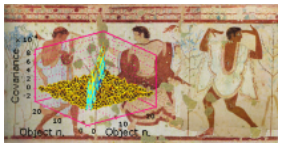

Quantitative Structure Activity Relationship Modelling of Environmentally Important Diphenyl Ether Herbicides Using MLR and PLS

A. Rouhollahi, E. Babaee and S. Aghajani

Chemistry Department, Faculty of Sciences, K. N. Toosi University of Technology, Tehran, Iran

Abstract

A quantitative structure – activity relationship (QSAR) study has been applied to a data set of 33 environmentally important diphenyl ether herbicides with theirs inhibition data on protoporphyrinogen oxidase (PPO) enzyme. Diphenyl ethers (DPEs) have been a class of herbicides that strongly and competitively inhibit the protoporphyrinogen oxidase (PPO) enzyme. PPO is the last common enzyme in the biosynthetic pathway to heme and chlorophyll. Two linear correlating models, multiple linear regression (MLR) and partial least squares (PLS) regression methods were used. The physicochemical meaning of the descriptors and their relations to the dependent variable are discussed

Introduction

Diphenyl ethers (DPEs) have been a class of herbicides that strongly and competitively inhibit the protoporphyrinogen oxidase enzyme. Protoporphyrinogen oxidase (PPO) is a key enzyme in the chlorophyll/heme biosynthetic pathway. This enzyme catalyzes the oxidation of protoporphyrinogen IX to protoporphyrin IX (PPIX) [1]. Herbicides are the most widely used class of pesticides, accounting for more than 60% of all pesticides applied in agriculture [2]. A number of herbicides have been identified as acting through the inhibition of PPO in plants. These inhibitors are structurally very diverse; include the diphenyl ethers, *N*-aryltetrahydropthalimides, oxadiazon and flumioxazin. These herbicides are potent inhibitors of PPO from both plants and animals [3]. The attributes of low application rates, good crop selectivity, low residue and environmental safety exhibited by these compounds are important characteristics for modern agrochemicals, which have led to the rapid success of Protox inhibitors as herbicidal products and attracted a worldwide research commitment.

Materials & Methods

Protoporphyrinogen oxidase enzyme inhibition data (I_{50} is molar concentration that leads to 50% inhibition) for 33 diphenyl ether molecules were taken from literature sources [4, 5]. HyperChem software (version 7, Hypercube, Inc) was used to draw the chemical structure of the molecules. AM1 semi-empirical calculation was used to optimize the 3D geometry of the molecules.

Results

For the selection of important descriptors the linear regression technique was used based on the construction of a linear mathematical model relating the observed pI_{50} to numerically encode structural parameters. The parameters appearing in the best equation showed that four descriptors are the most important for the prediction of pI_{50}:

$$pI_{50} = -3.5475 - 0.8335 \times Mor03m + 0.8605 \times PCR - 0.1649 \times RDF075m + 0.4919 \times T(Cl..Cl) \quad (1)$$

For the constructed model, general statistical parameters were selected to evaluate the prediction ability of the model for pI_{50}. Statistical parameters (RMSEP, SEP, REP%, R^2_{pred}, Q^2_{ext}) calculated for MLR and PLS models represented in Table 1. By comparison of the mean effects of the descriptors appearing in MLR model, it is observed that the *PCR* of the molecules has the largest effect on the pI_{50} of the diphenyl ethers (Table 4). The descriptor *PCR* is the ratio of multiple path counts to path counts. Similar to the invariants derived from the distance/distance matrix **D**/**D**, atomic path/walk indices are defined for each atom as the ratio between atomic path count ${}^mP_i$ and atomic walk count awc_i^m of the same length *m*, i.e.

$$({}^{p}\!/\!_{w})_i^m = {}^mP_i / awc_i(m) \qquad (2)$$

whereas the number of paths in a molecule is bonded and determined by the molecule diameter, the number of walks is unbounded.

Table 1. Statistical parameters obtained by applying the MLR and PLS methods to the test set

Parameter	MLR	PLS
RMSEP[1]	0.3436	0.3620
SEP[2]	0.3711	0.3910
REP[3] **(%)**	5.7372	6.0814
$\mathbf{R^2_{pred}}$	0.9478	0.9403
$\mathbf{Q^2_{ext}}$	0.9519	0.9466
No. LVs[4]	-	3
No. DSs[5]	4	4

[1] Root mean squares error of prediction, [2] Standard error of prediction, [3] Relative error of prediction, [4] Number of latent variables, [5] Number of descriptors.

Results

Variation in the herbicidal activity of diphenyl ether compounds were analyzed using the theoretical molecular descriptors. The multilinear QSAR equations were obtained by MLR and PLS methods and the most relevant set of descriptors were also selected by stepwise variable selection method. According to obtained results it is concluded that the Mor03m, PCR, RDF075m and T(Cl..Cl) can be used successfully for modelling biological activity pI_{50} of the under study compounds. Combination of 2D and 3D descriptors allows distinguishing between activities of enantiomers. The high correlation coefficient (0.9478 and 0.9403 for MLR and PLS, respectively) and low prediction errors obtained confirm good predictive ability of both models. One more point is that useful to add the accuracy of the experimental data reflects in the model quality and we suggest for a more robust model do some additional experimental work on the suspect data entry.

References

1) J.M.Camadro, M. Matringe, R. Scalla, P. Labbe, Kinetic studies on protoporphyrinogen oxidase inhibition by diphenyl ether herbicides, Biochem. J., 277(1), (1991) 17-21
2) R.L. Zimdahl, My view, Weed Sci., 50(6), (2002) 687-687
3) C. Tomlin, The Pesticide Manual, 13th ed. Hampshire: BCPC Publications (2003) ISBN-10 1-90139613-4
4) G.L. Durst, Comparative Molecular Field Analysis (CoMFA) of Herbicidal Protoporphyrinogen Oxidase Inhibitors using Standard Steric and Electrostatic Fields and an Alternative LUMO Field, Quant. Struct. Act. Relat., 17(5), (1998) 419-426
5) F.E. Dayan, S.N. Allen, Predicting the activity of the natural phytotoxic diphenyl ether cyperine using Comparative Molecular Field Analysis, Pest Manag. Sci., 56(8), (2000) 717-722