Study of Ecdysone-Agonist-Based Insecticidal Activity of Dibenzoylhydrazine Derivatives by Computational Approaches†

Takahiro SUZUKI,[1] Simona FUNAR-TIMOFEI,[2] Alina BORA,[2]
Luminita CRISAN,[2] Ana BOROTA[2]

Abstract

The ecdysteroid receptor is an important potential target site for the development of more specific and selective insecticides. In this study the ecdysone agonistic activity, based on an ecdysone-dependent reporter assay using cell lines derived from one lepidopteran species (the cotton leafworm *Spodoptera littoralis*), of a series of dibenzoylhydrazine derivatives, was modeled using the multiple linear regression (MLR) approach. The dibenzoylhydrazine structures were first energy pre-optimized by molecular mechanics calculations using the MMFF94s force field. 0D, 1D, 2D and 3D descriptors were calculated for the minimum energy conformers and were related to the experimental ecdysone agonistic activity, expressed as pEC_{50} values, using MLR calculations. Genetic algorithm was used for variable selection. Seven out of the total number of thirty three compounds were included in the test set. The more stable and predictive MLR model had the following statistical parameters: $r^2_{training} = 0.797$, $r^2_{test} = 0.753$, $q^2_{LOO} = 0.711$, $RMSE_{tr} = 0.499$, $RMSE_{ext} = 0.449$, $r^2_{adj} = 0.758$. Geometric dibenzoylhydrazine structural features influence the insecticidal activity.

Keywords : Ecdysone agonistic activity, Dibenzoylhydrazines, Insecticide, MLR, Omega, QSARINS

1. Introduction

The ecdysteroid receptor is an significant potential target site for more specific and selective insecticides (Harmatha et al, 2002). Nonsteroidal bisacylhydrazines were found

† Dedicated to the 150th anniversary of the Romanian Academy

1 Natural Science Laboratory, Toyo University, 5–28–20 Hakusan, Bunkyo-ku, Tokyo 112–8606, JAPAN

2 Institute of Chemistry Timisoara of the Romanian Academy, 24 Mihai Viteazul Blvd., RO–300223 Timişoara, Romania

to effective as agonists of the ecdysteroid receptor complex, being successful insect control agents. These compounds contribute to the efforts on the reduction effects on other animal groups and on lowering the environmental millstone (Dinan, 2003). The ecdysteroids themselves have very limited application as pesticides, as they are too polar, too complex and metabolically and environmentally too inconstant. Therefore ecdysteroid analogues (steroidal or non-steroidal) were studied as potent pesticides for trading applications.

Dibenzoylhydrazine derivatives are insect growth regulators that operate through the induction of an early and lethal larval molting process in susceptible insects that belong to the species of Lepidoptera and Coleoptera (Swevers et al, 2008). It was reported (Soin et al, 2009) the importance of the unusual high affinity for the ecdysone receptor of lepidopteran insects of dibenzoylhydrazine non-steroidal ecdysone agonists. The ecdysone receptor belongs to the large class of nuclear receptors which are known to act as transcription factors controlling a wide range of signaling pathways in many organisms (Holmwood and Schindler, 2009). When activated by the steroid hormone ecdysone, or more precisely by its active metabolite 20-hydroxyecdysone, the ecdysone receptor is responsible for initiating the moulting of insects by binding to ecdysteroid binding elements. As result, the insect stays permanently trapped in the molting process and is unable to feed, it dies in the period of a few days from desiccation and starvation.

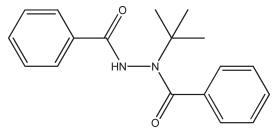
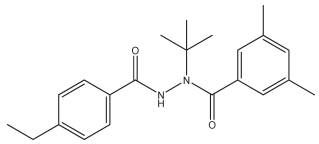
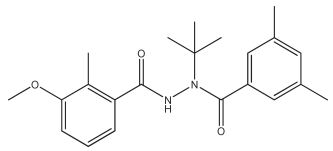
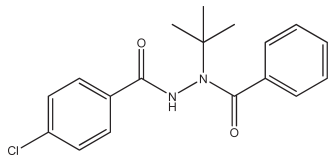
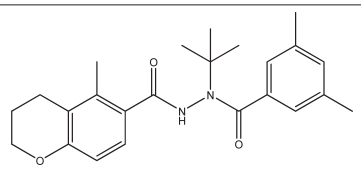
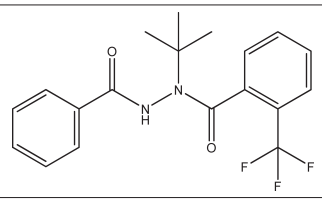
Dibenzoylhydrazines share the molting hormone receptor, which belong to the superfamily of nuclear receptors, with endogenous/endocrine-active ecdysteroids working in arthropods and nonarthropod invertebrates (Fujita and Nakagawa, 2009). The dibenzoylhydrazine derivatives include two benzene rings variously substituted. Several classical QSARs (quantitative structure-activity relationships) were applied to model the insecticidal activity of substituted dibenzoylhydrazines by correlating their insecticidal activity to rice stem borer larvae with several structural parameters, like: the 1-octanol/water partition coefficient, the inductive/field electronic effect of ortho substituents, the van der Waals volume (V) and ΔV value used as the difference from the reference V value of hydrogen and scaled by 0.1 (calculated for different substituent positions). The authors concluded an important participation to the activity of hydrophobic effect for each substituent, and of electronic effect especially for ortho substituents situated in the proximity of 'side-chain' carbonyl groups and unfavorable steric effects of meta and para substituents. Homology modeling and CoMFA (comparative molecular field analysis) approaches were applied as well to model the ligand-binding domain of ecdysone receptor.

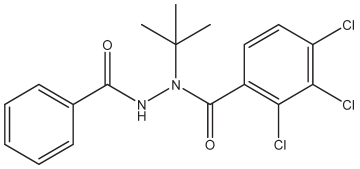
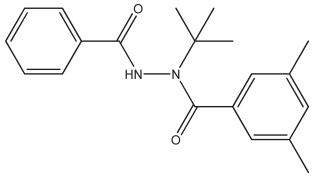
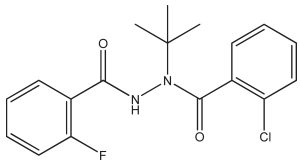
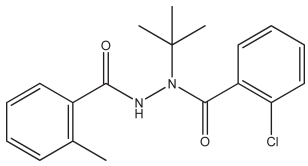
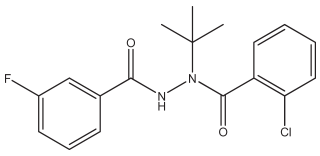
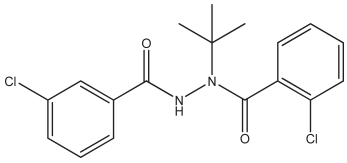
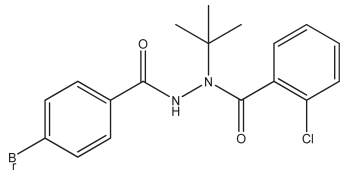
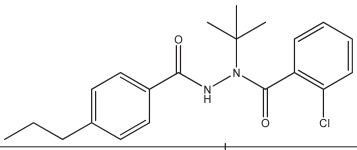
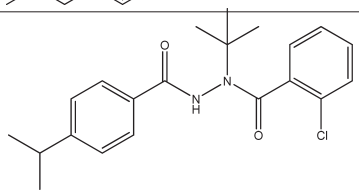
Hydrogen bonding was found to be important to the ecdysteroids-ecdysone receptors interactions by classical QSAR and CoMFA analysis (Harada et al, 2009). The binding of ecdysteroids to the ecdysone receptors of D. melanogaster was significantly correlated

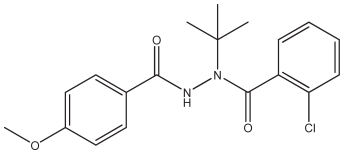
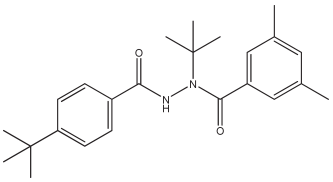
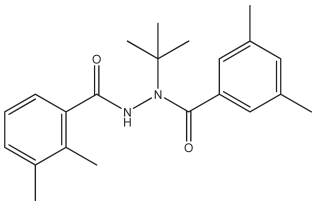
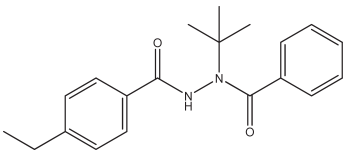
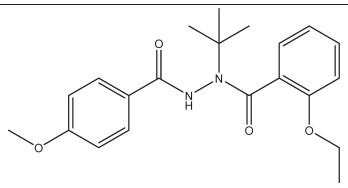
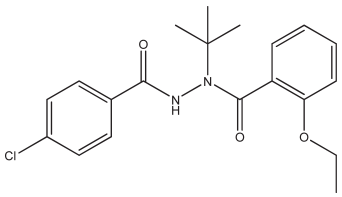
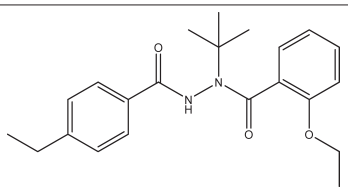
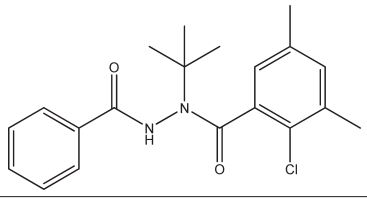
with the number of hydrogen bonds. Addition of steric effects slightly improved the correlation, even though the contribution of the steric effect was not as large as that of HB according to 3-D QSAR analysis.

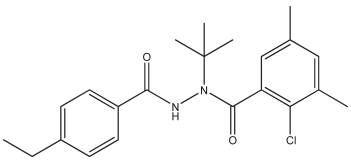
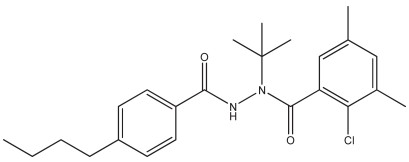
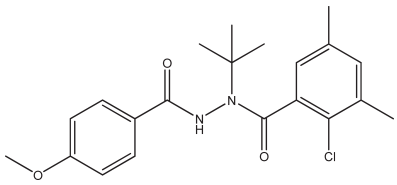
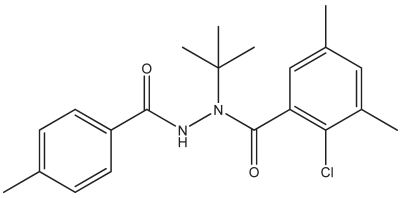
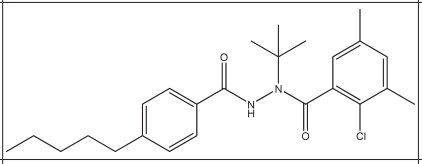
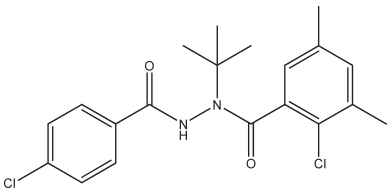
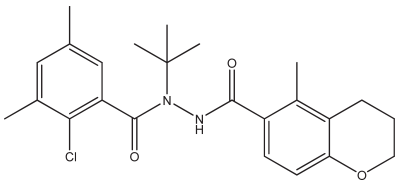
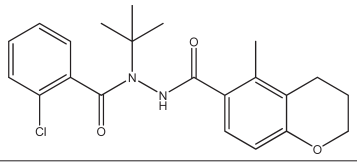
In this paper structural features of a series of 33 dibenzoylhydrazine ecdysone agonists (Soin et al, 2010) (Table 1), which influence the lethal larval molting process in susceptible insects that belong to the orders of one lepidopteran species, namely the cotton leafworm *Spodoptera littoralis* are studied. The quantitative relationship between chemical features and the ecdysone agonistic activity was determined by means of the multiple linear regression (MLR) approach.

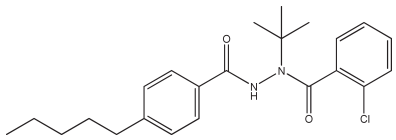
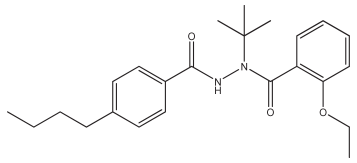
Table 1. The dibenzoylhydrazine structures, the experimental (pEC_{50}) and predicted (pEC_{50pred}) insecticidal activity values and descriptors included in the best the MLR1 model

No	Structure	pEC_{50}	pEC_{50pred}	EEig04r	RDF140m	Mor32p	L3s
1		5.89	5.79	2.91	0.00	-0.38	0.92
2		8.66	7.25	3.38	0.00	-0.28	1.25
3		8.22	7.75	3.36	0.00	-0.22	1.20
4		6.34	6.71	3.19	0.00	-0.36	0.90
5		8.58	7.82	3.57	0.00	-0.30	1.13
6		6.1	6.05	3.53	0.00	-0.42	1.65

7*		5.28	5.89	3.38	0.00	-0.46	1.31
8		6.3	6.86	3.31	0.00	-0.30	1.30
9		6.54	5.74	3.07	0.00	-0.31	1.55
10		6.36	6.58	3.09	0.00	-0.30	1.07
11*		6.34	5.62	3.19	0.00	-0.40	1.44
12*		5.77	5.55	3.15	0.00	-0.39	1.44
13		6.36	6.29	3.18	0.00	-0.35	1.20
14*		7.13	7.14	3.40	0.00	-0.33	1.13
15		7.76	7.42	3.60	0.00	-0.33	1.28

16		6.47	6.68	3.33	0.00	-0.36	1.14
17*		8.15	8.02	3.76	0.00	-0.30	1.32
18		7.79	8.04	3.35	0.00	-0.18	1.17
19		6.96	7.22	3.37	0.00	-0.36	0.88
20		4.66	4.94	3.33	0.00	-0.45	1.90
21		5.02	4.50	3.20	0.00	-0.44	1.99
22		5.16	5.18	3.38	0.00	-0.43	1.94
23		5.76	6.14	3.32	0.00	-0.35	1.55

24*		6.47	6.83	3.38	0.00	-0.29	1.50
25		5.95	6.39	3.41	0.09	-0.29	1.42
26		5.69	6.35	3.35	0.00	-0.34	1.49
27		5.87	6.51	3.33	0.00	-0.30	1.55
28		5.45	5.78	3.41	0.19	-0.29	1.38
29*		5.97	6.54	3.32	0.00	-0.30	1.55
30		7.17	7.78	3.57	0.00	-0.26	1.32
31		8.27	8.21	3.56	0.00	-0.24	1.13

32		6.49	6.07	3.40	0.19	-0.32	1.06
33		5.11	4.86	3.40	0.09	-0.42	1.80

*test compounds

2. Material and methods

Definition of target property and molecular structures

The ecdysone agonistic activity, measured *in vitro* based on an ecdysone-dependent reporter assay using cell lines derived from the lepidopteran species the cotton leafworm *Spodoptera littoralis*, expressed as pEC_{50} values (Table 1), and was used as dependent variable.

33 dibenzoylhydrazine insecticides (Table 1) were energy pre-optimized by molecular mechanics calculations using the MMFF94s force field included in the OMEGA (version 2.5.1.4, OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com) software (Hawkins et al, 2010; Hawkins & Nicholls, 2012). Structural 0D, 1D, 2D and 3D descriptors were calculated for the minimum energy structures using the DRAGON (Dragon Professional 5.5 (2007), Talete S.R.L., Milano, Italy) and InstantJchem (which was used for structure database management, search and prediction) (InstantJchem 15.10.0, 2012, ChemAxon (http://www.chemaxon.com) software .

MLR method

Multiple linear regression (Wold and Dunn III, 1983) (MLR) calculations were combined with a genetic algorithm for variable selection included in the QSARINS v.2.2 program (Chirico et al., 2012; Gramatica et al., 2013). The Genetic Algorithm (Depczynski et al, 2000) (GA) was used for the 1416 structural descriptors calculated for the 33 dibenzoylhydrazine compounds to select multiple linear regression models. GA is a reliable and extensively approach which uses adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. Fitness criteria are employed during the optimization processes illustrated by the evolution principles of Darwin of "survival of the fittest". The genetic algorithm evolves through other operators: crossover and mutation. In the QSARINS package the following parameters were used: the RQK fitness function (Todeschini et al, 2004 with leave-one-out cross-validation

(Hawkins et al, 2003) correlation coefficient as constrained function to be optimized, a crossover/mutation trade-off parameter of T = 0.5 and a model population size of P = 50.

The dibenzoylhydrazine derivatives were divided into training and test sets by splitting by response (Gramatica et al, 2012; Gramatica, 2014), for verifying the model on chemicals in the response domain (chemicals were ordered according to their increasing activity, and one out of every three chemicals was put in the prediction set, always including the most and the least active compounds in the training set). Seven compounds were taken out of the total number of compounds: compounds 7, 11, 12, 14, 17, 24, 29.

Model validation

The MLR models were internally validated using the following robustness were employed: leave-one-out cross-validation (Q^2_{LOO}), Y-scrambling (Todeschini et al, 1999) and Q^2_{LMO} leave-more-out (LMO) cross-validation (carried out for 21% of data out of training, each run). In Y-scrambling the process was randomly mixed 2000 times.

The domain of applicability was checked using the Williams plots (standardized crossvalidated residuals versus leverage (Hat diagonal) values) (Gramatica, 2013). A threshold of residual value greater than 2.5 times the value of standard error in calculation was employed for outlier detection.

The root-mean-square error (RMSE) of training and validation sets was compared to check the data over fitting and model applicability.

The Multi-Criteria Decision Making (MCDM) (Keller et al., 1991) is a technique that summarizes the performances of a certain number of criteria simultaneously, as a single number (score) between 0 and 1. This is done associating to every validation criteria a desirability function which values range from 0 to 1 (where 0 represents the worst validation criteria value and 1 the best). The geometric average of all the values obtained from the desirability functions gives the MCDM value. The ‚MCDM all' scores were calculated based on the fitting, cross validated and external criteria and were used to choose the best MLR models.

The model's predictive power was tested using the Q^2_{F1} (Shi et al, 2001); Q^2_{F2} (Schüürmann et al, 2008); Q^2_{F3} (Consonni et al, 2009) - external validation parameters and the concordance correlation coefficient (CCC) (Chirico & Gramatica, 2011) (having the thresholds values higher than 0.85, as they have been rigorously determined by a simulation study (Chirico & Gramatica, 2012)) and r^2_m (with a lowest threshold value of 0.5 to be accepted) (Roy et al, 2009).

3. Results and Discussion

A statistical analysis of the dibenzoylhydrazine analogues was performed using the multiple linear regression method.

A training set of 26 compounds and 7 test compounds (no.: 7, 11, 12, 14, 17, 24, 29) were used in the MLR calculations (Table 1). Structural parameters were calculated by InstantJChem and Dragon programs from the structures of the minimum energy obtained by molecular mechanics calculations. Variable selection was carried out by the genetic algorithm, using the leave-one-out fit criterion as constrained function to be optimized. Several satisfactory MLR models were obtained (Tables 2 to 4).

Table 2. Internal validation parameters of the MLR models (training set)*

Model	$r^2_{training}$	q^2_{LOO}	q^2_{LMO}	r^2_{adj}	$RMSE_{tr}$	MAE_{tr}	CCC_{tr}	r^2_{scr}	q^2_{scr}	SEE	F
MLR1	0.797	0.711	0.688	0.758	0.500	0.403	0.887	0.163	-0.313	0.556	20.560
MLR2	0.786	0.682	0.647	0.745	0.513	0.400	0.880	0.164	-0.297	0.570	19.296
MLR3	0.778	0.673	0.318	0.735	0.523	0.406	0.875	0.161	-0.310	0.582	18.365
MLR4	0.776	0.690	0.661	0.734	0.524	0.409	0.874	0.161	-0.304	0.583	18.218
MLR5	0.756	0.650	0.618	0.710	0.547	0.429	0.861	0.159	-0.304	0.609	16.303
MLR6	0.750	0.655	0.620	0.702	0.555	0.422	0.857	0.163	-0.296	0.617	15.731

*$r^2_{training}$ - correlation coefficient; q^2_{LOO} - leave-one-out correlation coefficient; q^2_{LMO} - leave-more-out correlation coefficient; $RMSE_{tr}$-root-mean-square errors; MAE_{tr}-mean absolute error; CCC_{tr}-the concordance correlation coefficient; r^2_{scr} - scrambled r^2; q^2_{scr} - scrambled cross-validated q^2; SEE-standard error of estimates; F-Fischer test.

Table 3. External validation parameters calculated for the MLR models (test set)*

Model	Q^2_{F1}	Q^2_{F2}	Q^2_{F3}	$RMSE_{ext}$	MAE_{ext}	CCC_{ext}
MLR1	0.741	0.740	0.836	0.449	0.374	0.863
MLR2	0.629	0.628	0.765	0.537	0.463	0.817
MLR3	0.728	0.727	0.828	0.460	0.402	0.843
MLR4	0.826	0.826	0.890	0.368	0.342	0.901
MLR5	0.747	0.747	0.840	0.443	0.381	0.861
MLR6	0.766	0.766	0.852	0.426	0.294	0.828

*Q^2_{F1} (Shi et al., 2001), Q^2_{F2} (Schüürmann et al., 2008), Q^2_{F3} (Consonni et al., 2009)-external validation parameters; $RMSE_{ext}$-root-mean-square errors; MAE_{ext} -mean absolute error; CCC_{ext}-the concordance correlation coefficient

Table 4. Other predictivity parameters (r^2_m), the Multi-Criteria Decision Making (MCDM) values and final descriptors selected in the MLR models.

Model	r^2_m	MCDM	Descriptors included in the model*
MLR1	0.700	0.773	EEig04r RDF140m Mor32p L3s
MLR2	0.700	0.728	TIC0 EEig05d EEig04r Mor32p
MLR3	0.730	0.707	TIC0 EEig04r VRD2 Mor32p
MLR4	0.809	0.784	EEig04r Mor32p L3s Am
MLR5	0.729	0.745	IDMT EEig04r Mor32p L3s
MLR6	0.589	0.742	EEig05d L3s Am F03[C-C]

*EEig04r represents the eigenvalue 4 from edge adj. matrix weighted by resonance integrals; RDF140m- Radial Distribution Function - 14.0 / weighted by atomic masses; Mor32p- D-MoRSE - signal 32 / weighted by atomic polarizabilities; L3s- 3rd component size directional WHIM index / weighted by atomic electrotopological states; TIC0- total information content index (neighborhood symmetry of 0-order); EEig05d- Eigenvalue 05 from edge adj. matrix weighted by dipole moments; VRD2- average Randic-type eigenvector-based index from distance matrix; Am- A total size index / weighted by atomic masses; IDMT- total information content on the distance magnitude; F03[C-C]- frequency of C-C at topological distance 3.

The MLR models included in Tables 2 to 4 are completely satisfactory in the fitting, and have good predictive power. They have been assessed by internal (LOO and LMO) cross-validation, Y-scrambling.

Best statistical results for model fitting and predictive power were obtained for the MLR1 model. Experimental versus predicted pEC_{50} values, Williams plots and Y-scramble plots this model are presented in Figure 1, 2 and 3, respectively. The Williams plots validate the absence of outliers and influential points in the final selected MLR1 model.

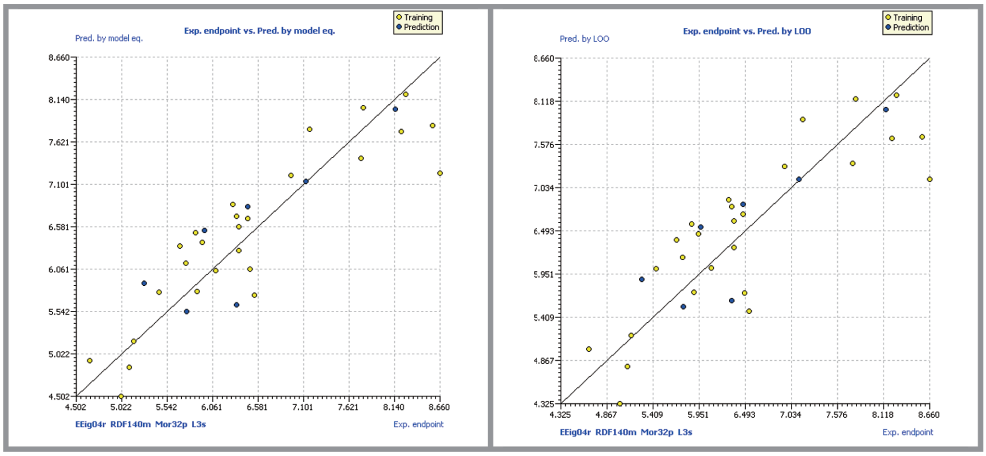


Figure. 1. Experimental versus calculated pEC_{50} values for the MLR1 model predicted by the model (left) and leave-one-crossvalidation method (right).

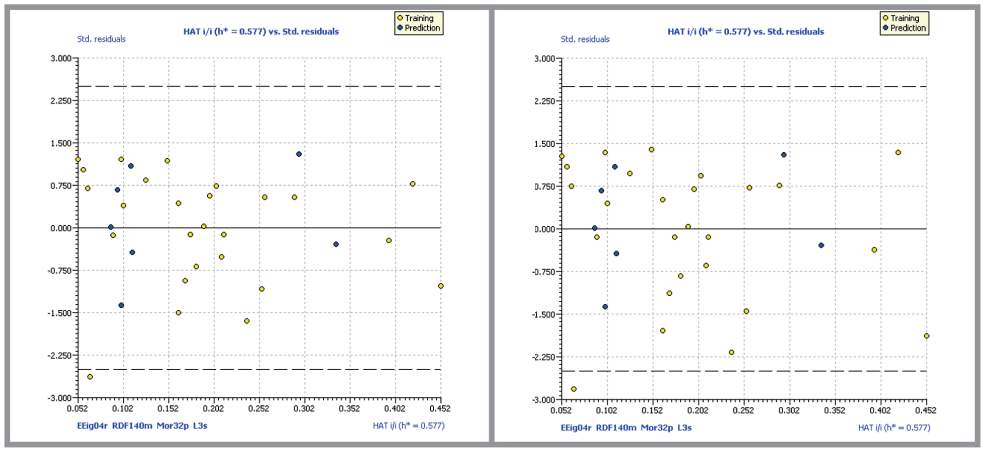


Figure. 2. Williams plot predicted by the final MLR1 (left) model and by the leave-one crossvalidation approach (right).

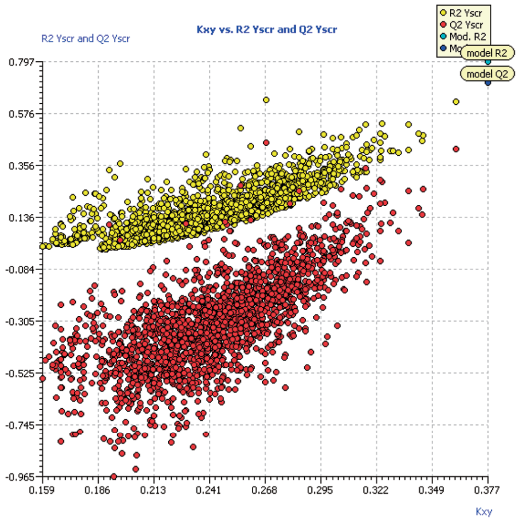


Figure. 3. Y-scramble plots for the MLR1 model.

The RMSE values for the training and validation sets are similar for the MLR1 model. The chosen models demonstrate a satisfactory stability in internal validation, have high fitting, internal and external predictivity (verified by different validations). The small difference of CCC values between the training and test sets of 2.4% (model 1) demonstrates that this model is able to predict the response for chemicals not used in the model development (validation set) just as they do for chemicals used to find the relationship (training set).

The risk of chance correlation was verified by the Y-scrambling procedure. The

extremely low calculated scrambling values (Table 2) indicate no chance correlation for the chosen models.

Edge adjacency indices are deduced from the edge (i.e. atom) adjacency matrix encoding the molecular connectivity between graph edges. The EEig04r (eigenvalue 4 from edge adj. matrix weighted by resonance integrals) descriptor in the MLR1 linear equation has positive coefficient value. Increase of its value will increase the pEC_{50} values.

RDF (The radial distribution function) descriptors are geometrical descriptors based on the geometrical interatomic distance and constitute a radial distribution function code. They represent 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. Low values of the RDF140m (Radial Distribution Function - 14.0 / weighted by atomic masses) descriptor increase the pEC_{50} values.

Other geometrical descriptors are the 3D-MoRSE (Molecule Representation of Structure based on Electron) descriptors, which are the sums of atom weights with different angular scattering function. High values of the Mor32p (3D-MoRSE - signal 32 / weighted by atomic polarizability) descriptor yield high pEC_{50} values.

WHIM (Weighted Holistic Invariant Molecular) descriptors are geometrical descriptors based on statistical indices calculated on the projections of the atoms along principal axes. In case of the L3s (3rd component size directional WHIM index / weighted by atomic electrotopological states) descriptor the atomic electrotopological states are one of the weighting schemes that are used for computing the weighted covariance matrix. L3s descriptor in the MLR1 linear equation has negative coefficient value; therefore increase of its value will decrease the pEC_{50} values.

4. Conclusion

Dibenzoylhydrazines share the molting hormone receptor, which belong to the superfamily of nuclear receptors, with endogenous/endocrine-active ecdysteroids working in arthropods and nonarthropod invertebrates. A series of 33 dibenzoylhydrazine derivatives were pre-optimized using molecular mechanics calculations and the calculated structural features were then related to the ecdysone agonistic activity (based on an ecdysone-dependent reporter assay using cell lines derived from the the cotton leafworm *Spodoptera littoralis* lepidopteran species using the multiple linear regression (MLR) approach. Good correlations with the ecdysone agonistic activity were found and good predictive models. Geometrical descriptors related to the ligand molecular conformation in 3D space weighted by electropological space and

polarizability influence the insecticidal activity. Based on the proposed MLR models new active insecticides can be designed.

Acknowledgements

This project was financially supported by Project 1.1 of the Institute of Chemistry of the Romanian Academy. Access to the Chemaxon Ltd., OpenEye Ltd. and QSARINS software are greatly acknowledged by the authors.

References

Chirico N., Gramatica P. (2012). 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.* **52**, 2044-2058.

Chirico N., Gramatica P.(2011). 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.*, **51**, 2320-2335.

Chirico N., Papa E., Kovarich S., Cassani S., Gramatica P. (2012). *QSARINS, software for QSAR MLR model development and validation. QSAR Res. Unit in Environ. Chem. and Ecotox., DiSTA, University of Insubria, Varese, Italy. http://www.qsar.it.*

Consonni V., Ballabio D., Todeschini R. (2009). Comments on the Definition of the Q^2 Parameter for QSAR Validation, *J. Chem. Inf. Model.* **49**, 1669-1678.

Depczynski U., Frost V. J. , Molt K. (2000). Genetic algorithms applied to the selection of factors in principal component regression, *Anal. Chim. Acta* **420**, 217-227.

Dinan L. (2003). Ecdysteroid structure-activity relationships, Atta-ur-Rahman (Ed.) *Studies in Natural Products Chemistry*, Vol. 29., Elsevier B.V., pp. 3-71.

Fujita T., Nakagawa Y. (2009). SAR and QSAR analyses of substituted dibenzoylhydrazines for their mode of action as ecdysone agonists, in: Endocrine Disruption Modeling, J. Devillers, ed., CRC Press, Boca Raton, FL, pp. 357-377.

Gramatica P. (2014). External Evaluation of QSAR Models, in Addition to Cross-Validation: Verification of Predictive Capability on Totally New Chemicals, *Mol. Inf.*, **33**, 311-314.

Gramatica P., Cassani S., Roy P. P., Kovarich S., Yap C. W., Papa E. (2012). QSAR Modeling is not "Push a Button and Find a Correlation": A Case Study of Toxicity of (Benzo-)triazoles on Algae, *Mol. Inf.*, **31**, 817-835.

Gramatica, P., Chirico, N., Papa, E., Cassani, S.,Kovarich, S.(2013). QSARINS: A new software for the development, analysis, and validation of QSAR MLR models, *J.*

Comput. Chem., **34**, 2121-2132.

Harada T., Nakagawa Y., Akamatsu M., Miyagawa H. (2009). Evaluation of hydrogen bonds of ecdysteroids in the ligand–receptor interactions using a protein modeling system, Bioorgan. *Med. Chem.* **17**, 5868-5873.

Harmatha J., Dinan L., Lafont R. (2002). Biological activities of a specific ecdysteroid dimer and of selected monomeric structural analogues in the BII bioassay, *Insect Biochem. Molec.* **32**, 181-185.

Hawkins D.M., Basak S.C., Mills D. (2003). Assessing Model Fit by Cross-Validation, *J. Chem. Inf.Comput. Sci.* **43**, 579-586.

Hawkins P.C.D., Nicholls A. (2012). Conformer Generation with OMEGA: Learning from the Data Set and the Analysis of Failures. *J. Chem. Inf. Model.* **52**, 2919-2936.

Hawkins P.C.D., Skillman A.G., Warren G.L., Ellingson B.A., Stahl M.T. (2010). Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **50**, 572-584.

Holmwood G., Schindler M. (2009). Protein structure based rational design of ecdysone agonists, *Bioorgan. Med. Chem.* **17**, 4064-4070.

Keller H.R., Massart D.L., Brans J.P. (1991). Multicriteria decision making: a case study. *Chemom. Intell. Lab. Syst.*, **11**, 175-189.

Roy P.P., Paul S., Mitra I., Roy K. (2009). On Two Novel Parameters for Validation of Predictive QSAR Models, *Molecules* **14**, 1660-1701.

Schüürmann G., Ebert R.U., Chen J., Wang B., Kuhne R. (2008). External Validation and Prediction Employing the Predictive Squared Correlation Coefficient s Test Set Activity Mean vs Training Set Activity Mean. *J. Chem. Inf. Model.* **48**, 2140-2145.

Swevers L., Soin T., Mosallanejad H., Iatrou K., Smagghe G. (2008). Ecdysteroid signaling in ecdysteroid-resistant cell lines from the polyphagous noctuid pest Spodoptera exigua. *Insect Biochem.* **38**, 825-833.

Shi L.M., Fang H., Tong W., Wu J., Perkins R., Blair R.M., Branham W.S., Dial S.L., Moland C.L., Sheehan D.M. (2001). QSAR models using a large diverse set of estrogens. *J. Chem. Inf. Model.* **41**, 186-195.

Soin T., De Geyter E., Mosallanejad H., Iga M., Martín D., Ozaki S., Kitsuda S., Harada T., Miyagawa H., Stefanou D., Kotzia G., Efrose R., Labropoulou V., Geelen D., Iatrou K., Nakagawa Y., Janssen C.R., Smagghe G., Swevers L. (2010). Assessment of species specificity of moulting accelerating compounds in Lepidoptera: comparison of activity between Bombyx mori and Spodoptera littoralis by in vitro reporter and in vivo toxicity assays. *Pest. Manag. Sci.* **66**, 526-535.

Soin T., Iga M., Swevers L., Rougé P., Janssen C.R., Smagghe G. (2009). Towards Coleoptera-specific high-throughput screening systems for compounds with ecdysone activity: development of EcR reporter assays using weevil (Anthonomus grandis)-

derived cell lines and in silico analysis of ligand binding to A. grandis EcR ligand-binding pocket. *Insect Biochem.* **39**, 523-534.

Todeschini R., Consonni V., Maiocchi A. (1999). The K correlation index theory development and its application in chemometrics. *Chemometr. Intell. Lab.* **46**, 13-29.

Todeschini R., Consonni V., Mauri A., Pavan M. (2004). Detecting "bad" regression models: multicriteria fitness functions in regression analysis, *Anal. Chim. Acta* **515**, 199-208.

Wold, S., Dunn III, W. J. (1983). Multivariate quantitative structure activity relationships (QSAR): Conditions for their applicability, *J. Chem.Inf. Comput. Sci.* **23**, 6-13.

和文要旨

エクジソン-アゴニストによる殺虫活性を有するジベンゾイルヒドラジン誘導体の計算化学による解析

鈴木孝弘，Simona Funar-Timofei，Alina Bora，
Luminita Crisan，Ana Borota

エクジステロイドレセプターは、選択的で効力の高い殺虫剤を開発するための重要な標的である。本研究では、ジベンゾイルヒドラジン誘導体のエクジソン－アゴニストによる殺虫活性をエクジソン・レポーターアッセイによって求め、その構造活性相関を重回帰分析（MLR）モデルによって解析した。ジベンゾイルヒドラジン誘導体の構造は、Omegaパッケージに含まれているMMFF94sを使った分子力場計算による分子の立体配座の安定性を検討し、２種の*E*体、*Z*体のシス－トランス異性体について種々の0D、1D、2D、3D記述子を計算した。それらの記述子と実験により求められたエクジソン－アゴニスト活性pEC_{50}との間の定量的構造活性相関を変数選択のための遺伝的アルゴリズムを用いた重回帰分析（MLR）によりモデル化した。全13化合物のうち７種をテストセットに用い、MLRモデルの安定性と予測性を評価したところ、最も統計的に有意なMLRモデルは、$r^2_{training} = 0.797$, $r^2_{test} = 0.753$, $q^2_{LOO} = 0.711$, $RMSE_{tr} = 0.499$, $RMSE_{ext} = 0.449$, $r^2_{adj} = 0.758$の統計的な指標を有した。ジベンゾイルヒドラジン誘導体の幾何学的な特徴が、殺虫活性に最も大きな寄与があることが明らかになった。