Comparative Assessment of Multiresponse Regression Methods for Predicting the Mechanisms of Toxic Action of Phenols

Shijin Ren*,†,‡ and Hyunjoong Kim†,§,||

Department of Statistics and Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee 37996-2200, and Department of Applied Statistics, Yonsei University, Shin-Chon Dong, Seo-Dae-Mun Gu, Seoul, Korea 120-749

Received May 9, 2003

The use of regression methods for classifying and predicting the mechanisms of toxic action of phenols was investigated in this study. Multiresponse regression was conducted using a total of six linear and nonlinear regression methods: simple linear regression (LinReg), logistic regression (LogReg), generalized additive model (GAM), locally weighted regression scatter plot smoothing (LOWESS), multivariate adaptive regression splines (MARS), and projection pursuit regression (PPR). A database containing phenols acting by four mechanisms (polar narcosis, weak acid respiratory uncoupling, proelectrophilicity, and soft electrophilicity) was used to assess the performances of the six regression methods in the multiresponse regression approach. For comparison purposes, traditional linear discriminant analysis (LDA) was also conducted as a baseline method to study the potential improvement of prediction accuracy by the multiresponse regression approach. Results showed that compared to LDA, the overall mechanism prediction error rate could be reduced to below 10% by multiresponse regression based on PPR. In addition to prediction accuracy, interpretability of the resultant models was discussed.

INTRODUCTION

In aquatic toxicology, it is known that chemical compounds exhibit toxicity via different mechanisms of toxic action. Several mechanisms are thoroughly studied and well established, such as narcosis and respiratory uncoupling (e.g., ref 1). Knowledge regarding mechanism of action is important because it enhances the understanding of the interaction between a xenobiotic compound and a living organism. This information has also been shown to be useful in developing and using mechanism-based quantitative structure−activity relationships (QSARs) for toxicity modeling and prediction. Such QSARs require the identification of the mechanism of a compound in order to choose a QSAR suitable for predicting its toxicity. While a current point of view is that high quality QSARs can only be developed for compounds with a common mechanism,[2] correctly determining toxicity mechanisms of chemical compounds is not an easy task.[3] As Schultz et al.[2] pointed out, the difficulty lies in the concept that the exhibition of toxic effects depends on several interactions, many of which are poorly understood or characterized.

A number of methods exists for determining the mechanisms of toxic action of chemical compounds. These methods include fish acute toxicity syndromes or FATS,[4] joint toxic action studies,[5] and some other methods described in the study of Russom et al.[6] The most frequently and conveniently used method is perhaps to assign a mechanism based on inspection of the chemical structure of a compound (e.g., the presence or absence of certain characteristic substituents). This method is sometimes referred to as rule-based mechanism designation (see, e.g., refs 6 and 7) and is popular partly because no experimental work is required to assign a mechanism to a compound, unlike methods such as FATS and joint toxic action studies. The rule-based method can work well for compounds with simple structures but may become unreliable for compounds with complex structures. Additionally, some characteristic structures are associated with more than one mechanism,[8] which cannot be properly accounted for by the rule-based method.

Recently, attention has been placed on the use of statistical classification methods in predicting the mechanisms of chemical compounds, especially phenols (e.g., refs 9−11). This approach is based on the analysis of the molecular descriptors of compounds with known or designated mechanisms in a learning database to develop a statistical model. This model can then be used for compounds not in the database for mechanism predictions. So far, the statistical method used for this purpose is mainly discriminant analysis. Other methods such as logistic regression is also used (e.g., ref 12) but to a lesser extent. The advantages and disadvantages of the statistical classification approach have been discussed previously.[9,13]

As with any other methods, there are uncertainties associated with the mechanism predictions made by the statistical classification approach. Oftentimes these uncertainties are quantified by the overall error rates of the classification methods. Naturally a low error rate is desired. Research is underway at the University of Tennessee to reduce the error

* Corresponding author phone: (865)974-9004; fax: (865)974-7076; e-mail: sjren@utk.edu. Current address: Department of Chemical Engineering, 419 Dougherty Engineering Building, University of Tennessee, Knoxville, TN 37996-2200.

† Department of Statistics, University of Tennessee.

‡ Center for Environmental Biotechnology, University of Tennessee.

§ Current address: Department of Applied Statistics, Yonsei University, Shin-Chon Dong, Seo-Dae-Mun Gu, Seoul, Korea 120-749.

|| Yonsei University.

10.1021/ci034092y CCC: $25.00 © 2003 American Chemical Society
Published on Web 09/11/2003

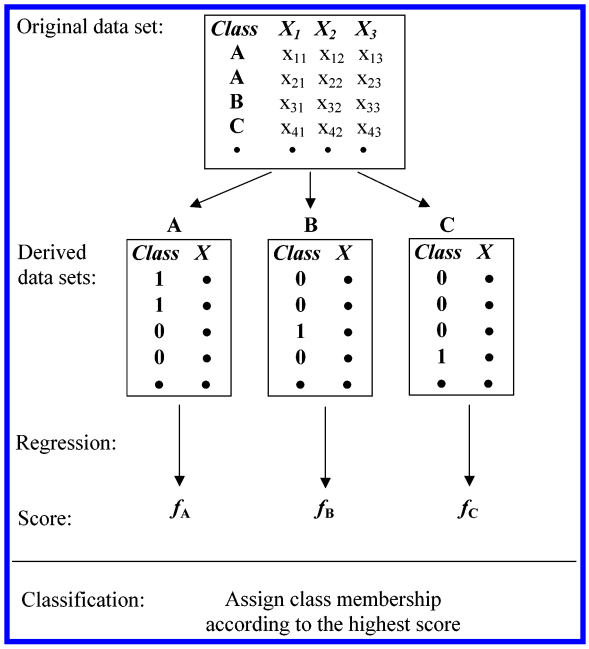

Figure 1. Multiresponse regression approach for classification.

rate of the statistical classification approach. To this end, a suitable database has been identified. The performances of different statistical classification models are assessed and compared with the performance of discriminant analysis. One of the attempts that has been made in this regard is the use of regression methods (both linear and nonlinear) for classification purposes. In toxicological research, regression methods have been extensively used for toxicity potency modeling, but they have not been used for mechanism classifications. In this manuscript, the progress in this aspect is reported. A specific aim of this study was to investigate the potential improvement of mechanism prediction accuracy of the regression methods over discriminant analysis.

METHODS

Database. The phenol database presented in the study of Aptula et al.[9] was identified as suitable for developing statistical classification models. The database contained 153 polar narcotic, 18 weak acid respiratory uncoupling, 27 proelectrophilic, and 23 soft electrophilic phenols. Five molecular descriptors were involved in the database, and they were $\log K_{ow}$, pK_a, E_{LUMO}, E_{HOMO}, and N_{hdon} (descriptions of these molecular descriptors are provided in the Appendix). These molecular descriptors had been frequently used in QSAR modeling and were known to be toxicologically and mechanistically relevant. Statistical classification models developed on this database had been used to assign mechanisms to other phenols.[14]

Classification using Regression. To use regression methods for classification purposes, the procedure described by Frank et al.,[15] i.e., multiresponse regression, was modified and followed in this study. The modified procedure is schematically shown in Figure 1 which uses a data set with three classes and three predictor variables as an example (the procedure can be directly generalized for data sets with more than three classes and/or predictor variables). Based on the original data set, three derived sets are created, each of which has a "primary" class. The primary class differs from one derived set to another such that each of the classes in the original data set is designated as the primary class in only one derived set. In these derived sets, the original class variable (in this study, the class variable was mechanism) is coded into 1 for the primary class and into 0 for *all* other classes. In this way, the original class variable, which is categorical, has numerical values of either 0 or 1 in the derived sets. Regression is then performed on each of the derived data sets with the *coded* class variable being the dependent variable and the predictors (in this study, the predictors were the molecular descriptors) being the independent variables. This regression step is referred to as the model step. Given values of the predictor variables, the regression models thus developed can be used to generate scores, i.e., predicted values for the coded variables in different derived sets. The scores obtained from models built on different derived data sets are then compared. The primary class corresponding to the coded variable with the highest score is assigned to an observation. For an observation that is not in the training set, the assigned mechanism serves as the predicted mechanism designation.

Regression Methods. The following regression methods were used in the model step described above: multiple linear regression (LinReg), logistic regression (LogReg), generalized additive model (GAM), locally weighted regression scatter plot smoothing (LOWESS), multivariate adaptive regression splines (MARS), and projection pursuit regression (PPR). LogReg fits nonlinear sigmoid function and GAM, LOWESS, MARS, and PPR, which are generally considered modern regression methods, fit unrestricted nonlinear regression. Since some of the regression methods may not be familiar to the readers, a brief introduction to those methods and relevant references are provided in Table 1. The scope of the present study was limited to linear or nonlinear regression methods. Other methods such as neural networks are generally not considered regression methods, and their use will be investigated in a separate study.

Model Development. For each of the before-mentioned regression methods, four regressions on coded variables were performed since there were four mechanisms in the database. Note that explicit variable selection procedure such as stepwise variable selection is possible for LinReg and LogReg but not for other regression methods. To ensure a fair comparison of the performances of the methods, all five molecular descriptors were used as model predictors in all the regression methods. LinReg and LogReg were conducted using SAS 8.2 (SAS Institute, Cary, North Carolina, U.S.A.), and GAM, LOWESS, MARS, and PPR were conducted using R 1.4.1 (The R Foundation, Vienna, Austria, downloadable at http://www.r-project.org). Key R codes are presented in the Appendix. The descriptor N_{hdon} was not included in the LOWESS model because of a lack of degrees of freedom. Keeping N_{hdon} and removing one of the other descriptors from the LOWESS model resulted in singularity problems. In this part of the study, the entire phenol data set[9] was used to develop regression models in the "model step". Using the models developed, the mechanism of each phenol in the data set was predicted. In this case, the misclassification rate could be viewed as a measure of "badness-of-fit" for the data set under investigation. Therefore, the predicted and a priori designated mechanisms were compared, and the error rate for each of the mechanism categories as well as the overall error rate were calculated.

Table 1. Statistical Regression Methods

method	brief introduction	ref
generalized additive model (GAM)	GAM extends the notion of linear regression by allowing linear functions of the predictors to be replaced by smooth spline functions. The model assumes that the spline predictor function has an additive structure.	16
locally weighted regression scatter plot smoothing (LOWESS)	LOWESS fits local multiple regression in a neighborhood of each data point. This is implemented via weighted regression techniques. Connecting several local weighted regressions makes the fitting nonlinear.	17
multivariate adaptive regression splines (MARS)	MARS is an adaptive procedure for regression. It can be viewed as a generalization of stepwise linear regression, but regression is fitted using piecewise linear basis functions.	18
projection pursuit regression (PPR)	PPR is an additive model using ridge function estimated by some flexible smoothing method. It can be viewed as nonlinear regression using linear combination of variables.	19

These error rates were referred to as resubstitution error rates.

Model Validation. Due to chance alone a certain validation set may be more challenging to one classification method than the other. Thus, to obtain a more reasonable evaluation of the performance of each of the classification methods, the method of 5-fold cross-validation (CV) or leave-20%-out CV was employed. This method had been proven to be an efficient way for model validation[20] and had been a common choice to validate QSAR models (e.g., refs 21 and 22). As such, the original data set[9] was divided into five disjoint subsets. Stratification across the mechanism was used when creating the subsets, i.e., the proportions of phenols with each of the mechanism categories (polar narcosis, weak acid respiratory uncoupling, pre-electrophilicity, and soft electrophilicity) were similar in all the subsets and the original set. Each CV subset contained approximately 20% of the phenols from each of the four mechanism categories in the database. The selection of phenols within each mechanism category was completely random. Each of the five subsets was used as a CV validation set. Corresponding to a CV validation set, the remaining four subsets were combined and served as a CV training set. Since the five subsets were disjoint, each phenol in the original data set was used in a validation set once.

For each pair of CV training and validation sets, regression models were developed as described previously on the CV training set and externally validated on the corresponding CV validation set. The mechanism predictions for phenols in the CV validation sets were noted. At the end of the 5-fold CV process, the predicted and designated mechanisms of all phenols were compared. The error rate for each of the four mechanism categories as well as the overall error rate were calculated. These error rates were referred to as prediction error rates.

Linear Discriminant Analysis Model. Since one specific aim of the present study was to compare the mechanism prediction accuracy of the multiresponse regression approach and that of discriminant analysis, linear discriminant analysis (LDA) was also conducted using the entire phenol data set.[9] SAS 8.2 was employed for this purpose. Proportional prior probability was used when developing linear discriminant functions. The resultant discriminant functions were validated using the 5-fold CV method described above. The resubstitution and prediction error rates were calculated. The performance of the LDA model was compared with the performance of the multiresponse regression approach.

Table 2. Resubstitution and Prediction Error Rates of LDA and the Multiresponse Regression Approach Based on Several Regression Methods

method	PN[a]	RU[b]	PE[c]	SE[d]	overall
LDA	6.5	22.2	22.2	17.4	10.9
	(7.2)	(38.9)	(22.2)	(34.8)	(14.5)
LinReg	4.6	50.0	59.3	60.9	20.8
	(4.6)	(61.1)	(59.3)	(65.2)	(22.2)
LogReg	2.0	33.3	33.3	43.5	12.7
	(3.9)	(33.3)	(37.0)	(56.5)	(15.8)
GAM	0.7	38.9	29.6	4.4	7.7
	(3.9)	(44.4)	(33.3)	(17.4)	(12.2)
LOWESS	2.0	22.2	40.7	8.7	9.0
	(3.9)	(22.2)	(44.4)	(17.4)	(11.8)
MARS	1.3	33.3	29.6	4.4	7.7
	(4.6)	(44.4)	(29.6)	(8.7)	(11.3)
PPR	1.3	16.7	7.4	0	3.2
	(4.6)	(27.8)	(18.5)	(11.7)	(9.9)

[a] Polar narcosis. [b] Respiratory uncoupling. [c] Proelectrophilicity. [d] Soft electrophilicity.

RESULTS

Model Development. The resubstitution error rates for LDA and multiresponse regressions are shown in Table 2. LDA had an overall error rate of 10.9% which was very similar to the results obtained by Aptula et al.[9] who used canonical discriminant analysis (LDA in the canonical space). Among the regression methods used in the multiresponse regression approach, LinReg did not perform well with an overall error rate of approximately 21%. LogReg, which was able to account to some extent the nonlinear pattern in the data, had better performance than LinReg. However, the overall error rate of LogReg was 12.7%, still higher than that of LDA. Multiresponse regression using the other four regression methods (GAM, LOWESS, MARS, and PPR) produced overall error rates smaller than that of LDA, with the biggest decrease in the overall error rate achieved by PPR. Examining the resubstitution error rates for individual mechanism categories, it appeared that while multiresponse regression based on GAM, LOWESS, MARS, and PPR provided considerably lower error rates for polar narcosis and soft electrophilicity, only the model based on PPR yielded a smaller error rate for respiratory uncoupling and proelectrophilicity.

Model Validation. The prediction error rates for LDA and multiresponse regressions are also shown in Table 2. LDA had an overall prediction accuracy of 14.5% which was again very similar to the results obtained by Aptula et al.[8] Among the regression methods used for multiresponse regression, LinReg did not provide satisfactory mechanism predictions

(overall error rate = 22.2%). LogReg also had an overall error rate (15.8%) lager than that of LDA. However, the rest of the regression methods resulted in improvements in the overall mechanism prediction accuracy over LDA. Similar to the resubstitution error rates, the lowest overall prediction error rate was achieved by the multiresponse regression based on PPR which was below 10%. This model also yielded smaller error rates for all mechanism categories, whereas models based on GAM, LOWESS, and MARS were outperformed in terms of prediction accuracy by LDA for at least one of the mechanism categories.

DISCUSSION

As Schultz et al.[2] pointed out, the current capability to predict mechanism from molecular structure has limited the applications of the mechanism-based QSARs in predictive toxicology. To solve this problem, methods that can be used to reliably predict toxicity mechanisms are needed. The statistical approach relating molecular structures to toxicity mechanisms appears to be a promising method that can be added to more traditional methods such as FATS and joint toxic action. Results in this study showed that with the use of (nonlinear) regression methods for classification purposes, the total error rate of mechanism prediction could be reduced to below 10%. Although this error rate is relatively low, false mechanism predictions are still not completely avoidable. In this situation, the approach taken by Russom et al.[6] seems reasonable. In addition to making mechanism designations, these authors also assigned confidence levels to mechanism designations, i.e., if several highly reliable methods give the same mechanism prediction to a certain compound, then the predicted mechanism has a high confidence level. The key is that these methods must be highly reliable, otherwise their output simply adds "noise" to the results of other methods and compromise their credibility. For this reason, a statistical mechanism classification and prediction method with a low error rate is desired, and efforts to improve the mechanism prediction accuracy are necessary.

Results in this study indicated that using certain (nonlinear) regression methods for mechanism classification can effectively increase the prediction accuracy. Among the regression methods considered in this study, LinReg, whose prediction accuracy was even worse than that of LDA, attempts to place a linear plane between compounds acting by different mechanisms. Doing so greatly simplified the effect of chemical structure on mechanism. However, as Cronin and Schultz[23] pointed out, biology and the modeling of biology is intrinsically a nonlinear phenomenon and always expecting a linear relationship is not realistic. LogReg, which fits sigmoid function, offered a better performance than LinReg partly because of its increased ability to account for nonlinearity compared to LinReg. The fact that the largest decrease in prediction error rate was achieved by a multiresponse regression based on PPR suggests that the effect of chemical structure (captured by the molecular descriptors) on toxicity mechanism is highly nonlinear.

Several sources can contribute to the error rates of statistical mechanism classification and prediction. The values of the molecular descriptors can be calculated using several different computing algorithms. Some of them can also be measured experimentally. Just as biological data contain errors, molecular descriptors are also subject to variability.[23] In addition to errors in the molecular descriptors, the a priori designated mechanisms in the database may not be completely accurate. Aptula et al.[9] pointed out that the mechanisms were assigned to the phenols in the database using simple structural rules. As mentioned in the Introduction, there are uncertainties associated with mechanism designations made by this method, especially for complex structures. A third source for the error rates is the classification methods used. However, as demonstrated in this study, contribution from this third source to the error rates can be reduced by choosing suitable statistical methods, classification algorithms, or combinations of the two.

Although having higher prediction accuracy, the classification models based on nonlinear regression are very difficult to interpret. Many of the nonlinear regression models cannot be explicitly expressed, and therefore the relative contribution of each of the molecular descriptors is unclear. The classification method using coded variables further complicates the interpretation of the resulting models. By contrast, LDA is superior in this regard because the discriminant functions can be explicitly expressed and interpretations of the coefficients are possible.

Together with the difficulty of interpretation, the lack of variable selection procedure in the nonlinear regression methods makes their use in identifying toxicologically relevant molecular descriptors limited. As mentioned previously, many aspects of the interaction between a toxicant and a living organism are poorly characterized. The molecular descriptors can be thought of ways to characterize various interactions in a toxic effect. For example, it is well-known that $\log K_{ow}$ can be used to characterize the uptake of a toxicant into the biophase and E_{LUMO} can be used to characterize the reaction between a toxicant and the site of toxic action. The nonlinear regression methods use all input variables that are provided, without showing in some way the relative importance of each of the variables. The goal of making accurate predictions is, in most situations, at conflict with the goal to interpret a model. Therefore, it is understood that the reduction in mechanism prediction error rate by multiresponse regression (based on nonlinear regression) is achieved at the sacrifice of model interpretability.

The five molecular descriptor used in the regression methods in this study did not go beyond those used in the study of Aptula et al.[9] All five molecular descriptors were used because first of all they were all toxicologically relevant, and second the results in Aptula et al.[9] seemed to indicate that removing some of them led to increased error rates. Although hundreds of molecular descriptors can be calculated with computer software, no more molecular descriptors were employed in the present study. In developing a QSAR for toxicity potency, it is not appropriate to include more descriptors simply to improve the statistical fit of the model and doing so may result in spurious or false correlations.[23] Similarly, it is not considered a good course of action to include more descriptors without appropriate justification in a mechanism classification algorithm, especially given the fact that the classification models based on nonlinear regression are difficult to interpret.

CONCLUSIONS

The use of regression for the classification and prediction of mechanism of toxic action was investigated in this study using a database for phenols. The multiresponse regression approach depicted in Figure 1 was taken in which the model step was accomplished by LinReg, LogReg, GAM, LOWESS, MARS, and PPR. The classification models were validated using the 5-fold cross validation (CV) method. Compared with more traditional classification method, i.e., linear discriminant analysis (LDA), it was found that the multiresponse regression based on PPR resulted in a decreased prediction error rate for every mechanism category as well as the overall error rate (9.9%). Multiresponse regression based on other regression methods did not provide improvements over LDA for at least one of the mechanism categories. While the error rates could be reduced, the classification models based on multiresponse regression were difficult to interpret compared to a model based on LDA.

ACKNOWLEDGMENT

S.R. wishes to thank Helen M. Delgado for motivation and support. The work of H.K. was supported by the Scholarly Research Grant Program of the College of Business Administration and the Professional Development Award of the University of Tennessee.

APPENDIX

Molecular Descriptors. logK_{ow}, logarithm of the octanol/water partition coefficient; pK_a, negative logarithm of the first dissociation constant; E_{LUMO}, energy of the lowest unoccupied molecular orbital; E_{HOMO}, energy of the highest occupied molecular orbital; N_{hdon}, number of hydrogen bond donors; ***Key R codes*** (x = data matrix containing the molecular descriptors, y = coded variable with numerical values of either 0 or 1); GAM, obj_gam(y∼s(logKow)+s(pKa)+s(Elumo)+s(Ehomo)+s(Nhdon),data=x,family=gaussian()); LOWESS, obj <- loess(y∼logKow+pKa+Elumo+Ehomo,x,control=loess.control(surface="direct")); MARS, obj <- mars(x,y); PPR, obj <- ppr(x,y,nterms=5).

REFERENCES AND NOTES

(1) Schultz, T. W.; Sinks, G. D.; Bearden, A. P. QSAR in aquatic toxicology: a mechanism of action approach comparing toxic potency to *Pimephales promelas*, *Tetrahymena pyriformis*, and *Vibrio fischeri*. In *Comparative QSAR;* Devillers, J., Ed.; Taylor & Francis: New York, 1998; pp 51−109.

(2) Schultz, T. W.; Cronin, M. T. D.; Walker, J. D.; Aptula, A. O. Quantitative structure−activity relationships (QSARs) in toxicology: a historical perspective. *J. Mol. Struct. (THEOCHEM)* **2003**, *622*, 1−22.

(3) Bradbury, S. P. Predicting modes of toxic action from chemical structure: an overview. *SAR QSAR Environ. Res.* **1994**, *2*, 89−104.

(4) McKim, J. M.; Bradbury, S. P.; Niemi, G. J. Fish acute toxicity syndromes and their use in the QSAR approach to hazard assessment. *Environ. Health Persp.* **1987**, *71*, 171−186.

(5) Veith, G. D.; Broderius, S. J. Rules for distinguishing toxicants that cause type I and type II narcosis syndromes. *Environ. Health Persp.* **1990**, *87*, 207−211.

(6) Russom, C. L.; Bradbury, S. P.; Broderius, S. J.; Hammermeister, D. E.; Drummond, R. A. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (*Pimephales promelas*). *Environ. Toxicol. Chem.* **1997**, *16*, 948−967.

(7) Schultz, T. W.; Sinks, G. D.; Cronin, M. T. D. Identification of mechanisms of toxic action of phenols to *Tetrahymena pyriformis* from molecular descriptors. In *Quantitative Structure−Activity Relationships in Environmental Sciences − VII;* Chen, F., Schüürmann, G., Eds.; SETAC Press: Pensacola, FL, 1997; pp 329−342.

(8) Bearden, A. P.; Schultz, T. W. Structure−activity relationships for *Pimephales* and *Tetrahymena*: a mechanism of action approach. *Environ. Toxicol. Chem.* **1997**, *16*, 1311−1317.

(9) Aptula, A. O.; Netzeva, T. I.; Valkova, I. V.; Cronin, M. T. D.; Schultz, T. W.; Kühne, R.; Schüürmann, G. Multivariate discrimination between modes of toxic action of phenols. *Quant. Struct.-Act. Relat.* **2002**, *21*, 12−22.

(10) Ren, S. Determining the mechanism of toxic action of phenols to *Tetrahymena pyriformis*. *Environ. Toxicol.* **2002**, *17*, 119−127.

(11) Ren, S. Use of molecular descriptors in separating phenols by three mechanisms of toxic action. *Quant. Struct.-Act. Relat.* **2002**, *21*, 486−492.

(12) Ren, S.; Schultz, T. W. Identifying the mechanism of aquatic toxicity of selected compounds by hydrophobicity and electrophilicity descriptors. *Tox. Lett.* **2002**, *129*, 151−160.

(13) Ren, S. Predicting three narcosis mechanisms of aquatic toxicity. *Tox. Lett.* **2002**, *133*, 127−139.

(14) Cronin, M. T. D.; Aynur, O. A.; Duffy, J. C.; Netzeva, T. I.; Rowe, P. H.; Valkova, I. V.; Schultz, T. W. Comparative assessment of methods to develop QSARs for the prediction of the toxicity of phenols to *Tetrahymena pyriformis*. *Chemosphere* **2002**, *49*, 1201−1221.

(15) Frank, E.; Wang, Y.; Inglis, S.; Holmes, G.; Witten, I. H. Using model trees for classification. *Machine Learning* **1998**, *32*, 63−76.

(16) Hastie, T.; Tibshirani, R. *Generalized Additive Models*; Chapman and Hall: New York, U.S.A., 1990.

(17) Cleveland, W. S.; Devlin, S. J. Locally weighted regression: an approach to regression analysis by local fitting. *J. Am. Statist. Ass.* **1988**, *83*, 596−610.

(18) Friedman, J. H. Multivariate additive regression splines. *Anls. Stat.* **1991**, *19*, 1−141.

(19) Friedman, J. H.; Stuetzle, W. Projection pursuit regression. *J. Am. Statisit. Ass.* **1981**, *76*, 817−823.

(20) Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: New York, U.S.A., 2001.

(21) Niculescu, S. P. Artificial neural networks and genetic algorithms in QSAR. *J. Mol. Struct.* (*Theochem*) **2003**, *622*, 71−83.

(22) Kaiser, K. L. E.; Niculescu, S. P. Using probabilistic neural networks to model the toxicity of chemicals to the fathead minnow (*Pimephales promelas*): a study based on 865 compounds. *Chemosphere* **1999**, *38*, 3237−3245.

(23) Cronin, M. T. D.; Schultz, T. W. Pitfalls in QSAR. *J. Mol. Struct. (THEOCHEM)* **2003**, *622*, 39−51.

CI034092Y