatest correlation coefficient. Its validation was analyzed in training versus test sets and the results obtained are presented in table 2.

Table 1. Statistical characteristics of MDF-SAR models

Characteristic	Value			
Sample size (s_{size})	30	30	30	30
Number of descriptors used by the model (n)	4	3	2	1
Correlation coefficient (r)	0.9942	0.9915	0.9875	0.9754
95% confidence interval for correlation coefficient (95%CI_r)	[0.9877, 0.9972]	[0.9820, 0.9959]	[0.9735, 0.9940]	[0.9483, 0.9883]
Squared correlation coefficient (r^2)	0.9885	0.9830	0.9753	0.9514
Adjusted squared correlation coefficient (r^2_{adj})	0.9866	0.9810	0.9734	0.9497
Standard error (s_{est})	0.081	0.096	0.1137	0.1565
Fisher parameter (F_{est})	537*	501*	533*	549*
Cross-validation leave-one-out correlation score	0.9849	0.9768	0.9712	0.9458
Fisher parameter for leave-one-out analysis	409*	361*	449*	488*
Standard error for leave-one-out analysis	0.092	0.1128	0.1236	0.1654
MDF-SAR model stability	0.0035	0.0062	0.0041	0.0056

*$p < 0.0001$

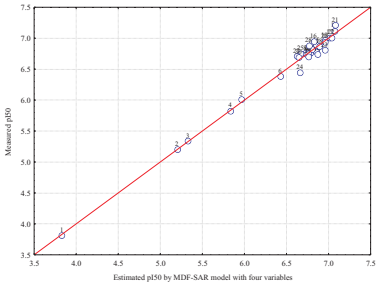
Table 2. The results obtained in training versus test analysis

Training set				Test set				
No_{tr}	r_{tr}	95% CI_{tr}	F_{tr}	No_{ts}	r_{ts}	95%CI_{rts}	F_{ts}	$Z_{rtr\text{-}rts}$
10	0.9474	[0.7873, 0.9877]	11*	20	0.9972	[0.9928, 0.9989]	18**	3.313*
11	0.9946	[0.9787, 0.9986]	138**	19	0.9929	[0.9812, 0.9973]	245**	0.321
12	0.9975	[0.9909, 0.9993]	353**	18	0.9839	[0.9562, 0.9941]	90**	2.242*
13	0.9957	[0.9851, 0.9987]	230**	17	0.9854	[0.9590, 0.9948]	79**	1.476
14	0.9937	[0.9796, 0.9980]	177**	16	0.9925	[0.9779, 0.9974]	143**	0.216
15	0.9946	[0.9832, 0.9982]	230**	15	0.9933	[0.9792, 0.9978]	126**	0.266
16	0.9951	[0.9856, 0.9983]	280**	14	0.9760	[0.9238, 0.9925]	38**	1.96*
17	0.9882	[0.9667, 0.9958]	125**	13	0.9945	[0.9812, 0.9984]	168**	0.93
18	0.9973	[0.9927, 0.9990]	612**	12	0.9861	[0.9494, 0.9962]	44**	1.975*
19	0.9966	[0.9910, 0.9987]	516**	11	0.9850	[0.9412, 0.9962]	43**	1.737*
20	0.9530	[0.8827, 0.9815]	37**	10	0.9974	[0.9888, 0.9994]	108**	3.268*
21	0.9940	[0.9849, 0.9976]	330**	9	0.9950	[0.9753, 0.9989]	55**	0.194
22	0.9905	[0.9768, 0.9961]	220**	8	0.9960	[0.9770, 0.9993]	69*	0.861
23	0.9946	[0.9870, 0.9977]	411**	7	0.9937	[0.9559, 0.9991]	21*	0.142

Abbreviations: number of compounds in training (No_{tr}) and test (No_{ts}) sets; correlation coefficient in training (r_{tr}) and test (rts) sets and associated 95% CI; Fisher parameter in training (F_{tr}) and test (F_{ts}) sets; Fisher Z-test parameter of comparison between correlation coefficient obtained in training set and the correlation coefficient obtained in test set ($Z_{rtr\text{-}rts}$)

** $p < 0.001$; * $0.05 < p < 0.001$

Figure 1. Estimated versus measured pI50



By applying of an correlated correlation analysis, the correlation coefficient obtained by the optimal MDF-SAR model (the model with four descriptors) proved to be statistical significant higher comparing with the correlation coefficient obtained by the previous reported QSAR model (Steiger's Z test parameter = 2.828, $p < 0.01$).

Conclusions

The herbicidal activities of 1,3,5-substituted-triazines proved to be of geometrical and topological nature, and strongly depended on partial charges and number of directly bonded hydrogen's.

Even if the MDF-SAR methodology is a time-consuming method it could be consider that open a new pathway in studying of herbicidal activities of 1,3,5-substituted-triazines, providing an virtual experimental environment.

Future studies are necessary in order to assess the influence of compounds number 1, 2, 3, 4 and 5 on the stability of the MDF-SAR model with four descriptors by the use of new 1,3,5-substituted-triazines with herbicidal activities.

Acknowledgement

Research was partly supported by UEFISCSU Romania through the project ET36/2005.

References

[1] W. J. G. M. Peijnenburg: Pure & Appl. Chem. 66, 1931-1941 (1994).
[2] V.M. Diudea, L. Jäntschi, and L. Pejov: Leonardo Electronic Journal of Practices and Technologies 1, 1-18 (2002).
[3] L. Jäntschi: Leonardo Electronic Journal of Practices and Technologies 6, 76-98 (2005).

XXIII International Biometric Conference
July 16 - 21, 2006 Montréal, Québec Canada