STRUCTURE-ACTIVITY RELATIONSHIPS ANALYSIS OF FUNGICIDAL AND HERBICIDAL THIOLCARBAMATES USING PLS METHOD

YOSHIKATSU MIYASHITA,^{*} HIROYUKI OHSAGO, CHIYOZO TAKAYAMA,[†] and SHIN-ICHI SASAKI

Department of Knowledge-based Information Engineering, Toyohashi University of Technology, Tempaku, Toyohashi 441, Japan, [†]Agricultural Research Laboratory, Takarazuka Research Center, Sumitomo Chemical Co. Ltd. Takatsukasa, Takarazuka 665, Japan

Abstract - Structure-activity correlations of 83 thiolcarbamates with fungicidal and herbicidal activities were investigated by the multivariate partial least squares (PLS) method. PLS2 model equations with three latent variables successfully classified fungicidal and herbicidal activity classes. The PLS2 modelling of multivariate activities gives useful information concerning optimization of an activity profile.

Key words PLS, multivariate QSAR, fungicidal and herbicidal thiolcarbamates, pattern recognition

The information obtained by the analysis of structure-activity data using mathematical and statistical (chemometric) methods is essential to drug design. In order to model activity data with chemical structures, chemometric methods play an important role in studying quantitative structure-activity relationships (QSAR). A compound exhibits several biological activities. If we establish a mathematical modelling of multivariate activity data, it is more easy to optimize an activity profile of a biologically active compound. The multiple linear regression analysis has been exclusively used in QSAR modelling. Recently, a novel method called PLS (partial least squares method in latent variables) proved to be quite useful for modelling biological data [1-3]. The PLS modelling method is a powerful tool for higher level pattern recognition [4,5].

In the present study, QSAR of eighty-three N-aralkyl thiolcarbamates with fungicidal and herbicidal activities was investigated by the PLS2 method. The PLS2 model equations were used for the classification of fungicidal and herbicidal activity classes.

Thiolcarbamate

METHOD

The fungicidal and herbicidal activity data of thiolcarbamates reported by Kirino et al. [6] were used. The fungicidal and herbicidal activities of the thiolcarbamates were rated 1 through 4. In this study the compounds rated at 4 were merged into those rated at 3, because the number of compounds rated at 4 is smaller than those rated at 1 through 3. Now, we have 42 weak (class 1), 22 moderate (class 2) and 19 strong fungicidal compounds (class 3), and 25 weak (class 1), 22 moderate (class 2) and 36 strong herbicidal compounds (class 3).

In the PLS2 analysis, $E_s^C(R_1)$, $E_s^C(R_2)$, $E_s^C(R_3)$, $\pi(R_1)$ and R_2H are used as structural parameters which are the same as used in the work by Kirino et al. [6]. $E_s^C(R_1)$, $E_s^C(R_2)$, and $E_s^C(R_3)$ are the Hancock corrected steric substituent constants for R_1 , R_2 , and R_3 , respectively. $\pi(R_1)$ is a hydrophobic constant for R_1 . R_2H is an indicator descriptor, which is equal to 1 when the substituent R_2 is a hydrogen atom, and zero for the other substituents. Five physico-chemical parameters were autoscaled to unit variance to give each parameter equal importance in the PLS analysis. In the PLS2 analysis, the biological data $Y = (Y_F \ Y_H)$ were related to structural data X by the same latent variables t .

A complete description of the PLS method has been given elsewhere[1-3]. The PLS model is given by the following equations,

$$X = \sum_{h=1}^A t_h p'_h + E \quad (1)$$

$$Y = \sum_{h=1}^A t_h q'_h + F \quad (2)$$

where t_h is a latent variable which is a linear combination of descriptor variables X . p_h and q_h are loading vectors for X and Y matrix in component h . E and F are residuals. A is the number of components of the PLS model which was determined by the cross-validation method[7].

RESULT AND DISCUSSION

A three component PLS2 model based on 5 physico-chemical parameters explains 75.6% of the fungicidal and herbicidal activity data. The PLS model equations for fungicidal activity y_F and herbicidal activity y_H are given as follows,

$$y_F = -0.482t_1 + 0.278t_2 - 0.184t_3 + 1.723 \quad (3)$$

$$y_H = 0.329t_1 + 0.498t_2 + 0.467t_3 + 2.133 \quad (4)$$

The model equations were used to classify a compounds into 3 classes. The recognition rates of fungicidal and herbicidal activities are 86% and 80%, respectively. The model equations show structural requirements for fungicidal and herbicidal activities. From Eqs. (3) and (4) it is clear that if we wish to design more potent fungicidal and herbicidal thiolcarbamates, t_2 should be increased while t_1 and t_3 should be restricted to near zero. This means that the PLS model gives more information to design optimum multivariate activity through latent variables. Since the latent variables t_1, t_2 , and t_3 are expressed in terms of original parameters, Eqs. (3) and (4) are represented by the 5 physico-chemical parameters.

$$y_F = 0.325E_s^C(R_1) + 0.131E_s^C(R_2) - 0.016E_s^C(R_3) + 0.731\pi(R_1) + 0.712R_2H - 0.401 \quad (5)$$

$$y_H = 0.185E_s^C(R_1) + 0.573E_s^C(R_2) + 0.510E_s^C(R_3) + 0.186\pi(R_1) - 1.898R_2H + 2.670 \quad (6)$$

These equations could be used to predict both more potent fungicidal and herbicidal thiolcarbamates.

The PLS2 method has a great advantage that it can handle structural parameters X and multivariate activity data Y and get a total picture of relationships between X and Y blocks. Therefore, valuable information on the design of a drug with multivariate activities could be obtained by the PLS2 method.

REFERENCES

1. B. Norden, U. Edlund, D. Johnels and S. Wold, *Quant. Struct.-Act. Relat.* **2**, 73-76 (1983)
2. H. Martens and T. Naes, *Multivariate Calibration*, John Wiley & Sons, Chichester 1989.
3. P. Geladi and B. R. Kowalski, *Anal. Chim. Acta*, **185**, 1-17 (1986)
4. Y. Miyashita, Y. Takahashi, C. Takayama, K. Sumi, K. Nakatsuka, T. Ohkubo, H. Abe and S. Sasaki, *J. Med. Chem.*, **29**, 906-912 (1986).
5. T. Okuyama, Y. Miyashita, S. Kanaya, H. Katsumi, S. Sasaki and M. Randic, *J. Comput. Chem.*, **9**, 636-646 (1988).
6. O. Kirino, C. Takayama, M. Yoshida, S. Inoue and R. Yoshida, *Agric. Biol. Chem.*, **52**, 561-568 (1988).
7. S. Wold, *Technometrics*, **20**, 397-406 (1987).