Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2007) 327–335
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Mode of action-based local QSAR modeling for the prediction of acute toxicity in the fathead minnow

Hua Yuan, Yong-Yan Wang, Yi-Yu Cheng *

Pharmaceutical Informatics Institute, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310027, China

Received 12 July 2006; received in revised form 13 November 2006; accepted 13 December 2006
Available online 16 December 2006

Abstract

The ultimate intention of quantitative structure–activity relationship (QSAR) study in toxicology is to predict the toxic potential of untested compounds with great accuracy. As QSAR has been based on the assumption that compounds from the same chemical domain will behave in similar manner, the QSAR model built upon the analogical chemicals is hypothesized to exhibit better performance than that derived from the miscellaneous data set. In this paper, the acute toxicity, 96 h LC_{50} (median lethal concentration) for the fathead minnow from database EPAFHM_v2a_617_1Mar05 served as the interested toxicity endpoint, and the mode of action (MOA) in toxic response was employed as a criterion to compartmentalize the chemical domains. MOA-based local QSAR models were built by partial least squares (PLS) regression for each subset with single mode of action such as Narcosis I, Narcosis II or Reactive, and global model was also developed for the combined data set containing several subsets above. By comparing the performances of these two types of models, the local models were superior to the global model in that the relative standard error (R.S.E.) of the former was much lower for both the training set and the test set of any subset. In addition, the influence of the reliability of MOA determination on the performance of local model was also investigated and the statistical results for subsets with MOAs at A and B confidence level were better than those at C and D confidence level. Therefore, the MOA-based local QSAR models are promising to improve the accuracy of toxicity prediction as long as the assessment of MOA is of high reliability.

Keywords: Quantitative structure–activity relationship (QSAR); Acute toxicity; Mode of action (MOA); Local/global models; Partial least squares (PLS)

1. Introduction

The toxicity of drug candidates is a factor as essential as the efficacy to be assessed in the drug development phase. Current preclinical safety evaluation programs use a combination of computational methods, mechanistic *in vitro* screening and *in vivo* animal tests to predict human toxicity [1]. In order to reduce the financial costs and protect animals, the European Commission has stated that the animal tests for the assessment of acute human health hazards should be replaced by the use of quantitative structure–activity relationship (QSAR) and *in vitro* test in the near future. Therefore, *in silico* prediction of toxicity from molecular structure has been receiving increasing attention. Numerous QSAR models for the prediction of various drug toxicities have been reported in recent years [2–5]. Because the data sets available for modeling toxicity are generally diverse in both molecular structure and mechanism of toxic action, few of the existing global QSAR models based on the miscellaneous data set with various mechanisms show good predictive ability. It is acknowledged that the QSAR with all the compounds involved acting by the same mechanism will model the biological data well [6]. Therefore, a clear mechanistic understanding of drug toxicity and the corresponding risk factors should improve the accuracy of the prediction of drug toxicity [7]. However, the mechanisms of toxic action for most compounds are unclear or quite complex, so the mechanism-based QSAR models cannot be set up for such data sets. As an alternate, QSAR modeling based on the mode of toxic action is a promising approach.

Mode of toxic action (MOA) is defined as a set of physiological and behavioral signs characterizing a type of adverse biological response, while a toxic mechanism is defined as the biochemical process underlying a given MOA [8,9]. Although inferior to the mechanistic model, the MOA-based local QSAR model is considered to be more exact than

* Corresponding author. Tel.: +86 571 879 52509; fax: +86 571 879 51138.
E-mail addresses: yh_cathy@zju.edu.cn (H. Yuan), chengyy@zju.edu.cn (Y.-Y. Cheng).

1093-3263/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2006.12.009

the global model in that it covers a more similar chemical domain. Acute toxicity is one of the most important toxicological endpoints in the routine evaluation of drug safety, especially for the environment related compounds. Russom et al. [10] have made great effort in assessing the modes of action for about 470 chemicals with toxic endpoint of 96 h LC_{50} (median lethal concentration) for the fathead minnow through a broad range of investigations, such as dose-response relationships and behavioral responses associated with 96 h LC_{50} bioassays, joint toxic action studies, fish acute toxicity syndrome (FATS) investigations, and an examination of toxicological literature specific to the issue of toxicodynamic classifications. The chemicals were classified as several subsets with different MOAs of narcosis (I, II or III), oxidative phosphorylation uncoupling, respiratory inhibition, electrophilic/proelectrophilic reactivity, acetylcholinesterase (AChE) inhibition, or central nervous system (CNS) seizure mechanisms. According to the amount of available information for a given compound, four different levels of confidence (A, B, C and D in descent trend) were assigned to the determination of each mode of action. A level, the highest level of confidence, required the information about FATS, joint toxic action determination, or chemical-specific literature information. To determine the B level of confidence, the required information included behavior syndrome, LC_{50} ratio, and excess toxicity (Te) values that were consistent with that observed for structurally similar compounds whose mode of action assignment was based on A level information. C level of certainty was attained when there were less than three level B components, but information such as the concentration/response slope, behavior comments, and/or chemical similarity to prototypical compounds was available to support the assessment. A level D certainty was indicated when there was no confidence in a mode of action classification due to insufficient data. An empirical database, EPAFHM_v2a_617_1Mar05 [11], was developed with chemical structures and their corresponding modes of toxic action. This database made it possible for the MOA-based QSAR investigation for the aquatic toxicity of environmental and industrial chemicals. Based on the knowledge about the mode of action and substructural rules, an expert system, ASTER (assessment tools for the evaluation of risk) [12], was developed with the ability of assigning acute modes of toxic action to compounds, associating with QSAR investigation for narcosis and oxidative phosphorylation uncoupling modes of action. However, the QSAR models in ASTER were mainly derived from the relationship between log P (octanol/water partition coefficient) and acute toxicity, more structural factors may be taken into account to improve the correlation and prediction capabilities of those QSARs. Though the MOA-based QSAR modeling has been explored [13], no reports have studied the reliability and validity about the improvement of such local models over global model in the prediction of toxicity.

Motivated by the above uncertainties, this paper designed the scheme as follows (shown in Fig. 1): (1) to build local models for each subset with single MOA at A and B or C and D confidence level, respectively, together with global model based on the combined data set at the same confidence level; (2) to compare the performance of local model with that of global model; (3) to evaluate the effect of the reliability of MOA determination on the performance of local model by comparing the MOA-based models at A and B confidence level with those at C and D confidence level.

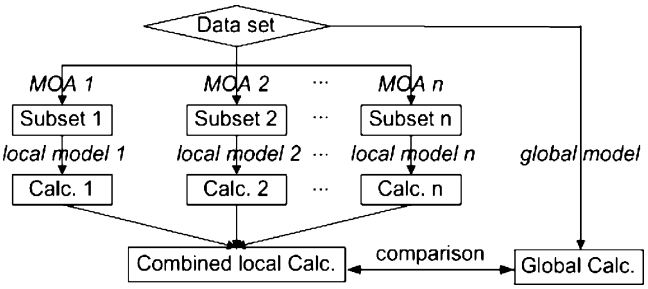

Fig. 1. The scheme of MOA-based QSAR modeling.

2. Materials and methods

2.1. Data set

Database, EPAFHM_v2a_617_1Mar05, contains 617 compounds with 96 h LC_{50} (mg/L), about 470 of which were assigned to eight modes of action with an A, B, C, or D level of confidence, respectively. After the elimination of stereo-isomers, salts, metal organic compounds and mixtures, the compounds with same MOA at A and B confidence level or C and D confidence level were gathered together to form subsets. In order to build a valid QSAR model, the subsets with compounds less than 20 were ignored. Therefore, three subsets with MOA of Narcosis I, Narcosis II and Reactive at A and B confidence level and two subsets with MOA of Narcosis I and Reactive at C and D confidence level were obtained for MOA-based QSAR modeling. In addition, the subsets at A and B or C and D confidence level were integrated into a whole set for the generation of global model. One fifth of each data set was used as the test set, and the remained compounds served as the training set. The details about each data set were shown in Table 1. The LC_{50} values of the concerned compounds varied between 0.0032 and 75200.0000 mg/L. In order to attain a more uniform range and make the distribution of the data more normal, LC_{50} was transformed to $\log(LC_{50})$ with the range from −2.4949 to 4.8762.

2.2. Generation of molecular descriptors

Dragon v. 5.2 [14] was employed to calculate 926 molecular descriptors (except 3D) contained in blocks 1–10, 17–18 and 20 of the descriptor package. These descriptors included constitutional descriptors, topological descriptors, walk and path counts, connectivity indices, information indices, 2D autocorrelations, edge adjacency indices, BCUT descriptors, topological charge indices, eigenvalue-based indices, functional group counts, atom-centered fragments, and molecular

Table 1
Data sets based on MOA at different confidence level

Confidence level	Data set	MOA	n[a]	n_C[b]	n_P[c]
A and B	Subset 1	Narcosis I	131	105	26
	Subset 2	Narcosis II	29	23	6
	Subset 3	Reactive	39	31	8
	Whole set 1	Combined[d]	199	159	40
C and D	Subset 4	Narcosis I	97	78	19
	Subset 5	Reactive	52	42	10
	Whole set 2	Combined[d]	149	120	29

[a] Total number of samples in each data set.
[b] Number of samples in the training set for calibration.
[c] Number of samples in the test set for prediction.
[d] The combination of several types of MOA other than a single MOA.

properties. The 2D molecular structures with hydrogens (H) were used as input.

2.3. Variable selection and preprocessing

Based on the calculated 926 molecular descriptors, objective variable selection was carried out to delete the redundant and highly correlated information according to the following steps: (1) eliminating descriptors which have same values for greater than 90% of the compounds by same test; (2) eliminating descriptors whose pairwise correlation with other descriptors exceeds 0.85 by pairwise correlation test. After objective variable selection based upon each training set, the number of descriptors remained was 156, 87, 134, 153, 133, 165 and 162 for subsets 1–5 and whole sets 1 and 2, respectively.

Because these descriptors characterized the molecular structural information from extensive perspectives, their magnitudes were largely different. In order to prevent the descriptors with larger ranges from outweighing those with smaller ranges, the original descriptors needed to be transformed into a more consistent range prior to QSAR modeling. Standardization was one of the most used approaches to transform descriptors. Therefore, the descriptors remained after variable selection were standardized with Eq. (1), i.e., each descriptor was divided by its standard deviation:

$$x_{ij}^{\mathrm{pre}} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2/(n-1)}} \quad (1)$$

where x_{ij} was the jth descriptor of the ith compound and $\bar{x}_j$ was the mean of the jth descriptor of n compounds.

2.4. Partial least squares (PLS)

Partial least squares (PLS) regression is one of the most-used algorithms for multivariate calibration. It models the dependent variable with a small number of new, independent latent variables extracted by a linear transformation of the original independent descriptors to a limited number of orthogonal factors. Therefore, it is particularly useful when the independent variables are much more than the number of samples. Many articles have reported the wide application of PLS to chemistry, economics, social sciences, etc. [15–17]. The original molecular descriptors left after variable selection were standardized prior to PLS. The optimal number of latent variables (optlv) was determined by 10-fold cross validation against the training set with the aim of achieving the minimum root mean square error. Because the size of each MOA-based data set was not very large, the maximum number of latent variables was set as 8. All these calculations together with variable selection were carried out in the Matlab environment with programs written in-house.

2.5. Performance evaluation

The square of correlation coefficient (R^2), root mean square error (RMSE) and relative standard error (R.S.E.) were used for the evaluation of QSAR performance.R_C^2, $\mathrm{RMSE_C}$ and $\mathrm{R.S.E._C}$ were for the training set, and R_P^2, $\mathrm{RMSE_P}$ and $\mathrm{R.S.E._P}$ for the test set. The formulations of RMSE and R.S.E. were shown as follows:

$$\mathrm{RMSE_C} = \sqrt{\frac{\sum_{i=1}^{n_C} (y_{\mathrm{exp}} - y_{\mathrm{calc}})^2}{n_C - \mathrm{optlv}}} \quad (2)$$

$$\mathrm{RMSE_P} = \sqrt{\frac{\sum_{i=1}^{n_P} (y_{\mathrm{exp}} - y_{\mathrm{calc}})^2}{n_P}} \quad (3)$$

$$\mathrm{R.S.E._C} = \sqrt{\frac{\sum_{i=1}^{n_C} (y_{\mathrm{exp}} - y_{\mathrm{calc}})^2}{\sum_{i=1}^{n_C} y_{\mathrm{exp}}^2}} \quad (4)$$

$\mathrm{R.S.E._P}$ is as same as $\mathrm{R.S.E._C}$ in the form. In Eqs. (2)–(4), y_{exp} and y_{calc} are the experimental and calculated $\log(\mathrm{LC_{50}})$s, respectively; n_C and n_P are the numbers of samples in the training set and test set, respectively.

3. Results and discussion

3.1. Results of subsets 1–3 and whole set 1 with MOAs at A and B confidence level

The details about subsets 1–3 and whole set 1 were listed in Table 1. Taking subset 1 as an example, the statistical analysis and results were illustrated. The test set was formed by taking every five compounds of subset 1, and the remained made up of the training set. The training set and test set included 105 and 26 compounds, respectively. After objective variable selection, 156 descriptors were left and then preprocessed for PLS modeling. Based on the training set, 10-fold cross validation was performed and the value of optlv was determined as 8. PLS model was built for the training set with 8 latent variables, with which $\log(\mathrm{LC_{50}})$s of the test set were predicted. Such PLS model based on a subset with single MOA was called as MOA-based local model, and the corresponding statistical quantities (R^2, RMSE and R.S.E.) were listed in Table 2 marked by local model. The first eight latent variables have explained 94.04% of the 156 descriptors. According to the loading weight vectors for

Table 2
The statistical results of subsets 1–3 and whole set 1

Data set	Model[a]	R_C^2	$RMSE_C$	$R.S.E._C$	optlv	R_P^2	$RMSE_P$	$R.S.E._P$
Subset 1	Local	0.9702	0.1970	0.0871	8	0.8497	0.4673	0.2030
	Global	0.7593	0.6199	0.2810	3	0.6934	0.7676	0.3334
Subset 2	Local	0.9002	0.2093	0.1016	3	0.7908	0.4059	0.2355
	Global	0.3576	0.5379	0.2610	3	0.6465	0.7212	0.4184
Subset 3	Local	0.8846	0.3792	0.2916	4	0.1681	0.7375	0.7793
	Global	0.5626	1.2544	0.9823	3	0.0690	1.7775	1.8784
Whole set 1	Local	0.9589	0.2337	0.1166	–[b]	0.8151	0.5248	0.2601
	Global	0.5709	0.7555	0.3769	3	0.2456	1.0455	0.5182

[a] Local model based on a single subset, while global model based on whole set 1 consisting of subsets 1–3.
[b] The optlv of local model for whole set 1 can not be determined because the results were the integration of three local models based on subsets 1–3 other than the PLS model based on whole set 1.

the first two latent variables, such descriptors as *Me*, *ATS3v*, *HNar*, *PW2* and *BEHm2* were the most important descriptors responsible for $\log(LC_{50})$ of subset 1. The description about each descriptor was listed in Table 3. Subsets 2 and 3 were similar to subset 1 in the statistical process.

Whole set 1 was composed of subsets 1–3. PLS modeling was carried out against 159 compounds of the training set, and the optlv was determined as 3. Just corresponding to the local model, this PLS model was called as the global model. For comparison, the $\log(LC_{50})$s of subsets 1–3 were also calculated by this global model and the statistical quantities were displayed below those of each local model in Table 2. In addition, the calculated $\log(LC_{50})$s of subsets 1–3 resulted from three local models were combined to a whole, and the statistical quantities were also calculated and shown in Table 2 marked by the local model of whole set 1. The calculated $\log(LC_{50})$s by local and global models were all listed in Tables S1 and S2 of Supplementary Information.

3.2. Results of subsets 4-5 and whole set 2 with MOAs at C and D confidence level

For MOAs at C and D confidence level, only two subsets have more than 20 samples, i.e., subset 4 containing 97 compounds with MOA of Narcosis I and subset 5 containing 52 compounds with MOA of Reactive (see Table 1). The combination of subsets 4 and 5 made up of whole set 2. As the statistical analysis described in the preceding section, local models for subsets 4 and 5 were constructed with optlv of 8 and 2, respectively. Global model based on whole set 2 was built with optlv of 3. The statistical results of local and global models for subsets 4–5 and whole set 2 were all shown in Table 4. Five

Table 3
The most important molecular descriptors in QSAR models of subsets 1–3 and whole set 1

Data set	Symbol	Description	Block
Subset 1	Me	Mean atomic Sanderson electronegativity (scales on Carbon atom)	Constitutional descriptors
	ATS3v	Broto-Moreau autocorrelation of a topological structure–lag 3/weighted by atomic van der Waals volumes	2D autocorrelations
	HNar	Narumi harmonic topological index	Topological descriptors
	PW2	Path/walk 2, Randić shape index	Topological descriptors
	BEHm2	Highest eigenvalue no. 2 of Burden matrix/weighted by atomic masses	Burden eigenvalues
Subset 2	Me	Mean atomic Sanderson electronegativity (scales on Carbon atom)	Constitutional descriptors
	PCR	Ratio of multiple path count over path count	Walk and path counts
	Mv	Mean atomic van der Waals volume (scaled on Carbon atom)	Constitutional descriptors
	JhetZ	Balaban-type index from Z weighted distance matrix (Barysz matrix)	Topological descriptors
	HNar	Narumi harmonic topological index	Topological descriptors
Subset 3	Me	Mean atomic Sanderson electronegativity (scales on Carbon atom)	Constitutional descriptors
	BELm1	Lowest eigenvalue n. 1 of Burden matrix/weighted by atomic masses	Burden eigenvalues
	AAC	Mean information index on atomic composition	Information indices
	GATS2p	Geary autocorrelation-lag 2/weighted by atomic polarizabilities	2D autocorrelations
	IC2	Information content index (neighborhood symmetry of 2-order)	Information indices
Whole set 1	Me	Mean atomic Sanderson electronegativity (scales on Carbon atom)	Constitutional descriptors
	BELm1	Lowest eigenvalue no. 1 of Burden matrix/weighted by atomic masses	Burden eigenvalues
	HNar	Narumi harmonic topological index	Topological descriptors
	ATS3m	Broto-Moreau autocorrelation of a topological structure, lag 3/weighted by atomic masses	2D autocorrelations
	BEHm2	Highest eigenvalue no. 2 of Burden matrix/weighted by atomic masses	Burden eigenvalues

Table 4
The statistical results of subsets 4–5 and whole set 2

Data set	Model[a]	R^2_C	$RMSE_C$	$R.S.E._C$	optlv	R^2_P	$RMSE_P$	$R.S.E._P$
Subset 4	Local	0.9661	0.2405	0.1224	8	0.7813	0.6448	0.2757
	Global	0.7851	0.6353	0.3346	3	0.6349	0.9829	0.4203
Subset 5	Local	0.5273	0.7273	0.5407	2	0.6724	0.6426	0.6440
	Global	0.5043	0.8861	0.6505	3	0.6153	0.8602	0.8620
Whole set 2	Local	0.8547	0.4642	0.2712	–[b]	0.7638	0.6441	0.3251
	Global	0.6490	0.7214	0.4215	3	0.5050	0.9424	0.4756

[a] Local model based on a single subset, while global model based on whole set 2 consisting of subsets 4–5.
[b] The optlv of local model for whole set 2 can not be determined because the results were the integration of two local models based on subsets 4–5 other than the PLS model based on whole set 2.

main descriptors in the QSAR models of each subset were displayed in Table 5. The calculated log (LC_{50})s were listed in Tables S3 and S4 of Supplementary Information.

3.3. Comparison of local model with global model at the same confidence level

As seen from Tables 2 and 4, the performance of local model was superior to that of global model in that R^2 of the former was much higher, while RMSE and R.S.E. were much lower for ether the training set or the test set of each subset. The calculated log(LC_{50})s of whole set 1 by local models and global model were plotted against the experimental ones, as shown in Figs. 2 and 3, respectively. Just like whole set 1, the plots of calculated log(LC_{50})s by local models and global model versus experimental ones for whole set 2 were also displayed in Figs. 4 and 5, respectively. It was observed that the training set and test set were evenly distributed across the toxicity value range. By comparison of Figs. 2 with 3, and Figs. 4 with 5, the calculated log(LC_{50})s by local models were apparently closer to the experimental ones than those by global model for both the training set and the test set, which also confirmed the advantage of local model over the global model.

The selection of appropriate QSAR models is one of the crucial factors for accurate prediction of toxicity. Although the application domain of global model is wide enough to cover with the compounds with various MOAs included in the whole data set, the global model is too general and coarse to characterize the exact structure–activity relationship of each subset. However, the MOA-based local model can describe the QSAR of the interested subset more precisely and elaborately because the compounds with the same toxic mode of action would possibly behave in a toxicologically similar manner. Therefore, the performance of local model is superior to that of global model.

3.4. Comparison of local models with MOAs at A and B with those at C and D confidence level

Is MOA-based local model always much better than global model despite the reliability of MOA determination, that is to say, does the accuracy of MOA determination have influence on the performance of local model? This paper investigated two groups of chemicals separately, one with MOA at A and B confidence level such as subsets 1–3 and whole set 1, the other with MOA at C and D confidence level such as subsets 4–5 and whole set 2. MOAs at A and B confidence level were more

Table 5
The most important molecular descriptors in QSAR models of subsets 4–5 and whole set 2

Data set	Symbol	Description	Block
Subset 4	Me	Mean atomic Sanderson electronegativity (scales on Carbon atom)	Constitutional descriptors
	BELm1	Lowest eigenvalue no. 1 of Burden matrix/weighted by atomic masses	Burden eigenvalues
	PW2	Path/walk 2, Randić shape index	Topological descriptors
	ESpm02d	Spectral moment 02 from edge adjacency matrix weighted by dipole moments	Edge adjacency indices
	HNar	Narumi harmonic topological index	Topological descriptors
Subset 5	Me	Mean atomic Sanderson electronegativity	Constitutional descriptors
	BELm1	Lowest eigenvalue n. 1 of Burden matrix/weighted by atomic masses	Burden eigenvalues
	PCR	Ratio of multiple path count over path count	Walk and path counts
	HNar	Narumi harmonic topological index	Topological descriptors
	EEig01d	Eigenvalue 01 from edge adjacency matrix weighted by dipole moments	Edge adjacency indices
Whole set 2	Me	Mean atomic Sanderson electronegativity	Constitutional descriptors
	BELm1	Lowest eigenvalue n. 1 of Burden matrix/weighted by atomic masses	Burden eigenvalues
	PW2	Path/walk 2, Randić shape index	Topological descriptors
	HNar	Narumi harmonic topological index	Topological descriptors
	EEig01d	Eigenvalue 01 from edge adjacency matrix weighted by dipole moments	Edge adjacency indices

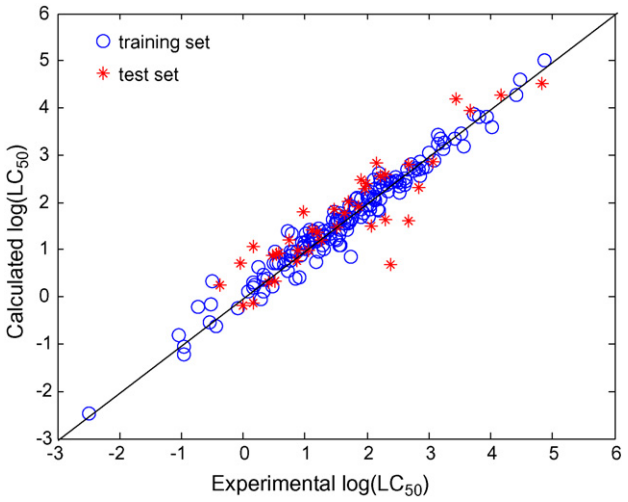

Fig. 2. The plot of calculated vs. experimental log(LC_{50})s for whole set 1 with local model.

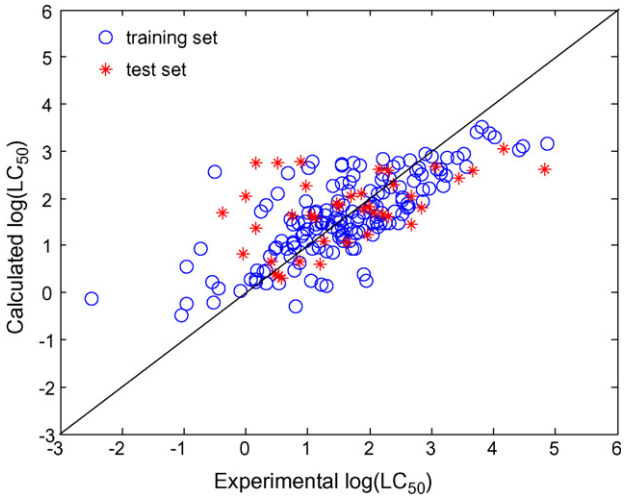

Fig. 3. The plot of calculated vs. experimental log(LC_{50})s for whole set 1 with global model.

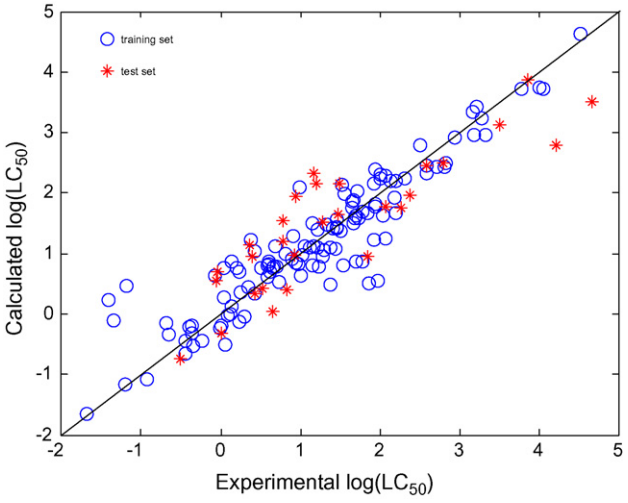

Fig. 4. The plot of calculated vs. experimental log(LC_{50})s for whole set 2 with local model.

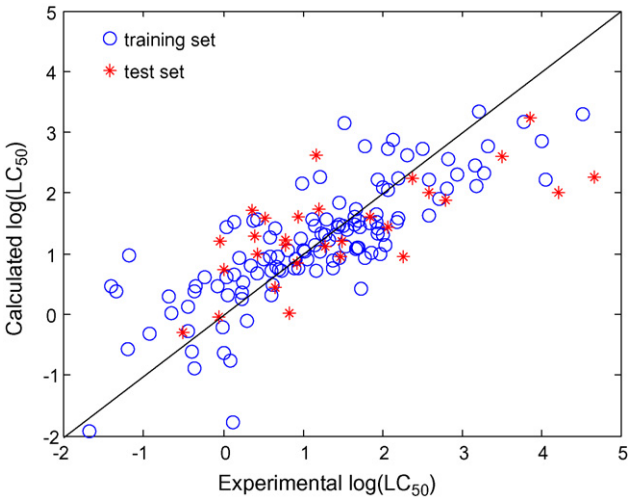

Fig. 5. The plot of calculated vs. experimental log(LC_{50})s for whole set 2 with global model.

credible than those at C and D confidence level due to more information available for the determination of the former. To evaluate such effect, an evaluation criterion EC was defined as follows:

$$\mathrm{EC} = \frac{\mathrm{R.S.E.}^{\mathrm{l}}}{\mathrm{R.S.E.}^{\mathrm{g}}} \tag{4}$$

where $\mathrm{R.S.E.}^{\mathrm{l}}$ and $\mathrm{R.S.E.}^{\mathrm{g}}$ were the relative standard error (R.S.E.) calculated by local model and global model for the same data set, respectively. The smaller value of EC indicated the more superiority of the local model over the global model.

Narcosis I and Reactive modes of action together with the combination of them at A and B and C and D confidence levels were used as the objects for comparison. The R.S.E. and EC values calculated by local model and global model for the training set and test set of subsets 1, 3, 4, 5 and whole sets 1 and 2 were listed in Table 6. For the same MOA at different confidence level, the common trend could be observed that the EC values of subsets with MOAs at A and B confidence level were smaller than those with the corresponding MOAs at C and D confidence level, which indicated that more improvements of the local model over the global model could be achieved if the

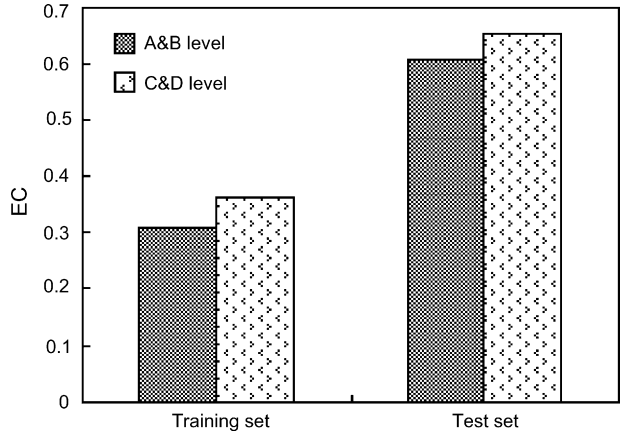

Fig. 6. Plot of EC for Narcosis I MOA at A and B level and C and D level.

Table 6
The EC values of subsets with MOAs at A and B or C and D confidence level

MOA	Subset[b]	Confidence[c]	R.S.E.$_C^{l}$[d]	R.S.E.$_C^{g}$[e]	EC$_{Cf}$[f]	R.S.E.$_P^{l}$[g]	R.S.E.$_P^{g}$[h]	EC$_P$[i]
Narcosis I	1	A and B	0.0871	0.2810	0.3100	0.2030	0.3334	0.6089
	4	C and D	0.1224	0.3346	0.3658	0.2757	0.4203	0.6560
Reactive	3	A and B	0.2916	0.9823	0.2969	0.7793	1.8784	0.4149
	5	C and D	0.5407	0.6505	0.8312	0.6440	0.8620	0.7471
Combined[a]	W-1	A and B	0.1166	0.3769	0.3094	0.2601	0.5182	0.5019
	W-2	C and D	0.2712	0.4215	0.6434	0.3251	0.4756	0.6836

[a] The combination of several types of MOA other than a single MOA.
[b] Subsets 1, 3, 4, and 5 are corresponding to the subsets in Table 1, while W-1 and W-2 are the abbreviations of whole set 1 and whole set 2, respectively.
[c] The confidence level of MOA.
[d] R.S.E. calculated by local model for the training set.
[e] R.S.E. calculated by global model for the training set.
[f] EC of the training set.
[g] R.S.E. calculated by local model for the test set.
[h] R.S.E. calculated by global model for the test set.
[i] EC of the test set.

mode of action was determined at higher confidence level. It could be understood that the reliability of MOA assignment had significant influence on the validity of application domain of the local model. If some compounds in subset were wrongly defined in MOA, the performance of the QSAR model would be impaired for both the calibration and the prediction. Therefore, the credible assessment of mode of toxic action is the prerequisite for the accuracy of MOA-based QSAR modeling and prediction of toxicities. Although an expert system, ASTER [12] has been developed to predict the modes of acute toxic action based on the substructural fragments, more thorough and comprehensive understanding of modes of toxic action need to be proceeded and higher correctness of MOA assessment is to be pursued.

To display more intuitively the influence of reliability of MOA determination on the superiority of MOA-based local models over global models, the ECs of training set and test set for Narcosis I, Reactive and combined MOAs at A and B confidence level were plotted for comparison with those at C and D confidence level in Figs. 6–8. In the histograms, the black rectangles were all lower than the adjacent white floscular ones for these three types of MOA studied, which indicated that the MOA-based local models at A and B confidence were better than those at C and D confidence level, especially for Reactive mode of action (see Fig. 7). By comparison of Fig. 6 with Fig. 7, it may also be inferred that the correctness of MOA assignment for Reactive at C and D confidence level was not as good as that for Narcosis I at C and D confidence level.

3.5. Discussion of the molecular descriptors

Seen from Tables 3 and 5, the most important molecular descriptors for each local model differed from each other, which conformed to the fact that the chemical–biological interactions were distinct for different mode of toxic action. The crucial molecular descriptors for each local model were basically from several blocks of descriptors, such as constitutional descriptors, topological descriptors, Burden eigenvalues, 2D autocorrelations and edge adjacency indices. The main structural information reflected by these descriptors was related to molecular polarity, size and shape. According to McFarland [18], toxicity was the result of the penetration of a toxicant into the biophases and the interaction of the toxicant with the site of action. As such, acute toxicity in the fathead minnow also

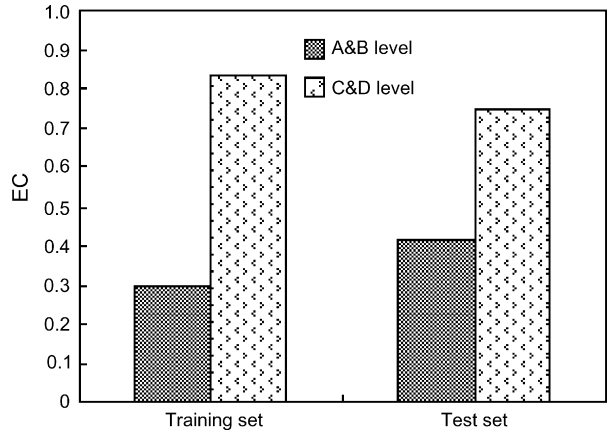

Fig. 7. Plot of EC for Reactive MOA at A and B level and C and D level.

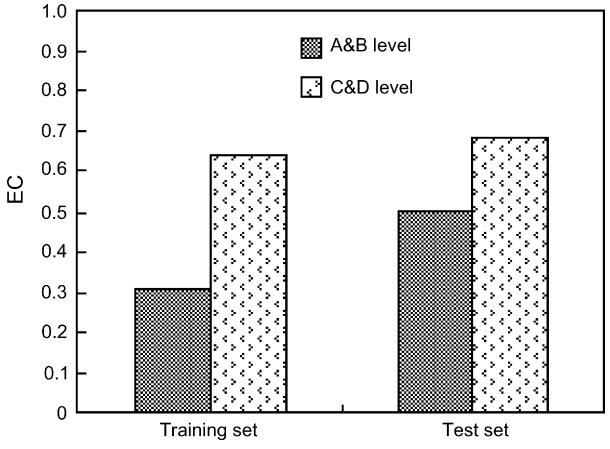

Fig. 8. Plot of EC for combined MOAs at A and B level and C and D level.

included two processes: the penetration into the cell membrane and the interaction with the target. The molecular polarity was closely related to the hydrophobicity which facilitated the penetration of chemicals into the lipid bilayer. *Me* (mean atomic Sanderson electronegativity), characterized the molecular polarity and played an important role in the QSAR models of all the subsets and whole sets. In addition, *GATS2p*, *EEig01d* and *ESpm02d* were also related to the electronic effect of molecule. The molecular size was another crucial factor determining the penetration process, which could be proved by the evidence that descriptors describing molecular size, such as *Mv*, *BELm1*, *BEHm2* [19], *ATS3m* and *ATS3v* [20] appeared among the most important descriptors of each local and global model. Although the detailed mechanism remains unclear, the molecular shape was thought to have effect on the binding of chemicals to the specific protein target, so the descriptors describing molecular shape such as *PW2* [21], *IC2* [22] and *PCR* [23] might be responsible for this interaction. As assumed that 3D molecular descriptors are more capable of characterizing molecular size, shape and electronic features than 2D molecular descriptors, the performances of QSAR models involving 3D descriptors may be improved despite of the more complex calculations. Although the QSAR investigation with the consideration of 3D structural information goes beyond the scope of this paper, it is a good subject for further studies.

By comparison between the subsets with same MOA at different confidence level, such as subsets 1 and 4, subsets 3 and 5, the principal descriptors were partially consistent. For example, *Me*, *ATS3v*, *HNar*, *PW2* and *BEHm2* were the five most important descriptors for subset 1 with Narcosis I MOA at A and B confidence level, while *Me*, *BELm1*, *PW2*, *Espm02d* and *HNar* for subset 4 with Narcosis I MOA at C and D confidence level. There were three descriptors coexisting in both local models. The other two different descriptors might result from the fact that not all the compounds in subset 4 really kept to the mode of action of Narcosis I due to the low confidence level, which could also be inferred by the large discrepancy of statistical results between subsets 1 and 4. Similar phenomenon could be observed in subsets 3 and 5 with MOA of Reactive at different confidence levels. For the combined data sets (whole sets 1 and 2), the main descriptors were almost from the important descriptors of the component subsets. For example, five main descriptors of whole set 2 were all included in the descriptor sets of subsets 4 and 5, and four descriptors (except *ATS3m*) of whole set 1 were consistent with the important descriptors of subsets 1–3.

4. Conclusions

In silico prediction of toxic potential of chemicals by QSAR models has become an increasingly important approach for risk assessment. The prediction accuracy of QSAR models is significant for its wide application. From the investigation of this paper, the performance of MOA-based local model is much superior to that of global model as long as the assessment of MOA is of high reliability. Therefore, the local models based on the mode of toxic action are preferred to global model for the improvement of the accuracy of toxicity prediction. However, the modes of action for most toxicities other than acute toxicity are not available due to the inadequately systematic understanding of toxicological behavior. Comprehensive work on the determination and high reliability of modes of action for a variety of toxicities, similar to that with the fathead minnow, is an essential requirement for the wide application of MOA-based QSAR modeling. The results and conclusions of this study are beneficial to researchers attempting to model other biological activities in addition to toxicity.

Acknowledgements

This work was financially supported by the key Program in Major Research Plan of National Natural Science Foundation of China (No. 90209005) and National Basic Research Program of China (No. 2005CB523402).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2006.12.009.

References

[1] D.E. Johnson, G.H.I. Wolfgang, Predicting human safety: screening and computational approaches, Drug Discov. Today 5 (2000) 445–454.
[2] R. Benigni, A. Giuliani, R. Franke, A. Gruska, Quantitative structure–activity relationships of mutagenic and carcinogenic aromatic amines, Chem. Rev. 100 (2000) 3697–3714.
[3] L. He, P.C. Jurs, Assessing the reliability of a QSAR model's predictions, J. Mol. Graph Model 23 (2005) 503–523.
[4] E. Estrada, G. Patlewicz, M. Chamberlain, D. Basketter, S. Larbey, Computer-aided knowledge generation for understanding skin sensitization mechanisms: the TOPS-MODE approach, Chem. Res. Toxicol. 16 (2003) 1226–1235.
[5] J. Padmanabhan, R. Parthasarathi, V. Subramanian, P.K. Chattaraj, Group philicity and electrophilicity as possible descriptors for modeling ecotoxicity applied to chlorophenols, Chem. Res. Toxicol. 19 (2006) 356–364.
[6] J.C. Dearden, *In silico* prediction of drug toxicity, J. Comput. Aided Mol. Des. 17 (2003) 119–127.
[7] A.P. Li, Accurate prediction of human drug toxicity: a major challenge in drug development, Chem. Biol. Interact. 150 (2004) 3–7.
[8] G.M. Rand, P.G. Wells, L.S. McCarty, Introduction to aquatic toxicology, in: Fundamentals of Aquatic Toxicology: Effects, Environmental Fate and Risk Assessment, Taylor & Francis, Washington, DC, 1995, pp. 3–67.
[9] S. Spychera, M. Nendzab, J. Gasteigera, Comparison of different classification methods applied to a mode of toxic action data set, QSAR Comb. Sci. 23 (2004) 779–791.
[10] C.L. Russom, S.P. Bradbury, S.J. Broderius, D.E. Hammermeister, R.A. Drummond, Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (*Pimephales promelas*), Environ. Toxicol. Chem. 16 (1997) 948–967.
[11] http://www.epa.gov/nheerl/dsstox.
[12] C.L. Russom, E.B. Anderson, B.E. Greenwood, A. Pilli, ASTER: An integration of the AQUIRE data base and the QSAR system for use in ecological risk assessments, Sci. Total Environ. 109/110 (1991) 667–670.
[13] S.P. Bradbury, C.L. Russom, G.T. Ankley, T.W. Schultz, J.D. Walker, Overview of data and conceptual approaches for derivation of quantitative

structure–activity relationships for ecotoxicological effects of organic chemicals, Environ. Toxicol. Chem. 22 (2003) 1789–1798.
[14] R. Todeschini, V. Consonni, A. Mauri, M. Pavan, Dragon, Version 5.2, Milano Chemometrics and QSAR Research Group, University of Milano-Bicocca, Milan, Italy, 2004.
[15] O.E. De Noord, E.H. Theobald, Multilevel component analysis and multilevel PLS of chemical process data, J. Chemometr. 19 (2005) 301–307.
[16] V. Lengard, M. Kermit, 3-Way and 3-block PLS regressions in consumer preference analysis, Food Qual. Prefer. 17 (2006) 234–242.
[17] W.W. Chin, B.L. Marcelin, P.R. Newsted, A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study, Inform. Syst. Res. 14 (2003) 189–217.
[18] J.W. McFarland, On the parabolic relationship between drug potency and hydrophobicity, J. Med. Chem. 13 (1970) 1092–1196.
[19] F.R. Bruden, Molecular identification number for substructure searches, J. Chem. Inform. Comput. Sci. 29 (1989) 221–225.
[20] P. Broto, G. Moreau, C. Vandycke, Molecular structures: perception, autocorrelation descriptor and sar studies, Eur. J. Med. Chem. 19 (1984) 66–70.
[21] M. Randić, Novel shape descriptors for molecular graphs, J. Chem. Inform. Comput. Sci. 41 (2001) 607–613.
[22] V.R. Magnuson, D.K. Harriss, S.C. Basak, Information indices, in: R.B. King (Ed.), Studies in Physical and Theoretical Chemistry, Elsevier, Amsterdam, 1983, pp. 178–191.
[23] G. Rücker, C. Rücker, Counts of all walks as atomic and molecular descriptors, J. Chem. Inform. Comput. Sci. 33 (1993) 683–695.