QSAR MODELING ON FUNGICIDAL ACTIVITY OF MANNICH BASES

Simona Funar-Timofei[1], Ana Borota[1], Alina Bora[1], Sorin Avram[1], Daniela Ionescu[2]

[1]Institute of Chemistry of Romanian Academy, Mihai Viteazul Bvd. 24, 300223, Timisoara, Romania

[2]University of Medicine and Pharmacy, Faculty of Pharmacy, P-ta Eftimie Murgu 2-4, 300034 Timisoara, Romania

ABSTRACT

In this study the fungicidal activity, expressed as relative inhibition rate, was correlated with trifluoromethyl-1,2,4-triazole derivative descriptors by the Partial Least Squares (PLS) method. For a data set containing 18 structures, previously modeled by the RM1 semiempirical quantum chemical method, various electronic and 0D, 1D, 2D and 3D descriptors were calculated. The test set includes five out of the eighteen Mannich bases containing trifluoromethyl-1,2,4-triazoles. The resultant two-components PLS model had acceptable statistical quality ($R^2X(Cum)$ = 0.805, $R^2Y(cum)$ = 0.823 and $Q^2(Cum)$ = 0.735) for predicting the fungicidal activity of the 1,2,4-triazole derivatives. Specific 1,2,4-triazole structural features supplying information about interatomic distances, topological distances, types of atoms and which encode chemical information influence the fungicidal activity.

Keywords: QSPR, fungicide, PLS, 1,2,4-triazoles, model validation

INTRODUCTION

1,2,4-triazole and its derivatives represent a versatile class of biologically active compounds, possessing a wide spectrum of activities, including anti-inflammatory, antiviral, analgesic, antimicrobial, anticonvulsant, anticancer, antioxidant, antitumoral and antidepressant activity [1]. Furthermore, some of the complexes containing 1,2,4-triazole ligands have rather peculiar structures and specific magnetic properties.

The 1, 2, 4-triazole nucleus has been incorporated into a wide variety of therapeutically interesting drug candidates [2]. Some of the modern day drugs having fused heterocycles with a triazole moiety are alprazolam, triazolam, estazolam (hypnotic, sedative, tranquilizer), trazodone

(antidepressant, anxiolytic), trapidil (hypotensive), terconazole (antifungal), hexaconazole (antifungal), etizolam (amnesic, anxiolytic, anticonvulsant, hypnotic, sedative and skeletal muscle relaxant), rilmazafon (hypnotic, anxiolytic) and rizatriptan (antimigrane agent).

Mannich bases, especially candidates with morpholine, 4-benzylpiperazine, N-methylpiperidine and trifluoromethylphenylpiperazine in the aminomethyl group, have very good antibacterial activity [3, 4].

Triazoles are used in the control of a variety of fungal diseases in fruits, vegetables, legumes and grain crops, both as pre- and postharvest applications [5]. The biochemical mechanism of their antifungal effect is based on the inhibition of ergosterol biosynthesis thereby interfering with fungal cell-wall formation. They also inhibit sterol 14α–demethylase and hence considered steroid demethylation inhibitor. 3- amino-1,2,4-triazole is an inhibitor of mitochondrial and chloroplast function. Commercial grade 3-amino-1,2,4-triazole is used as a herbicide and cotton defoliant. The triazole derivatives such as, S-3307, S-3308, triadimefon and paclobutrazol are recommended for use both as fungicides and plant growth regulators.

Molecules containing thiazole ring systems have low toxicity and very good biological activity [6].

A series of trifluoromethyl-substituted 1,2,4-triazole Mannich bases containing substituted benzylpiperazine ring with herbicidal and fungicidal activity have been synthesized [7] (Table 1).

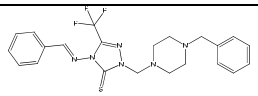
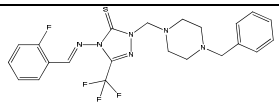
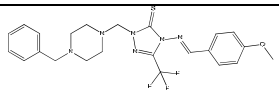
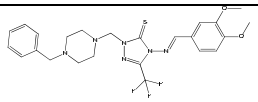
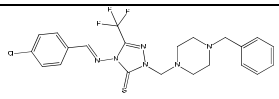
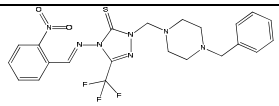
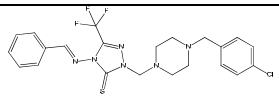
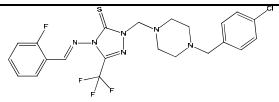
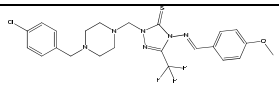
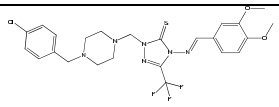
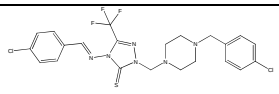
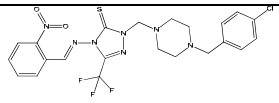
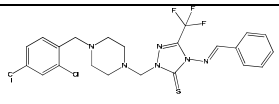
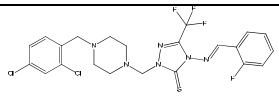
This paper presents a quantitative structure-activity relationships study for this series of 1-[(4-substituted-benzylpiperazin-1-yl)methyl]-4-(substituted)benzylideneamino-3-trifluoromethyl-1H-1,2,4-triazole-5(4H)-thiones, which were previously modeled [8] using the RM1 semiempirical molecular orbital method [9]. Descriptors calculated for the optimized structures were related to the mycelial growth inhibition activity against the *Fusarium oxysporum f. sp. cucumerinum* fungi test [7] using the partial least squares (PLS) approach.

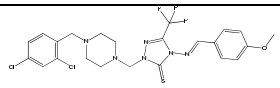
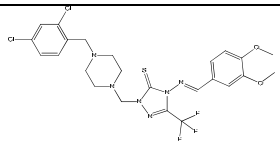
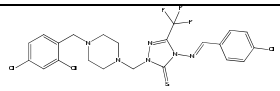
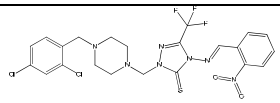
METHODS

Definition of target property and molecular structures

A series of 18 Mannich bases having trifluoromethyl-substituted 1,2,4-triazole containing substituted benzylpiperazine ring (Table 1) was used, having the fungicidal *Fusarium oxysporum f. sp. Cucumerinum* relative inhibition rate (RIR, expressed in %) as dependent variable.

Table 1. Mannich bases structures including trifluoromethyl-substituted 1,2,4-triazoles, *Fusarium oxysporum f. sp. Cucumerinum* experimental relative inhibition rate (RIR) and descriptors used in the final PLS model*

No	Structure	RIR	F10[C-F]	Mor26e	R7e+	RDF020e	RDF020p	RDF020u	RDF020v
1		0.101	3.00	-0.39	0.02	5.04	1.35	4.48	1.15
2		0.804	6.00	0.00	0.02	8.74	1.71	7.31	1.49
3		0.187	3.00	-0.43	0.02	4.60	1.17	4.02	1.01
4		0	3.00	-0.46	0.02	4.94	1.21	4.21	1.06
5		0	3.00	-0.57	0.02	4.47	1.13	3.89	0.98
6		0.402	5.00	-0.24	0.02	4.27	1.05	3.68	0.91
7		0.509	3.00	-0.24	0.02	6.53	1.40	5.53	1.22
8		0.719	6.00	-0.49	0.02	8.41	1.58	6.97	1.38
9		0.604	5.00	-0.20	0.03	4.89	1.29	4.32	1.10
10		0.401	3.00	-0.11	0.02	6.70	1.47	5.71	1.27
11		0.303	3.00	-0.53	0.02	5.56	1.35	4.86	1.17
12		0.502	3.00	-0.30	0.02	5.71	1.34	4.86	1.18
13		0.708	3.00	-0.29	0.02	8.09	1.81	7.00	1.57
14		0.826	6.00	0.07	0.03	9.84	1.97	8.33	1.71

15		0.504	3.00	-0.34	0.03	5.68	1.42	5.00	1.23
16		0.705	3.00	0.05	0.02	6.28	1.55	5.46	1.35
17		0.607	3.00	-0.10	0.03	5.83	1.47	5.16	1.27
18		0.608	5.00	-0.35	0.03	5.75	1.41	5.03	1.22

* F10[C-F] represents the frequency of C-F at topological distance 10 (2D frequency fingerprints descriptor), Mor26e – the 3D-MoRSE - signal 26 / weighted by atomic Sanderson electronegativities (3D-MoRSE descriptor), R7e+ - the R maximal autocorrelation of lag 7 / weighted by atomic Sanderson electronegativities (GETAWAY descriptor), RDF020e – the Radial Distribution Function - 2.0 / weighted by atomic Sanderson electronegativities (Radial Distribution Function descriptor), RDF020p – the Radial Distribution Function - 2.0 / weighted by atomic polarizabilities (Radial Distribution Function descriptor), RDF020u – the Radial Distribution Function - 2.0 / unweighted (Radial Distribution Function descriptor), RDF020v – the Radial Distribution Function - 2.0 / weighted by atomic van der Waals volumes (Radial Distribution Function descriptor).

The title fungicides were previously energy optimized [8] by the RM1 semiempirical quantum chemical approach, performed using semiempirical NDDO module of the Schrödinger software (Schrödinger, LLC, New York, NY, 2008). Following quantum chemical descriptors were derived for the energy optimized structures: electronegativity, hardness, chemical potential, electrophilicity, HOMO and LUMO molecular orbital energies, heat of formation, dipole moment, molecular surface area, softness, maximum average local ionization energy on the molecular surface, minimum average local ionization energy on the molecular surface, mean average local ionization energy on the molecular surface, maximum electrostatic potential on the molecular surface, minimum electrostatic potential on the molecular, mean electrostatic potential on the molecular surface, electrophilic superdelocalizability, nucleophilic superdelocalizability, radical superdelocalizability, atom self polarizability.

Structural 0D, 1D, 2D and 3D descriptors were calculated for the lowest energy structures using the DRAGON (Dragon Professional 5.5 (2007), Talete S.R.L., Milano, Italy), InstantJchem (which was used for structure database management, search and prediction) (InstantJchem 15.7.27, 2015, ChemAxon (http://www.chemaxon.com) and ChemProp (UFZ Department of Ecological Chemistry 2014. ChemProp 6.2, http://www.ufz.de/index.php?en=6738) software. Structural 0D,

1D, 2D and 3D descriptors in conjunction with the electronic parameters were used for QSAR models development.

The Partial Least Squares (PLS) method

Projections to Latent Structures (PLS) represent a regression technique for modeling the relationship between projections of dependent factors and independent responses. PLS (Partial Least Squares) regression is a statistical modeling technique with data analysis features linking a block (or a column) of response variables to a block of explanatory variables [10]. The PLS approach leads to stable, correct and highly predictive models even for correlated descriptors [11].

This method describes the **X** matrix of chemical descriptors of the training set (N compounds) defining a number of F significant principal components (PC), i.e. t_{if} columns formed by equations (1), when i = 1, ..., N.

$$x_{ik} = \bar{x}_k + \sum_{f=1}^{F} p_{fk} \cdot t_{if} + e_{ik} \qquad (1)$$

where $\bar{x}_k$ denotes the mean of variable k, p_{fk} the loading of variable k in dimension (factor) f, and e_{ik} the residuals [12]. The consecutive orthogonal latent variables (t_f) are deduced assuring maximal covariance of them with y. The linear PLS inner relation is described by equation (2):

$$y_i = \bar{y} + \sum_{f=1}^{F} b_f \cdot t_{if} + e_i \qquad (2)$$

where $\bar{y}$ represents the average of the y-variable and b_f the regression coefficients. These can be transformed to express the biological activity y as a function of the original x_k descriptors.

PLS calculations were performed by the SIMCA package (SIMCA P+ 12.0.0.0, May 20 2008, Umetrics, Sweeden, http://www.umetrics.com/).

Model validation

The data over fitting and model applicability was controlled by comparing the root-mean-square errors (RMSE) and the mean absolute error (MAE) [13] of the training and validation sets.

For internal validation results several measures of robustness were employed: Y-scrambling [14] and Q^2 (leave-seven-out) cross-validation coefficient.

To test the robustness and predictive power of the model the concordance correlation coefficient (CCC) [15] (having the thresholds values higher than 0.85, as they have been rigorously determined by a simulation study [16]).

In addition, to test the predictive power of the model, the predictive r^2 (r^2_{pred}) test [17] was employed. It is considered that for a predictive QSAR model, the value of r^2_{pred} should be higher than 0.5.

RESULTS AND DISCUSSION

The series of 18 trifluoromethyl-1,2,4-triazole derivatives was previously modeled [8] by the RM1 semiempirical molecular orbital approach and further used to calculate structural descriptors. The trifluoromethyl-1,2,4-triazole derivative structures and the descriptors used in the final PLS model are presented in Table 1.

A PCA model was built for the entire **X** matrix (including $N = 18$ compounds and $X = 1494$ descriptors). From five significant principal components resulting from this analysis, the first three components already explained 65.7 % of the information content of the descriptor matrix.

PLS calculations were performed to correlate the RIR values with all the calculated descriptors. From the entire set of trifluoromethyl-1,2,4-triazole derivative, a training set of 13 compounds and a random test set of 5 compounds: no. 4, 8, 11, 14, 16 (Table 1) were considered.

The PLS model was constructed using the training set and an acceptable PLS model with 2 principal components was obtained: $R^2X(Cum) = 0.805$, $R^2Y(cum) = 0.823$, $Q^2(Cum) = 0.735$, where $R^2Y(CUM)$ are the cumulative sum of squares of the entire Y's explained by all extracted principal components and $Q^2(CUM)$ is the fraction of the total variation of the Y's that can be predicted for all the extracted principal components.

From the entire descriptor matrix following descriptor types were included in the final PLS model: 2D frequency fingerprints, 3D-MoRSE, GETAWAY and Radial Distribution Function, which were found to be significant for the model (Table 1). The score and loading plots are presented in Figures 1 and 2.

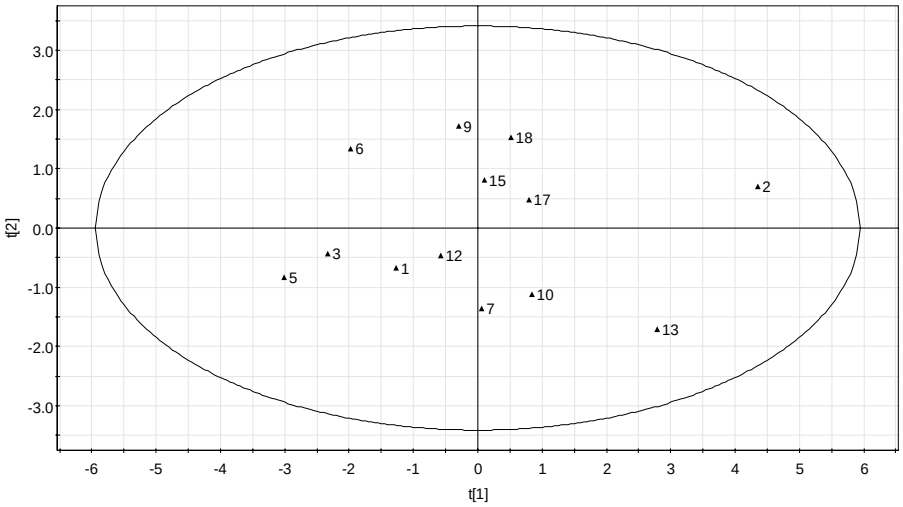


Figure 1. Score scatter plot of the final PLS model

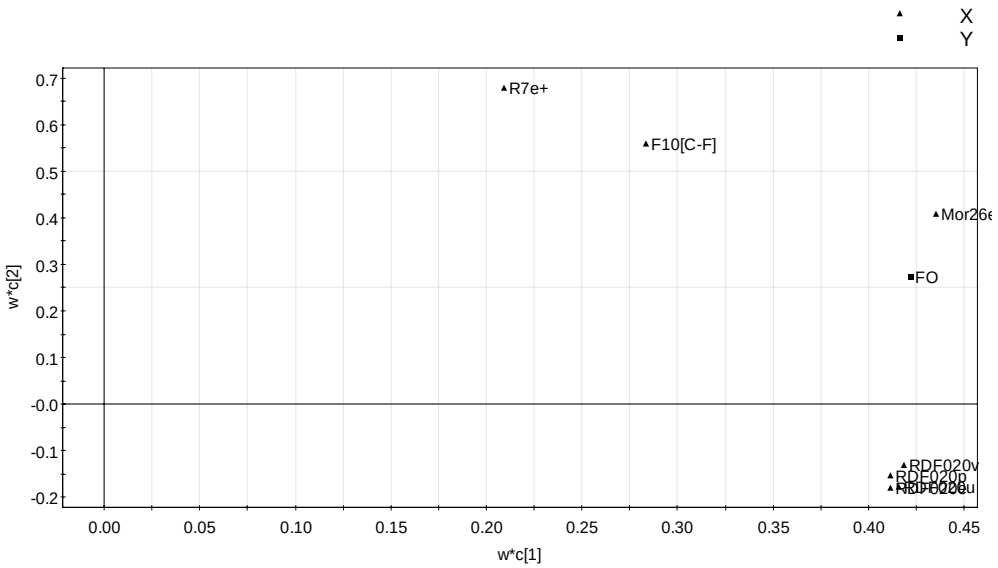


Figure 2. Loading scatter plot of the final PLS model.

The loading scatter plot confirms that the first component is dominated by the 3D-MoRSE and RDF descriptors and the GETAWAY and 2D-frequency fingerprint descriptors are dominant for the second component.

The importance of a given x variable for the **Y** matrix is proportional to its distance from the origin in the loading space. These lengths correspond to the PLS regression coefficients after F = 2 dimensions. The importance of descriptors was evaluated by the VIP (Variable Influence on Projection) values [18], which summarizes the importance of the x variables for both **Y** and **X** matrices in the model. This is a weighted sum of squares of the PLS-weights, with the weights

calculated from the amount of Y variance of each PLS component. The noise variables (variable coefficient values close to 0) were excluded to reduce the model over fit. The PLS coefficients and VIP plots are presented in Figures 3 and 4.

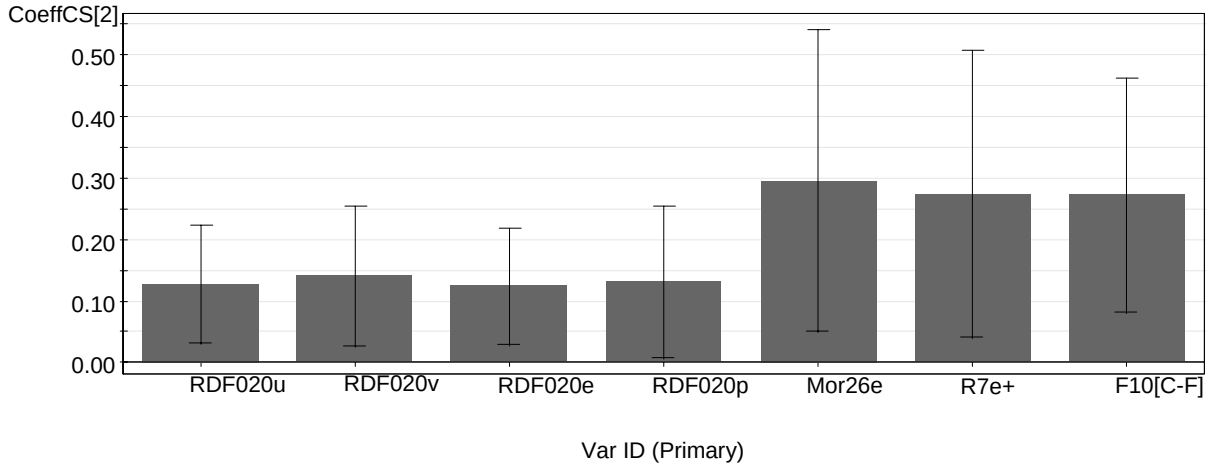


Figure 3. PLS regression coefficients plot of the two-components PLS model. The bars indicate 95% confidence intervals based on jack-knifing.

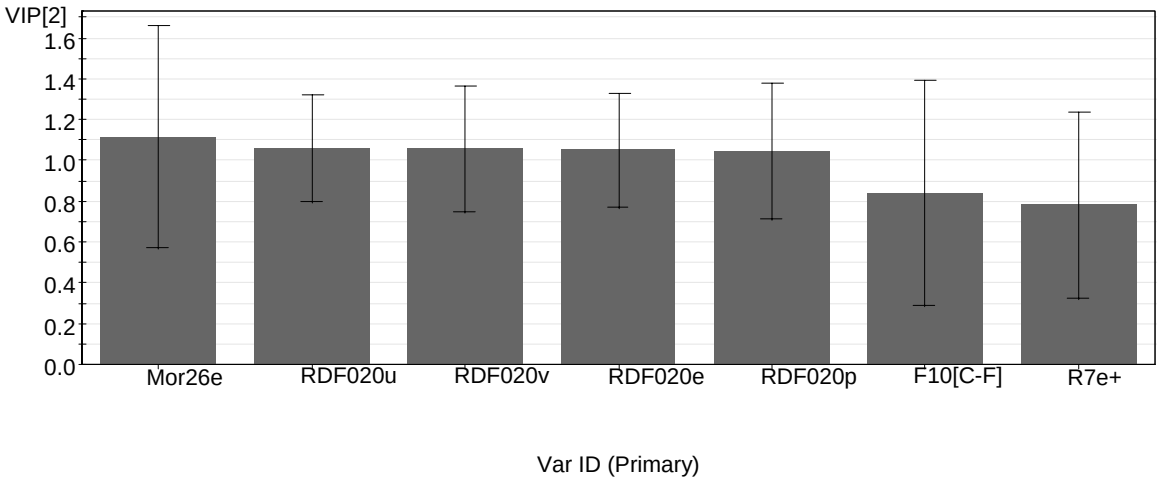


Figure 4. VIP plot of the x-variables of the two-component PLS model.

Y-randomization test and leave-seven-out crossvalidation runs were performed to check the robustness and internal predictive ability of the PLS models. The risk of chance correlation was verified by the Y-scrambling procedure, which was repeated 999 times. The extremely low calculated scrambled R^2 (0.158) and Q^2 (-0.346) values indicate no chance correlation for the chosen model. The Y-scrambling plot is presented in Figure 5.

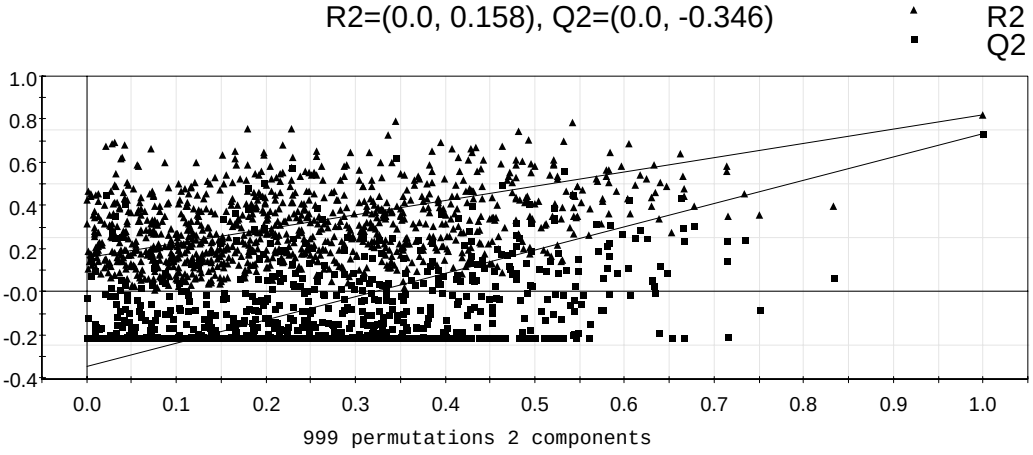


Figure 5. Y-randomization results for the final PLS model. The *x*-axis reports the correlation coefficient between original and permuted response data, while on the *y*-axis R^2 (black triangles) and Q^2 (black squares) values for the 999 randomized models are reported.

The data over fitting and model applicability was controlled by comparing the root-mean-square errors (RMSE) and the mean absolute error (MAE) [13] calculated for the training ($RMSE_{tr}$ = 0.096, MAE_{tr} = 0.079) and validation ($RMSE_{ext}$ = 0.178, MAE_{ext} = 0.140) sets.

The leave-seven-out crossvalidation results highlights that the model is stable, not obtained by chance: the difference between $R^2Y(cum)$ and $Q^2(cum)$ is small: 8.8 % and the calculated RMSE and MAE values indicate an internally predictive model.

The calculated concordance correlation coefficient values for the training (CCC_{tr} = 0.903), crossvalidation (CCC_{L7O} = 0.832) and test (CCC_{ext} = 0.853) sets indicate a robust model with predictive power, which was confirmed by the r^2_{pred} value of 0.681 too.

The PLS model is satisfactory in the fitting and has predictive power. The dependence between the experimental and predicted RIR values is presented in Figure 6.

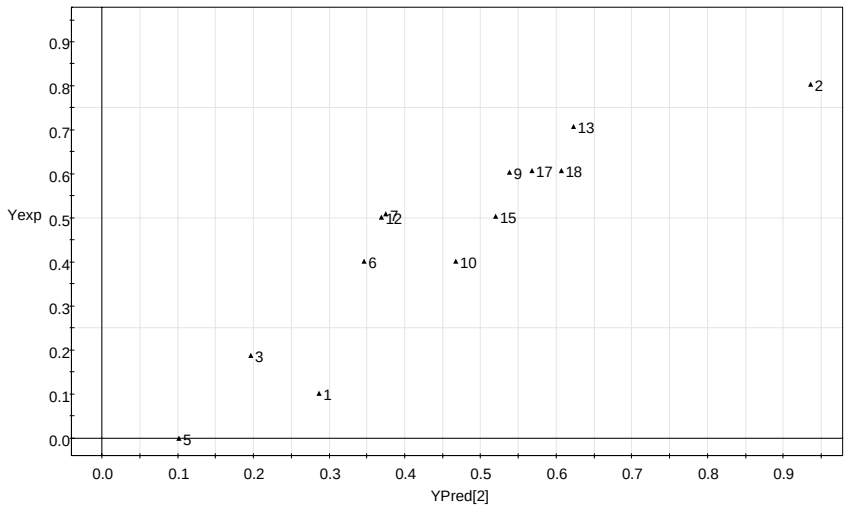


Figure 6. Experimental *versus* predicted RIR values obtained by the PLS model.

The Table 1, Figures 3 and 4 describe seven descriptors as the most relevant variables for the present set of 1,2,4-triazoles which influence the fungicidal activity. Regarding the selected descriptors used to build the PLS model, some considerations can be pointed out:

GETAWAY (Geometry, Topology, and Atom-Weights AssemblY, descriptor R7e+ in this case) are geometrical descriptors which encode geometrical and topological information on the effective position of substituents and fragments in the molecular space [19]. They are independent of molecule alignment and account also for information on molecular size and shape as well as for specific atomic properties. Thereby, R7e+ is weighted by Sanderson electronegativities with positive contribution to the model (Table 2).

The 3D-MoRSE (3D-Molecule Representation of Structures based on Electron diffraction, descriptor Mor26e in this case) descriptors afford the possibility for choosing an appropriate atomic property [20]. They present the advantage that they code with fixed-length representation of 3D molecular structure, allowing the comparison of data sets comprising molecules of different size, and number of atoms.

The RDF (radial distribution function, descriptors RDF020e, RDF020p, RDF020u and RDF020v in this case) descriptors are calculated from the molecular geometry, representing the molecular conformation in 3D with a series of weighting schema, including weighted by atomic masses, atomic van der Waals volumes, atomic Sanderson electronegativities and atomic polarizabilities. These descriptors are based on the geometrical interatomic distance and constitute a radial distribution function code [21]. These atomic properties enable the discrimination of the atoms of a molecule for almost any property that can be attributed to an atom. All RDF descriptors

had a positive influence on the studied property, generally increasing the fungicidal activity of 1,2,4-triazole analogues.

Most frequently used in drug research are structural keys and hashed fingerprints (F10[C-F], in the present case) [22]. The former make use of a predefined fragment dictionary and record the presence or absence of a number of small generic or specific fragments, whereas hashed fingerprints allow to reduce the length of the bitstring. In fact, a 'molecular fingerprint' is a binary bitstring of 1's (for the presence) and 0's (for the absence) of specific fragments and the use of hash functions allows to break the fingerprint into smaller strings, easier to handle. The main advantage of the 2D-descriptors is the speed of their calculation, but their meaning is more difficult to be interpreted. Atomic, atom-type and total non-stochastic quadratic indices have shown a great ability to encode chemical information, which can be used for the development of QSARs [23]. The frequency (F10[C-F]) of fragments that have electron negative atom at a topological distance of ten bonds display major positive contributions to the final PLS model, also highlights the role of electrostatic interactions for 1,2,4-triazole activity.

In the final PLS model the selected molecular descriptors capture 3D information (3D-Morse, GETAWAY, RDF), supplying information about interatomic distances, topological distances, types of atoms and which encode chemical information (2D-frequency fingerprints).

CONCLUSIONS

The fungicidal activity of Mannich bases containing trifluoromethyl-1,2,4-triazole, in terms of relative inhibition rate, was correlated with structural descriptors by the Partial Least Squares (PLS) method. The title compounds were previously modeled by quantum mechanics and 0D, 1D, 2D and 3D descriptors were derived from the optimized structures. An acceptable PLS model with predictive power was obtained. The fungicidal activity can be raised by molecular conformation in 3D descriptors weighted by atomic van der Waals volumes, atomic Sanderson electronegativities and atomic polarizabilities, geometrical descriptors referring to the effective position of substituents and fragments in the molecular space.

ACKNOWLEDGEMENT

This project was financially supported by the Project No. 1.1 of the Institute of Chemistry of Romanian Academy, Timisoara. The authors are indebted to the Chemaxon Ltd. and Prof. Gerrit Schüürmann from Helmholtz Centre for Environmental Research (UFZ, Leipzig, Germany) for giving access to their software.

REFERENCES

1. M. Koparir, C. Orek, P. Koparir, K. Sarac, Synthesis, experimental, theoretical characterization and biological activities of 4-ethyl-5-(2-hydroxyphenyl)-2H-1,2,4-triazole-3(4H)-thione, *Spectrochim. Acta A*, 2013, *105*, 522–531.

2. J. K. Sahu , S. Ganguly, A. Kaushik, Triazoles: A valuable insight into recent developments and biological activities, *Chin. J. Nat. Med.*, 2013, *11(5)*, 0456−0465.

3. G. Roman, Mannich bases in medicinal chemistry and drug design, *Eur. J. Med. Chem.*, 2015, *89*, 743-816.

4. M. Koparir, C. Orek, A. E. Parlak, A. Söylemez, P. Koparir, M. Karatepe, S. D. Dastan, Synthesis and biological activities of some novel aminomethyl derivatives of 4-substituted-5-(2-thienyl)-2,4-dihydro-3H-1,2,4-triazole-3-thiones, *Eur. J. Med. Chem.*, 2013, *63*, 340-346.

5. S.S. Kumar, H.P. Kavitha, Synthesis and Biological Applications of Triazole Derivatives - A Review, *Mini-Rev. Org. Chem.*, 2013, *10(1)*, 40-65.

6. X. Qin, H. B. Yu, H. Dai, Z. F. Qin, X. Zhang, G. F. Bing, T. T. Wang, J. X. Fang, Synthesis and plant-growth regulatory activities of novel imine derivatives containing 1H-1,2,4-triazole and thiazole rings, *Chinese Chem. Lett.*, 2010, *21*, 283–286.

7. B.–L. Wang, X.–H. Liu, X.–L. Zhang, J.–F. Zhang, H.–B. Song, Z.–M. Li, Synthesis, Structure and Biological Activity of Novel 1,2,4-Triazole Mannich Bases Containing a Substituted Benzylpiperazine Moiety, *Chem. Biol. Drug. Des.* 2011, *78*, 42–49.

8. S. Funar-Timofei, A. Borota, A. Bora, R. Curpan, S. Avram, Modeling of Mannich bases fungicidal activity by the MLR approach, *21st International Symposium on Analytical and Environmental Problems (ISAEP)*, Szeged, Hungary, 28 September, 2015.

9. G. B. Rocha, R. O. Freire, A. M. Simas, J. J. P. Stewart, RM1: a Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I, *J. Comput. Chem.* 2006, *27(10)*, 1101-1111.

10. H. Wold, Partial Least Squares, in: S. Kotz and N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences* (Vol. 6), Wiley, New York, 1985, pp. 581-591.

11. A. Höskuldsson, PLS regression methods, *J. Chemometrics,* 1988, *2(3)*, 211-228.

12. S. Hellberg, S. Wold, W.J. Dunn III, J. Gasteiger, M.G. Hutchings, The Anesthetic Activity and Toxicity of Halogenated Ethyl Methyl Ethers, a Multivariate QSAR Modelled by PLS, *Quant. Struct.-Act. Relat.* 1985, *4*, 1-11.

13. P. Gramatica, In: Reisfeld B, Mayeno AN, editors. Computational Toxicology, Volume II, *Methods in Molecular Biology*, Vol. 930, "On the Development and Validation of QSAR Models", Springer, 2013, pp. 499-526.

14. R. Todeschini, V. Consonni, A. Maiocchi, The *K* correlation index: theory development and its application in chemometrics, *Chemometr. Intell. Lab*., 1999, *46*, 13-29.
15. N. Chirico, P. Gramatica, Real External Predictivity of QSAR Models: How To Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient, *J. Chem. Inf. Model*. 2011, *51*, 2320-2335.
16. N. Chirico, P. Gramatica, Real External Predictivity of QSAR Models. Part 2. New Intercomparable Thresholds for Different Validation Criteria and the Need for Scatter Plot Inspection, *J. Chem. Inf. Model*. 2012, *52*, 2044−2058.
17. P. P. Roy, S. Paul, I. Mitra, K. Roy, On Two Novel Parameters for Validation of Predictive QSAR Models, *Molecules* 2009*, 14,* 1660-1701.
18. L. Eriksson, E. Johansson, N. Kettaneh-Wold, S. Wold, *Multi- and Megavariate Data Analysis. Principles and Applications*, Umetrics AB, Umeå, Sweden, 2001, pp. 94-97.
19. R. Todeschini, V. Consonni, *Molecular Descriptors for Chemoinformatics*, Volumes I & II, WILEY-VCH, Weinheim, 2009.
20. L. Saíz-Urra,M. Pérez González, M. Teijeira, QSAR studies about cytotoxicity of benzophenazines with dual inhibition toward both topoisomerases I and II: 3D-MoRSE descriptors and statistical considerations about variable selection, *Bioorgan. Med. Chem.,* 2006, *14*, 7347–7358.
21. R. Todeschini, V. Consonni, In: J. Gasteiger, editor. *Handbook of Chemoinformatics: From Data to Knowledge in 4 Volumes*, Part VIII, Chapter VIII.2, "Descriptors from Molecular Geometry", Wiley-VCH, Weinheim, Germany, 2003, pp. 1004-1033.
22. S. Sciabola, E. Carosati, L. Cucurull-Sanchez, M. Baroni, R. Mannhold, Novel TOPP descriptors in 3D-QSAR analysis of apoptosis inducing 4-aryl-4H-chromenes: Comparison versus other 2D- and 3D-descriptors, *Bioorgan. Med. Chem.*, 2007,*15*, 6450–6462.
23. A. Montero-Torres, M. Celeste Vega, Y. Marrero-Ponce, M. Rolón, A. Gómez-Barrio, J. A. Escario, V. J. Arán, A. R. Martínez-Fernández, A. Meneses-Marcel, A novel non-stochastic quadratic fingerprints-based approach for the 'in silico' discovery of new antitrypanosomal compounds, *Bioorgan. Med. Chem.*, 2005, *13*, 6264–6275.