Archives of Environmental Contamination and Toxicology (2023) 84:214–226
https://doi.org/10.1007/s00244-023-00980-9

A Quantitative Structure-Activity Relationship Approach to Determine Biotoxicity of Amide Herbicides for Ecotoxicological Risk Assessment

Kexin Wang[1] · Yangzhou Lv[1] · Mei He[1,3] · Lei Tian[1,2] · Fan Nie[3] · Zhiguo Shao[3] · Zhansheng Wang[3]

Received: 9 October 2022 / Accepted: 5 January 2023 / Published online: 17 January 2023

Abstract

Amide herbicides have been widely applied in agriculture and found to be widespread and affect nontarget organisms in the environment. To better understand the biotoxicity mechanisms and determine the toxicity to the nontarget organisms for the hazard and risk assessment, five QSAR models were developed for the biotoxicity prediction of amide herbicides toward five aquatic and terrestrial organisms (including algae, *daphnia*, fish, earthworm and avian species), based on toxicity concentration and quantitative molecular descriptors. The results showed that the developed models complied with OECD principles for QSAR validation and presented excellent performances in predictive ability. In combination, the investigated QSAR relationship led to the toxicity mechanisms that eleven electrical descriptors (E_{HOMO}, E_{LUMO}, α_{xx}, α_{yy}, α_{zz}, μ, $q\mathrm{N}^-$, Q_{xx}, Q_{yy}, $q\mathrm{H}^+$, and q^-), four thermodynamic descriptors (Cv, S^θ, H^θ, and ZPVE), and one steric descriptor (V_m) were strongly associated with the biotoxicity of amide herbicides. Electrical descriptors showed the greatest impacts on the toxicity of amide herbicides, followed by thermodynamic and steric descriptors.

Introduction

Amide herbicides constitute the second largest proportion of the herbicides used in agriculture (Ding et al. 2011). Amide herbicides are a group of chemicals than can specifically interfere with biosynthesis of fatty acids, proteins and membrane, inhibit α- amylase and protease activities of germinating seeds, and suppress photosynthesis as inhibitors and uncoupling agents of electron chain transport in these plants (Qin et al. 2007; Robin et al. 2017). However, only a small amount of the applied herbicides reaches the target plants while an overwhelmingly larger portion is introduced into the environment (Cui et al. 2012) that adversely affect nontarget species, such as algae (Zhao et al. 2017), fish (Nassar et al. 2021), and earthworm (Li et al. 2021). These organisms occupy different trophic levels which constitute a significant part of the food web. Ecotoxicological risk assessment of such compounds is traditionally performed using standardized tests (e.g., according to OECD and ISO guidelines) which focus on biotoxicity toward some sensitive organisms. However, current methodologies for biotoxicity testing are expensive, time-consuming, laborious and poorly reproducible (Pavan and Worth 2008).

Quantitative Structure Activity relationship (QSAR) is an effective and low-cost alternative technology of biotoxicity testing that can accurately determine biotoxicity by developing mathematical models to establish the quantitative relationship between molecular structure and biotoxicity of chemicals (Kishor et al. 2019), which can also provide scientific insights into biotoxicity mechanisms of these chemicals (Pandey et al. 2020). Several studies have employed QSAR approach to investigate the quantitative structure activity relationship and predict the biotoxicity of pollutants such as natural medicine (Hamadache et al. 2016), herbicides (Gough and Hall 1999; Zakarya et al. 1996), food additive (Valerio et al. 2007), cosmetics (Hamadache et al. 2016), agriculture (Yang et al. 2020) and metal nanomaterials (Sizochenko and Leszczynski 2016). According to OECD Requirements and Guidelines (Netzeva et al. 2005), a valid

✉ Mei He
hemei-521@163.com

✉ Lei Tian
tianlei4665@163.com

[1] Hubei Key Laboratory of Petroleum Geochemistry and Environment (Yangtze University), Wuhan 430100, China

[2] School of Petroleum Engineering, Yangtze University, Wuhan 430100, China

[3] State Key Laboratory of Petroleum Pollution Control, CNPC Research Institute of Safety and Environmental Technology, Beijing 102200, China

and effective QSAR model conforms to the following characteristics: (1) a defined endpoint, (2) an unambiguous algorithm, (3) a defined domain of applicability, (4) appropriate measures of goodness-of-fit, robustness and predictability, (5) a mechanistic interpretation if possible. The QSAR models that complied with OECD principles for QSAR validation can be effectively applied into the ecotoxicological risk assessment of these compounds for management and pollution control.

In the present study, we assembled the toxicity concentrations of amide herbicides from Pubchem and Pesticide Properties Database. A series of electrical, thermodynamic, steric and hydrophobic molecular descriptors including E_{HOMO}, E_{LUMO}, α_{xx}, α_{yy}, α_{zz}, μ, qN^-, Q_{xx}, Q_{yy}, Q_{zz}, qH^+, q^-, V_m, E_{th}, C_v, S^θ, E_t, H^θ, G^θ, and ZPVE were calculated by density-functional theory calculation in EPIWEB4.1 and ORCA software. Five QSAR models toward algae, *daphnia*, fish, earthworm and avian species were respectively developed using a combination of Multiple Linear Regression (MLR) and Principal Component Analysis (PCA). The quantitative relationship between chemical structure and biotoxicity was then investigated. The main aim of this study were: (1) to develop valid and effective biotoxicity QSAR models of amide herbicides to different organisms; (2) to establish the quantitative relationship between structure and biotoxicity of amide herbicides; (3) to identify the effects of different molecular descriptors on the toxicity of amide herbicides. The results of this study provide an effective and low-cost measure for accurate biotoxicity determination of amide herbicides and give new insights that will helps to understand the biotoxicity mechanisms.

Materials and Methods

Biological Toxicity of Amide Herbicides

Amide herbicides are hazardous to the ecological environment. There is an extensive database on the effects of amide herbicides on ecosystems. Amide herbicides are thought to inhibit the growth, reproduction and development of many terrestrial and aquatic organisms (Coleman et al. 2000; Lunghini et al. 2020). In this study, acute toxicity results (50% effective concentration (EC_{50}), lethal concentration or dose (LC_{50} or LD_{50})) of amide herbicides on five organisms (algae, *daphnia*, fish, earthworm and avian species) were searched and collected as many as possible from online database or other references based on OECD guidelines. For instance, growth inhibition for algae, immobilization inhibition for *daphnia*, and mortalities and abnormalities appearance/behavior for fish, earthworm, and avian species. In this study, amide herbicides with higher toxicity that have been studied extensively in previous literatures were selected and collected as much as possible from PubChem and Pesticide Properties Database. Finally, the minimum value of toxicity data of 27 amide herbicides was collected. The name, chemical formula, CAS number and molecular weight of these amide herbicides are shown in Table 1.

Quantification of Molecular Structure

In order to characterize the molecular structure of these amide herbicides, twenty-one molecular descriptors were calculated to quantify these structures. The molecular descriptors involve one hydrophobic parameters (Octanol–Water Partition Coefficient, $logk_{ow}$), twelve electronic parameters (E_{HOMO}, E_{LUMO}, α_{xx}, α_{yy}, α_{zz}, qN^-, qH^+, q^-,μ, Q_{xx}, Q_{yy}, and Q_{zz}), seven thermodynamic parameters (E_{th}, E_t, C_v, S^θ, G^θ, H^θ, and ZPVE) and one steric parameters (V_m). The symbols and definitions of these molecular descriptors are shown as Table 2. The molecular descriptors were quantified as follows: Firstly, the molecular structure of each amide herbicide was initially constructed using ChemDraw 19.0 and then optimized by Chem3D software. Secondly, $LogK_{ow}$ of amide herbicides was calculated by EPIWEB4.1. Lastly, the rest of the molecular descriptors were calculated with the optimized structure at the B3LYP/6-311G^{++} (d, p) level using the density-functional theory (DFT) calculation by ORCA software, as described by Micera and Garribba (Giovanni and Eugenio 2011).

QSAR Model Development

MLR and PCA modeling methods were performed to develop QSAR models for biotoxicity prediction using SPSS26, Eviews7 and StataMP software, as described by Fadilah and Toropova (Fadilah et al. 2018; Toropova et al. 2015). MLR was employed to describe the quantitative linear correlations between the molecular descriptors and the biotoxicity. PCA was used to eliminate multicollinearity between the individual molecular descriptors during modeling. In this study, the collected toxicity data were categorized into five groups (algae, *daphnia*, fish, earthworm and avian species) and then the QSAR models were developed, respectively. The specific modeling steps were as follows:

Firstly, in the MLR analysis, Ordinary Least Squares Method (OLS) was applied to identify the most important descriptors contributing to the toxicity. F-tests and T tests were employed to eliminate insignificant descriptors in the OLS analysis. If F-tests and T tests can not pass, the regression equation recheck is necessary. According to the stepwise regression results, an initial QSAR model was developed. The statistical quality of QSAR model was evaluated by fitting coefficient (R^2) and root-mean-square error (RMSE). R^2 and RMSE are defined as Eqs. (1–2). Higher R^2 and lower RMSE

Table 1 The molecular structure information of amide herbicides

No.	Chemical name	Chemical formula	Molecular formula	CAS	Molecular weight
1	Acetamide	N,N-diallyl-2-chloroacetamide	$C_8H_{12}ClNO$	93-71-0	173.64
2	Amicarbazone	4-amino-N-tert-butyl-4,5-dihydro-3-isopropyl-5-oxo-1H-1,2,4-triazole-1-carboxamide	$C_{10}H_{19}N_5O_2$	129909-90-6	241.29
3	beflubutamid	(RS)-N-benzyl-2-[(α,α,α,4-tetrafluoro-m-tolyl)oxy]butyramide	$C_{18}H_{17}F_4NO_2$	113614-08-7	355.30
4	butanamide	(RS)-2-bromo-3,3-dimethyl-N- (1-methyl-1-phenylethyl)butyramide	$C_{15}H_{22}BrNO$	74712-19-9	312.24
5	cafenstrole	N,N-diethyl-3- (mesitylsulfonyl)-1H-1,2,4-triazole-1-carboxamide	$C_{16}H_{22}N_4O_3S$	125306-83-4	350.40
6	dimethenamid-P	(S)-2-chloro-N- (2,4-dimethyl-3-thienyl)-N- (2-methoxy-1-methylethyl)acetamide	$C_{12}H_{18}ClNO_2S$	163515-14-8	275.80
7	Diphenamid	N,N-dimethyl-2,2-diphenylacetamide	$C_{12}H_{18}ClNO_2S$	957-51-7	239.31
8	Fentrazamide	4- (2-chlorophenyl)-N-cyclohexyl-N-ethyl-4,5-dihydro-5-oxo-1H-tetrazole-1-carboxamide	$C_{16}H_{20}ClN_5O_2$	158237-07-1	349.81
9	Flucarbazone	4,5-dihydro-3-methoxy-4-methyl-5-oxo-N-{[2- (trifluoromethoxy)phenyl]sulfonyl}-1H-1,2,4-triazole-1-carboxamide	$C_{12}H_{11}F_3N_4O_6S$	145026-88-6	396.30
10	Flupoxam	1-[4-chloro-α- (2,2,3,3,3-pentafluoropropoxy)-m-tolyl]-5-phenyl-1H-1,2,4-triazole-3-carboxamide	$C_{19}H_{14}ClF_5N_4O_2$	119126-15-7	460.80
11	Fomesafen	5-[(2-chloro-α,α,α-trifluoro-p-tolyl)oxy]-N-mesyl-2-nitrobenzamide	$C_{15}H_{10}ClF_3N_2O_6S$	72178-02-0	438.80
12	Huangcaoling	N-mesyl-N-methyl-2-[(phosphonomethyl)amino]acetamide	$C_5H_{13}N_2O_6PS$	98565-18-5	190.00
13	Isoxaben	N-[3- (1-ethyl-1-methylpropyl)isoxazol-5-yl]-2,6-dimethoxybenzamide	$C_{18}H_{24}N_2O_4$	82558-50-7	332.40
14	Napropamide	(RS)-N,N-diethyl-2- (1-naphthyloxy)propionamide	$C_{17}H_{21}NO_2$	15299-99-7	271.35
15	Napropamide-M	(R)-N,N-diethyl-2- (1-naphthyloxy)propionamide	$C_{17}H_{21}NO_2$	41643-35-0	271.35
16	Naptalam	N-1-naphthylphthalamic acid	$C_{18}H_{13}NO_3$	132-66-1	291.30
17	Pethoxamid	2-chloro-N- (2-ethoxyethyl)-N- (2-methyl-1-phenylprop-1-enyl) acetamide	$C_{16}H_{22}ClNO_2$	106700-29-2	295.80
18	Propyzamide	3,5-dichloro-N- (1,1-dimethylprop-2-ynyl)benzamide	$C_{12}H_{11}Cl_2NO$	23950-58-5	256.12
19	Quinonamid	2,2-dichloro-N- (3-chloro-1,4-naphthoquinon-2-yl)acetamide	$C_{12}H_6Cl_3NO_3$	27541-88-4	318.50
20	Saflufenacil	N′-{2-chloro-5-[3,6-dihydro-3-methyl-2,6-dioxo-4- (trifluoromethyl)pyrimidin-1 (2H)-yl]-4-fluorobenzoyl}-N-isopropyl-N-methylsulfamide	$C_{17}H_{17}ClF_4N_4O_5S$	372137-35-4	500.90
21	Tebutam	N-benzyl-N-isopropylpivalamide	$C_{15}H_{23}NO$	35256-85-0	233.35
22	Alachlor	2-chloro-N- (2,6-diethylphenyl)-N- (methoxymethyl)acetamide	$C_{14}H_{20}ClNO_2$	15972-60-8	269.77
23	Acetochlor	2-chloro-N- (ethoxymethyl)-N- (2-ethyl-6-methylphenyl)acetamide	$C_{14}H_{20}ClNO_2$	194992-44-4	265.30
24	Metolachlor	2-chloro-N- (2-ethyl-6-methylphenyl)-N- (1-methoxypropan-2-yl)acetamide	$C_{15}H_{22}ClNO_2$	51218-45-2	283.79
25	Pretilachlor	2-chloro-N- (2,6-diethylphenyl)-N- (2-propoxyethyl)acetamide	$C_{17}H_{26}ClNO_2$	51218-49-6	311.80
26	Butachlor	N- (butoxymethyl)-2-chloro-N- (2,6-diethylphenyl)acetamide	$C_{17}H_{26}ClNO_2$	23184-66-9	311.80
27	Propanil	N- (3,4-dichlorophenyl)propanamide	$C_9H_9Cl_2NO$	709-98-8	218.08

are considered to be one of the necessary conditions for QSAR model to meet the OECD standard (Pandey et al. 2020).

$$R^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}} \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{2}$$

where SSR is sum of squares due to regression, SSE is sum of squares due to error, SST is sum of squares total, $\hat{y}_i$ is the predicted value of the test set, and y_i is the experimental value of the test set.

Table 2 Symbols and definition of the molecular descriptors

No	Type of descriptors	Descriptors	Definition
1	Hydrophobic descriptors	$\mathrm{LogK_{ow}}$	Water partition coefficient of octanol
1	Electronic descriptors	E_{HOMO}	Energy of the highest unoccupied molecular orbital
2		E_{LUMO}	Energy of the lowest unoccupied molecular orbital
3		$q\mathrm{N}^-$	Charge of the most negative nitrogen atom
4		$q\mathrm{H}^+$	Highest positive charge of hydrogen
5		q^-	Negative charge of the most negative atom
6		μ	Dipole moment
7		α_{xx}	XX of Exact polarizability
8		α_{yy}	YY of Exact polarizability
9		α_{zz}	ZZ of Exact polarizability
10		Q_{xx}	XX of quadrupole moment
11		Q_{yy}	YY of quadrupole moment
12		Q_{zz}	ZZ of quadrupole moment
1	Thermodynamic descriptors	E_{th}	Sum of electronic and thermal Energies
2		E_{t}	Total energy
3		C_{V}	Heat at constant volume
4		S^{θ}	Entropy
5		G^{θ}	Sum of electronic and thermal Free Energies
6		H^{θ}	Sum of electronic and thermal Enthalpies
7		ZPVE	Zero-point vibrational energy
1	Steric descriptors	V_{m}	Van der Waals volume of molecule

Then, PCA was performed on the significant molecular descriptors in the MLS analysis to extract principal components of variables and eliminate multicollinearity. In this work, variance inflation factors (VIF) are defined as Eq. (3), which are adopted to evaluate the collinearity of descriptors in the model. If VIF > 10, the regression equation is unstable and recheck is necessary. MLR analysis was then used again on the extracted principal components of descriptors in the PCA, and a new regression equation was built and a QSAR model with eliminated multicollinearity was developed.

$$\mathrm{VIF} = \frac{1}{1 - R^2} \tag{3}$$

where R^2 is fitting coefficient of the regression equation.

Lastly, double cross-validation (internal validation and external validation) was conducted on the developed QSAR models for reliable estimation of prediction errors. In this study, leave-one-out method was used for internal validation and external validation to evaluate the reliability and accuracy of the models, as described by Baumann and Baumann (Désirée and Knut 2014). Internal stability of the developed models was evaluated by leave-one-out cross-validation coefficient (Q^2_{LOO}). The models' performance in predictions was evaluated by external validation correlation coefficient (Q^2_{EXT}). Q^2_{LOO} and Q^2_{EXT} are defined as Eq. (4–5).

$$Q^2_{\mathrm{LOO}} = 1 - \frac{\mathrm{PRESS}}{\mathrm{TSS}} \tag{4}$$

$$Q^2_{\mathrm{EXT}} = 1 \frac{\sum_{i=1}^{n_{\mathrm{EXT}}} \left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n_{\mathrm{EXT}}} \left(\hat{y}_i - \bar{y}_{\mathrm{EXT}}\right)^2} \tag{5}$$

where PRESS is prediction error sum of squares, TSS is sum of squares of deviations of the experimental values, $\hat{y}_i$ is the predicted value of the test set, y_i is the experimental value of the test set,$\bar{y}_{\mathrm{EXT}}$ is the mean of the experimental values of the test set.

According to the procedures described above, an effective and accurate QSAR model for biotoxicity prediction was finally established. The biological toxicity (Y) is described with the best combination of the most relevant descriptors used as independent variables ($x_1, x_2 \ldots x_n$), as follows (6):

$$Y = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + a_0 \tag{6}$$

where a_0 is the intercept and $a_1, a_2 \ldots a_n$, are the regression coefficients.

Biotoxicity Prediction Accuracy Verification of the Developed QASR Models

The predicted biotoxicities by QASR models and the measured biotoxicities of four amide herbicides (benzadox,

cyprazole, epronaz, and coconut diethanol amide CDEA), which were randomly selected, were compared to verify the accuracy of the developed QSAR models for biotoxicity prediction. In order to predict the biotoxicities by the developed QSAR models, the four amide herbicides were performed to structural quantification as "*Quantification of molecular structure*". Then, the biotoxicities of these amide herbicides were calculated and predicted based the quantified molecular structures and the quantitative relationship between molecular structure and biotoxicity in the developed models. In order to measure the biotoxicities of the four amide herbicides on algae, *daphnia*, fish, earthworm and avian species, biotoxicity tests were respectively carried out according to the following OECD guidelines: freshwater Alga and Cyanobacteria, Growth Inhibition Test (OECD 201); *Daphnia* sp., Acute Immobilisation Test (OECD 202); Fish, Acute Toxicity Test (OECD 203); Earthworm, Acute Toxicity Tests (OECD 207); Acute Avian Oral Sequential Toxicity Test (OECD 223).

Results and Discussion

Toxicity of Amide Herbicides

As shown in Table 3, five groups (algae, *daphnia*, fish, earthworm, and avian species) of acute toxicity data were collected for QSAR model development. The toxicity for each organism species showed significant differences among these amide herbicides. The EC_{50}, LC_{50} and LD_{50} concentrations of these amide herbicides were 0.0036–149 mg/L, 0.058–500 mg/L, 0.36–170 mg/L, 0.515–1000 mg/Kg and 180–30000 mg/Kg for algae, *daphnia*, fish, earthworm and avian species, respectively. The toxicity of amide herbicides followed in the order of algae > *daphnia* > earthworm > fish > avian species, as reflected by the change ranges and fold changes in EC_{50}, LC_{50} and LD_{50} concentrations. The maximum and minimum toxicity was separately observed in algae and avian

Table 3 Toxicity of amide herbicides on algae, *daphnia*, fish, earthworm and avian species

Amide herbicides Compound	EC_{50} (mg/L)		LC_{50} (mg/L)	LC_{50} (mg/Kg soil)	LD_{50} (mg/Kg body weight)
	Algae	*Daphnia*	Fish	Earthworm	Avian species
Allidochlor	10	–	2	–	**180**
Amicarbazone	–	40.8	120	–	1965
Beflubutamid	69.2	1.64	1.86	366	2000
Bromobutide	–	–	10.0	–	–
Cafenstrole	–	**500**	1.2	–	2000
Dimethenamid-P	0.019	3.2	2.6	294.4	1068
Diphenamid	–	**0.058**	97	–	**30,000**
Fentrazamide	0.068	112	122	750	5000
Flucarbazone	–	–	–	484	1061
Flupoxam	0.17	330	**170**	**1000**	5000
Huangcaoling	1.4	1.3	0.87	500	2000
Isoxaben	3.4	14.3	6.6	282	2250
Napropamide	28.18	19	11.2	122.02	2000
Napropamide-M	–	118.5	76.1	–	4640
Naptalam	**149**	23	2.2	316	1578
Pethoxamid	2.8	5.6	4.7	173	6578
Propyzamide	–	–	0.45	–	11,136
Quinonamid	–	98.2	98.0	**1000**	2000
Saflufenacil	82	5.6	19	200	5000
Tebutam	0.966	10	1.8	386.8	1536
Alachlor	**0.0036**	8.3	**0.36**	105.5	928
Acetochlor	57.1	23.5	3.9	140	2000
Metolachlor	9.29	13	0.9	19.23	10,000
Pretilachlor	0.2	2.4	0.44	**0.515**	4640
Butachlor	0.11	2.39	5.4	734	196
Propanil	94	95	87	147	–

* The data in bold are the maximum and minimum values for each organism; The endpoints of the EC_{50}, LC_{50} and LD_{50} are growth inhibition for algae, immobilization inhibition for *daphnia*, mortalities and abnormalities appearance/behavior for fish, earthworm and avian species

species, with a EC_{50} or LD_{50} concentration of 0.0036 mg/L and 30,000 mg/kg. It can be seen from these results that a great diversity of amide herbicides with large differences in toxicity to a variety of organisms were included in the QSAR model development.

Structural Quantification Information of Amide Herbicides

The three-dimensional graphics of the involved amide herbicides are listed in Figure S1 of the supplementary material. In this study, molecular structure characteristics of amide herbicides were quantified and characterized by a series of molecular descriptors, including one hydrophobic parameter, twelve electronic parameters, seven thermodynamic parameters and one steric parameter. These molecular descriptors showed the structural information and properties of different aspects of amide herbicides. For example, E_{HOMO} is an electronic descriptor directly related to the ionization potential, which characterizes the susceptibility of the molecule toward attack by electrophiles (Sun et al. 2013). q is another electronic descriptor characterizing atomic charges, which are connected with the reactive centers activity of a chemical (Niu and Yu 2004). S^{θ} is a thermodynamic descriptor that is a measure of resistance to thermal disturbance within a compound (Zhu et al. 2010). C_v and H^{θ} are thermodynamic descriptors that reflect changes in heat and energy within a molecular system (Xi et al. 2006).

In this study, 27 amide herbicides were collected from pervious references in PubChem and Pesticide Properties Database. As reflected by the molecular descriptors in Table S1–S3 of the supplementary material, the involved amide herbicides showed a large difference in hydrophobic, electronic, thermodynamic parameters and steric properties, indicating the molecular structure differs significantly among these amide herbicides compounds. For hydrophobic descriptor, $LogK_{ow}$ varied from − 3.3600 to 6.5100, which indicated a great difference for the aqueous solubility and hydrophobicity of these amide herbicides. The variation in electronical descriptors among these amide herbicides ranged from 1.3-fold to 26.8-fold for E_{HOMO} and μ, respectively. For thermodynamic descriptors, E_{th}, E_t, C_V, S^{θ}, G^{θ} and H^{θ} were, respectively, − 3288.8321 to − 715.2634, 110.2390–268.5050, 44.2020–116.1480, 115.5790–211.3120 and − 3288.8311 to − 715.2624. The steric descriptor V_m varied extensively, ranged from − 101.8673 to 3306.8850. Maximum and minimum values for these molecular descriptors occurred frequently in allidochlor and saflufenacil. These results supported that amide herbicides with a wide variety of molecular properties were involved in this study.

QSAR Model Development and Accuracy in Toxicity Prediction

Table 4 provides the overall summary of the developed QSAR models for each set of organism groups. Based on the fitting coefficient (R^2), root-mean-square error (RMSE) in the MLR and PCA analysis and the double cross-validation coefficients (Q^2_{LOO} and Q^2_{EXT}) in the model validation procedures, these QSAR models showed good robustness and prediction ability in the toxicity evaluation and prediction of amide herbicides. Less than 5% deviation between the model predicted biotoxicity values and the measured biotoxicity results was identified as accurate QSAR models. The QSAR model for earthworm was better than the other four models for algae, avian species, *daphnia*, and fish, as indicated by higher R^2, Q^2_{LOO}, Q^2_{EXT}, and lower RMSE in QSAR model development.

QSAR Models for Aquatic Organisms

The developed QSAR models for three aquatic organisms (algae, *daphnia* and fish) are shown in Table 4. According to OECD Requirements and Guidelines, if $R^2 > 0.6$, $Q^2_{LOO} > 0.6$ and $Q^2_{EXT} > 0.5$, the developed models are available, and if $Q^2_{LOO} > 0.9$ and $Q^2_{EXT} > 0.9$, the models are identified as excellent. Additionally, the closer R^2 gets to 1, the better the fitting effects of the developed models. Our results supported that all the three models were stable ($0.8869 < R^2 < 0.9666$) and robust ($0.7201 < Q^2_{LOO} < 0.8634$) and showed good performance in the toxicity prediction of amide herbicides ($0.5612 < Q^2_{EXT} < 0.7676$).

For predictive ability tests of the developed QSAR models, the model predicted EC_{50} or LC_{50} concentrations of the chosen four amide herbicides (benzadox, cyprazole, epronaz, and CDEA) on the three aquatic organisms (algae, *daphnia* and fish) was compared with the measured EC_{50} or LC_{50} concentrations, as shown in Fig. 1. The results showed that the deviations of the model predicted EC_{50} or LC_{50} values and the measure EC_{50} or LC_{50} concentrations varied from 0.7 to 3.5, from 1.0 to 4.9, and from 2.6 to 4.6, respectively, for algae, *daphnia* and fish, all of which were less than 5% deviation. It was supported that the measured EC_{50} or LC_{50} concentrations and the model predicted EC_{50} or LC_{50} concentrations showed good agreements, indicating the accurate predictive ability of the developed QSAR models for aquatic organisms.

QSAR Models for Terrestrial Animals

As shown in Table 4, the QSAR models for two terrestrial animals (earthworm and avian species) were developed, conforming to OECD Requirements and Guidelines ($R^2 > 0.6$, $Q^2_{LOO} > 0.6$ and $Q^2_{EXT} > 0.5$) (Tropsha

Table 4 Summary of the best QSBR model for each of the selected organisms

Organism	Model equation	Model statistics	Molecular descriptors involved in the model
Algae	$\ln EC_{50\text{–}72\ h\ (Algae)} = -134.312432E_{HOMO} - 12.477484qN^- + 30.252008qH^+ - 0.01420973V_m - 0.12127a_{xx} + 0.360051a_{zz} + 0.157306Q_{xx} + 0.231579Q_{yy} - 0.015182H^\theta - 0.9079 43\ S^\theta + 1.31474C_v + 33.445788$	$R^2 = 0.9666$, $F = 51.72$, $p_F = 5.7 \times 10^{-5}$, $t = -8.69$–11.48, $p_t = 0$–0.016, RMSE = 0.4966, $Q^2_{LOO} = 0.8134$, $Q^2_{EXT} = 0.5612$	S^θ, C_v, a_{zz}, H^θ, V_m, Q_{yy}, Q_{xx}, qN^-, a_{xx}, E_{HOMO} and qH^+
Daphnia	$\ln EC_{50\text{-}48\ h\ (Daphnia)} = -223.021358E_{LUMO} + 3.189945q^- + 1.208121\mu - 0.050095S^\theta - 0.048182a_{xx} - 0.064467a_{yy} - 0.039137a_{zz} + 2.86888E^{-05}ZPVE + 5.169833$	$R^2 = 0.9350$, $F = 61.72$, $p_F = 2.51 \times 10^{-8}$, $t = -10.99$–18.90, $p_t = 0$–0.01, RMSE = 0.5610, $Q^2_{LOO} = 0.9800$, $Q^2_{EXT} = 0.7676$	E_{LUMO}, ZPVE, S^θ, μ, a_{xx}, a_{yy}, q^- and a_{zz}
Fish	$\ln LC_{50\text{–}96\ h\ (Fish)} = 0.957012\mu - 0.00217V_m + 0.032194a_{zz} - 9.13631E-06ZPVE + 0.46328\ C_v - 0.31657S^\theta + 15.81153$	$R^2 = 0.8869$, $F = 40.22$, $p_F = 3.7 \times 10^{-7}$ $t = -8.72$–13.16, $p_t = 0$–0.01, RMSE = 0.6460, $Q^2_{LOO} = 0.7201$, $Q^2_{EXT} = 0.6189$	C_v, S^θ, ZPVE, μ, a_{zz}, and V_m
Earthworm	$LC_{50\text{–}14\ day\ (Earthworm)} = -6743.61E_{LUMO} - 511.09qN^- + 817.95q^- + 69.94\mu - 10.31Q_{xx} - 3.61a_{yy} - 1.153a_{zz} - 0.78H^\theta - 0.29V_m + 17.52Q_{yy} + 874.77$	$R^2 = 0.9700$, $F = 68.10$, $p_F = 0.00$, $t = -15.08$–33.75, $p_t = 0$–0.01, RMSE = 0.4962, $Q^2_{LOO} = 0.9800$, $Q^2_{EXT} = 0.9800$	Q_{yy}, Q_{xx}, H^θ, ZPVE, q^-, qN^-, E_{LUMO}, a_{yy}, a_{zz}, V_m and μ
Avian species	$\ln LD_{50\ (Avian)} = 0.129670S^\theta - 0.004391H^\theta - 66.630943 E_{LUMO} - 0.000008ZPVE + 0.039585Q_{XX} - 1.203028q^- + 3.617467qN^- - 14.967525E_{HOMO} - 3.074131$	$R^2 = 0.8931$, $F = 27.35$, $p_F = 9.77 \times 10^{-7}$, $t = -6.19$–79.99, $p_t = 0$–0.001, RMSE = 0.4274, $Q^2_{LOO} = 0.6821$, $Q^2_{EXT} = 0.6342$	S^θ, H^θ, E_{LUMO}, ZPVE, Q_{xx}, q^-, qN^-, and E_{HOMO}

* R the correlation coefficient, R^2 the adjusted correlation coefficient, F the F-test value, p_F the P value, t the t test value, p_t the P_t value, RMSE the standard error, Q^2_{LOO} the internal validation of correlation coefficients, Q^2_{EXT} the external validation statistics

2010). Comparatively, the QSAR model for earthworm showed more excellent performance, with high stability ($R^2 = 0.9700$), robustness ($Q^2_{LOO} = 0.9800$), and external predictive ability ($Q^2_{EXT} = 0.9800$). The QSAR model for avian species was stable ($R^2 = 0.8931$), robust ($Q^2_{LOO} = 0.6821$) and showed good predictive ability in toxicity ($Q^2_{EXT} = 0.6342$).

For further accuracy validation in toxicity prediction of the developed models for terrestrial organisms, the model predicted LC_{50}/LD_{50} concentrations and the measured LC_{50}/LD_{50} concentrations were compared, as indicated in Fig. 1. Our results indicated that the difference between the model predicted EC_{50} or LC_{50} values and the measured EC_{50} or LC_{50} results of all the chosen four amide herbicides (benzadox, cyprazole, epronaz, and CDEA) were lower than 5%. The deviations of the model predicted EC_{50} or LC_{50} values varied from 2.6 to 4.3 and from 1.9 to 4.6, respectively, for earthworm and avian species. The good consistency between the measured LC_{50}/LD_{50} concentrations and model predicted LC_{50}/LD_{50} concentrations supported the excellent predictive potential of the developed two models for terrestrial organisms.

The Effects of the Molecular Descriptors on Biotoxicity

In this study, five validated mathematical QSAR models were established for the toxicity prediction of amide herbicides. The molecular descriptors involved in these models can provide explanations and mechanisms of the toxicity caused by amide herbicides. It is possible to gain some insights into the interrelation of molecular structure and toxicity of amide herbicides through these molecular descriptors, which could provide solid foundation for the toxicity prediction and risk assessment.

The Molecular Descriptors Involved in the Models

Our results showed that electrical, thermodynamic and steric descriptors were included in the developed QSAR models, which made statistically significant contributions to the toxicity of amide herbicides. From the QSAR models for aquatic organisms in Table 4, eleven molecular descriptors, including seven electrical parameters (a_{xx}, a_{zz}, Q_{xx}, Q_{yy}, qN^-, E_{HOMO}, and qH^+), three thermodynamic parameters (S^θ, C_v, and H^θ), and one steric parameter (V_m), were associated with algal toxicity of amide herbicides. Eight molecular descriptors, including six electrical parameters (μ, a_{xx}, a_{yy}, a_{zz}, E_{LUMO} and q^-) and two thermodynamic parameters (ZPVE and S^θ), were related to *daphnia* toxicity of amide herbicides. Six molecular descriptors, involving two electronic parameters (μ and a_{zz}), three thermodynamic parameters (C_v, S^θ, and ZPVE), and one steric parameter (V_m), were

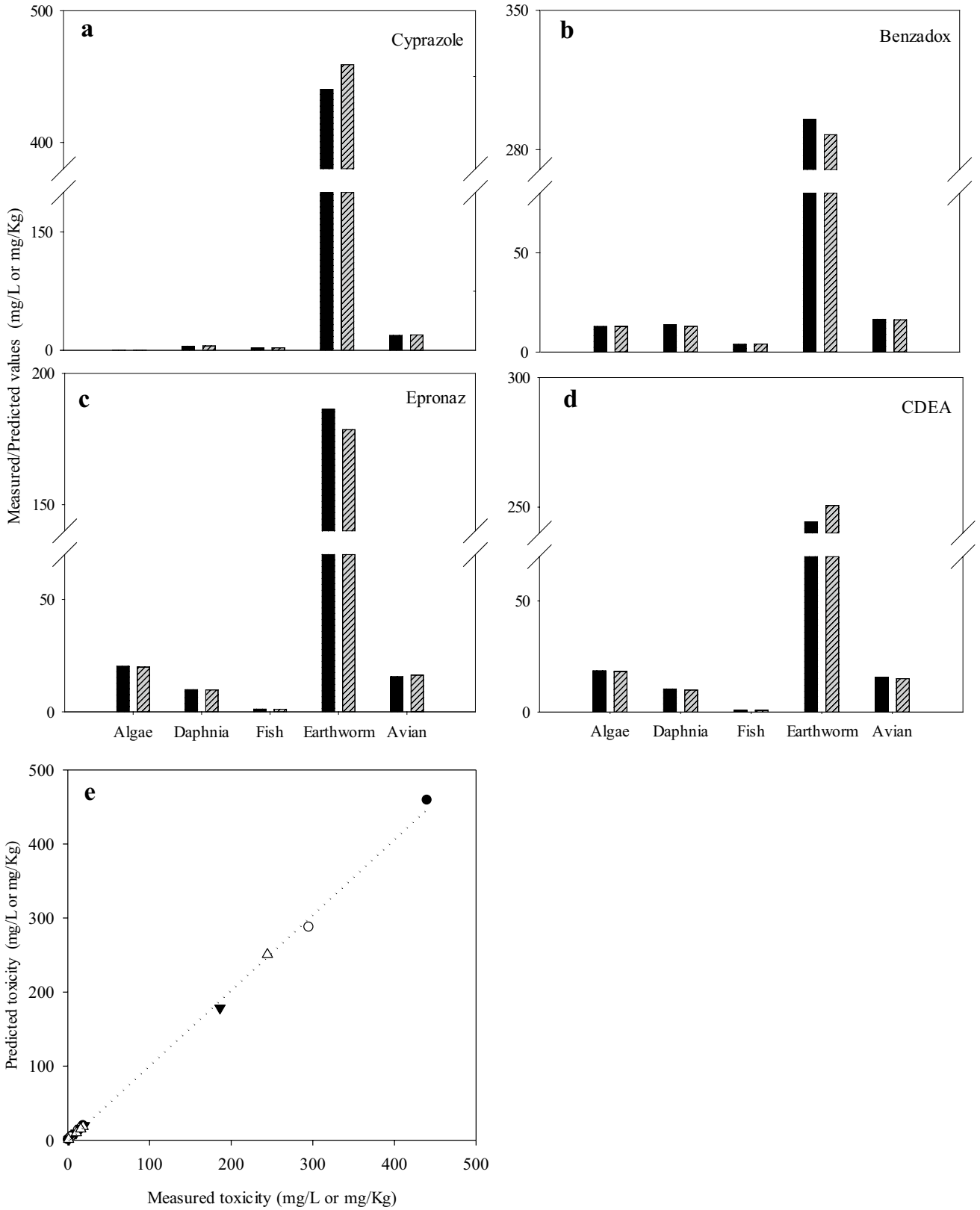

Fig. 1 The predictive performance of the developed QSAR models as reflected by the measured and the model predicted EC_{50} or LC_{50} values (mg/L or mg/Kg) of four amide herbicides on algae, *daphnia*, fish, earthworm and avian species. **a** Cyprazole, **b** benzadox, **c** epronaz, **d** CEDA, **e** the measured EC_{50} or LC_{50} values and the model predicted EC_{50} or LC_{50} values of four amide herbicides

connected with fish toxicity of amide herbicides. However, no influence of hydrophobic parameter was found over the aquatic toxicity of amide herbicides. The thermodynamic descriptor S^{θ} and the electrical descriptor a_{zz} were observed to be important molecular parameters affecting the toxicity of amide herbicides on all the three investigated aquatic organisms (*daphnia*, algae and fish), which respectively accounted for 11–37% and 5–16% of the weight in all the influencing molecular descriptors (Fig. 2).

There were differences in the molecular descriptors that associated to the toxicity of amide herbicides on different terrestrial animals, compared with aquatic organisms (Table 4). From the molecular descriptors involved in the QSAR models for terrestrial animals, ten molecular descriptors including seven electrical parameters (Q_{xx}, Q_{yy}, q^{-}, $q\mathrm{N}^{-}$, E_{LUMO}, a_{yy} and a_{zz}), two thermodynamic parameters (ZPVE, H^{θ}) and one steric parameter (V_{m}) were observed to be associated with the earthworm toxicity of amide herbicides. Eight molecular descriptors, including five electrical parameters (E_{LUMO}, q^{-}, $q\mathrm{N}^{-}$, Q_{xx} and E_{HOMO}) and three thermodynamic parameters (ZPVE, S^{θ} and H^{θ}), were relevant to the toxicity of amide herbicides on avian species. H^{θ}, ZPVE, Q_{xx}, $q\mathrm{N}^{-}$, q^{-} and $\mathrm{E}_{\mathrm{LUMO}}$ were observed to be important molecular descriptors affecting both the toxicity of amide herbicides on both earthworm and avian species. As shown in Fig. 2, the thermodynamic descriptors H^{θ} and ZPVE accounted for 14–19% and 12% of the weight for all the influencing molecular descriptors. ZPVE was also observed to be an important factor influencing the aquatic toxicity of amide herbicides, which accounted for 24% and 9% of the weight in all the influencing molecular descriptors, respectively, for *daphnia* and fish. Additionally, S^{θ} was found to affect the toxicity of terrestrial organisms (2% weight) as well as aquatic organisms.

The Underlying Mechanism of the Structure Descriptors Related to Biotoxicity

The molecular descriptors involved in the developed QSAR models demonstrate the mechanism underlying the toxicity of amide herbicides. Previous studies have reported that physicochemical, electrical, thermodynamic and steric properties are important factors influencing toxicity for many chemicals (such as phenols, organic phosphorus, benzenes, chlorophenols, PCBs (Duchowicz et al. 2008; Zvinavashe et al. 2009), chlorophenols, organophosphorus pesticide and aldehyde (Hadanu et al. 2015). In our study, the obtained results also indicated that the associated descriptors with toxicity of amide herbicides were related to the electrical, thermodynamic and steric properties. Electrical and thermodynamic properties had a larger impact on the toxicity of amide herbicides than steric properties.

Our results supported that the toxicity of amide herbicides was closely related to the molecular polarity of the herbicide molecules, as reflected by the sixteen electrical, thermodynamic and steric descriptors involved in the QSAR models. The electrical molecular descriptors (a_{xx}, a_{yy}, a_{zz}, $q\mathrm{N}^{-}$, q^{-}, and μ) and the thermodynamic molecular descriptors (C_{v} and S^{θ}) characterize the polarizability properties of molecule. a_{xx}, a_{yy}, and a_{zz} indicate the weight of polarizability in x, y, and z directions (Yang et al. 2021). $q\mathrm{N}^{-}$ represents the charge of the most electronegative atom in a compound (Qiu et al. 2013). q^{-} is related to atomic charge that influencing the binding of chemicals to the active site as well as the ability to form hydrogen bonds with biological receptors and therefore potentially affecting their toxicity. μ characterizes the average charge separation in a molecular system (Walker et al. 2002), which has negative contribution toward the toxicity as evidenced by the negative regression coefficient in the models of *daphnia*, fish and earthworm, which is consistent with previous reports. C_{v} indicates the constant capacity heat capacity of a chemical, which is directly proportional to the molecular polarity and the toxicity (Zuriaga et al. 2019). S^{θ} is the standard entropy representing the disorder degree of molecular reaction system of a compound which determines the difficulty of chemical reaction and thus has an impact on biological toxicity. Generally, the higher the S^{θ} value, the greater the biological toxicity (Ding et al. 2009). Previous studies have reported that chemicals might be more toxic with the increase of molecular polarity (Zhang et al. 2011). The polarizability represents the ability of a compound molecule to deform under an applied electric field and affects the interaction of electrons between the compound and atoms or molecules at the reaction site (Su et al. 2010). The higher the polarizability of the compound molecule, the deformation of the compound molecule in the corresponding direction of three-dimensional space will be enhanced, resulting in the increase in molecular polarity and the enhancement of toxicity (Wang et al. 2005).

The obtained results also indicated that the toxicity of amide herbicides was affected by the gain and loss of electrons. Energy of highest occupied molecular orbit (E_{HOMO}) and Energy of lowest unoccupied molecular orbit (E_{LUMO}) are electrical parameters describing energy of molecular orbital of chemicals, which respectively reflect the electron supply and loss capacity of a compound (Zhang et al. 2019). E_{HOMO} and E_{LUMO} relate with the sensitivity of compound molecules to external electrophilic/nucleophilic attack (Jiang et al. 2015). In this study, amide herbicides with larger energy of molecular orbital (higher E_{HOMO} and E_{LUMO}) produced higher toxicity to environmental organisms, which was consistent with previous discoveries (Clare 2004). The greater the E_{HOMO} values of one

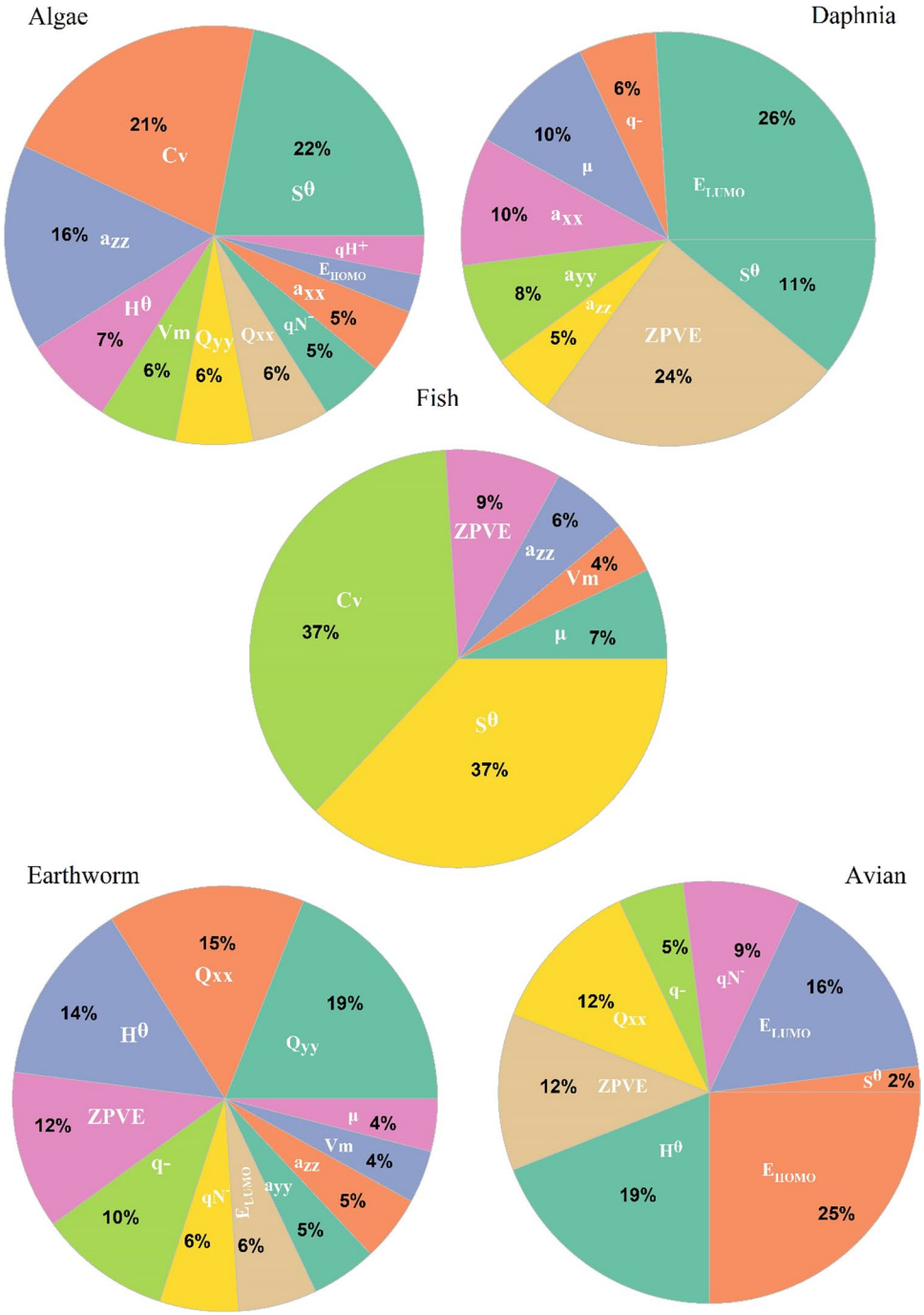

Fig. 2 The effects of molecular descriptors (involved in the developed QSAR models) on the toxicity of amide herbicides on algae, *daphnia*, fish, earthworm and avian species

chemical, the stronger its loss of electrons and reducing power and the greater its binding affinity for receptors and thus the lower its biological toxicities (Sun et al. 2013). The higher the E_{LUMO} value of one chemical, the stronger its gain of electrons and electrophilic ability and thus the greater its biological toxicity (Walker et al. 2002).

The results also showed that the toxicity of amide herbicides was associated with the electrostatic inductions across the herbicides molecules and the organisms. The electronic molecular descriptors (Q_{xx}, Q_{yy}, and $q\mathrm{H}^{+}$) relate to electrostatic inductions. Q_{xx} and Q_{yy} characterizes the non-spherical symmetry of three-dimensional charge distribution, which higher Q_{xx} and Q_{yy} are beneficial to form electrostatic inductions and thus increase the possibility of biological toxicity (Mhin et al. 2002). Our results indicated that Q_{xx} and Q_{yy} has a majority impact on the toxicity of amide herbicides on both aquatic and terrestrial organisms, which was consistent with previous reports (Zhang et al. 2008). $q\mathrm{H}^{+}$ represents the maximum charge of hydrogen atoms in a compound which could influence

the electrostatic attraction and thus affect the toxicity of chemicals (Niu and Yu 2004). Our results supported that the increase of qH^+ had a positive influence in decreasing the algal toxicity of amide herbicides.

Molecular volume was observed to be another important factor affecting the toxicity of amide herbicides in this study. V_m is Van der Waals volume that expresses the volume of per unit compound molecule, which may influence the biological toxicity by affecting other properties of a chemical, such as water solubility (Liu et al. 2015). It is reported that there is a toxicity threshold V_m in every compound system. The toxicity of a compound molecule enhances with the increase of its V_m but declines when exceeding the toxicity threshold V_m of this compound (Nohair et al. 2009). Our results showed that as V_m of amide herbicides increased, the EC_{50} value of amide herbicides decreased and the toxicity enhanced.

The water partition coefficient of octanol parameters ($LogK_{ow}$) characterizes the hydrophilic and hydrophobic properties of a chemical and is consistent with its lipid solubility (Wu et al. 2020). $LogK_{ow}$ is an important physicochemical parameter affecting the biological toxicity of a chemical, which has been mentioned in many chemicals such as polycyclic aromatic hydrocarbons (PAHs), nitrobenzene, dioxin and phenols (Bellifa and Mekelleche 2016; Ha et al. 2019). However, in this study, no significant effect was observed of $LogK_{ow}$ on toxicity of amide herbicides. Other influencing factors such as the actual effective concentration of amide herbicides on the target sites in the organisms may also affect the biotoxicity, which has been verified in previous study (Qiu et al. 2013).

Conclusions

The developed QSAR models showed excellent performance in predicting the biotoxicity of amide herbicides, supporting as an alternative approach to expensive laboratory toxicity tests. The QSAR relationship between electrical, thermodynamic, steric properties and toxicity can be easily interpreted with respect to potential mechanistic explanations of their effects on biotoxicity of amide herbicides. Strong association of electrical descriptors with the biotoxicity suggested electrical descriptors as the best predictive parameters for the biotoxicity prediction of amide herbicides.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00244-023-00980-9.

Acknowledgements This study was supported by the CNPC Research Institute Safety & Environmental Technology Programme (2021DJ6605), the National Natural Science Foundation of China (Grant No. 41472124), PetroChina Innovation Foundation (Grant No. 2015D-5006-0210, NO. 2016D-5007-0702), and the Yangtze Talents Fund [2020-2023].

Declarations

Conflict of interest The authors declare that we have no conflicts of interest. The manuscript is approved by all authors and has not been submitted to more than one journal for simultaneous consideration. The manuscript described has not been published before. The submitted work has not received any financial support from any third party, and there is no financial relationship with any entities. All of the financial organizations associated with this work have been disclosed. There is no patent, whether planned, pending or issued, broadly relevant to the submitted work.

References

Bellifa K, Mekelleche SM (2016) QSAR study of the toxicity of nitrobenzenes to Tetrahymena pyriformis using quantum chemical descriptors. Arab J Chem 9:S1683–S1689. https://doi.org/10.1016/j.arabjc.2012.04.031

Clare BW (2004) A novel quantum theoretic QSAR for hallucinogenic tryptamines: a major factor is the orientation of π orbital nodes. J Mol Struct Theochem 712(1–3):143–148. https://doi.org/10.1016/j.theochem.2004.08.050

Coleman S, Linderman R, Hodgson E, Rose RL (2000) Comparative metabolism of chloroacetamide herbicides and selected metabolites in human and rat liver microsomes. Environ Health Perspect 108(12):1151–1157. https://doi.org/10.1289/ehp.001081151

Cui ZL, Cui LX, Huang Y, Yan X, He J, Li SP (2012) Advances and application of microbial degradation in pesticides pollution remediation. J Nanjing Agric Univ 35(5):93-102. https://doi.org/10.7685/j.issn.1000-2030.2012.05.011

Désirée B, Knut B (2014) Reliable estimation of prediction errors for QSAR models under model uncertainty using double cross-validation. J Cheminformatics 6(1):47. https://doi.org/10.1186/s13321-014-0047-1

Ding GH, Li X, Zhang F, Chen JW, Huang LP, Qiao XL (2009) Mechanism-based quantitative structure-activity relationships on toxicity of selected herbicides to *Chlorella vulgaris* and *Raphidocelis subcapitata*. Bull Environ Contam Toxicol 83(4):520–524. https://doi.org/10.1007/s00128-009-9811-8

Ding L, Fu Y, Ye F (2011) Progress in research and application of amide herbicides. Pestic Sci Adm 32(9):22–26. https://doi.org/10.3969/j.issn.1002-5480.2011.09.011

Duchowicz PR, Mercader AG, Fernández FM, Castro EA (2008) Prediction of aqueous toxicity for heterogeneous phenol derivatives by QSAR. Chemometr Intell Lab Syst 90(2):97–107. https://doi.org/10.1016/j.chemolab.2007.08.006

Fadilah F, Arsianti A, Yanuar A, Andrajati R, Indah Paramita R, Hernawati Purwaningsih E (2018) Structure activity relationship analysis of antioxidant activity of simple benzene carboxylic acids group based on multiple linear regression. Orient J Chem 34(5):2656–2660. https://doi.org/10.13005/ojc/340558

Giovanni M, Eugenio G (2011) Is the spin-orbit coupling important in the prediction of the ^{51}V hyperfine coupling constants of $V^{(IV)}O^{2+}$ species? ORCA versus Gaussian performance and biological applications. J Comput Chem 32(13):2822–2835. https://doi.org/10.1002/jcc.21862

Gough JD, Hall LH (1999) Modeling the toxicity of amide herbicides using the electrotopological state. Environ Toxicol Chem 18(5):1069–1075. https://doi.org/10.1002/etc.5620180535

Ha H, Park K, Kang G, Lee S (2019) QSAR study using acute toxicity of *Daphnia magna* and *Hyalella azteca* through exposure to polycyclic aromatic hydrocarbons (PAHs). Ecotoxicology 28(3):333–342. https://doi.org/10.1007/s10646-019-02025-1

Hadanu R, Idris S, Sutapa I (2015) QSAR analysis of benzothiazole derivatives of antimalarial compounds based on AM1 semi-empirical method. Indones J Chem 15(1):86–92. https://doi.org/10.22146/ijc.21228

Hamadache M, Benkortbi O, Hanini S, Amrane A, Khaouane L, Si Moussa C (2016) A quantitative structure activity relationship for acute oral toxicity of pesticides on rats: validation, domain of application and prediction. J Hazard Mater 303:28–40. https://doi.org/10.1016/j.jhazmat.2015.09.021

Jiang L, Wen JY, Zeng YL, Li Y (2015) Investigation on aryl hydrocarbon receptor binding affinity QSAR model of polybrominated diphenyl ethers based on substituent descriptors/quantum chemical parameters. Asian Chem 27(2):575–581. https://doi.org/10.14233/ajchem.2015.17042

Kishor A, David W, Jan D, Maciej B, Paola M, Wojciech M, Oladapo K, Tomasz P, Russell JD (2019) A quantitative structure-biodegradation relationship (QSBR) approach to predict biodegradation rates of aromatic chemicals. Water Res 157:181–190. https://doi.org/10.1016/j.watres.2019.03.086

Li MY, Ma XX, Wang YR, Saleem M, Yang Y, Zhang QM (2021) Ecotoxicity of herbicide carfentrazone-ethyl towards earthworm *Eisenia fetida* in soil. Comp Biochem Physiol C Toxicol Pharmacol 253:109250–109250. https://doi.org/10.1016/J.CBPC.2021.109250

Liu HC, Lu S, Ran T, Zhang YM, Xu JX, Xiong X, Xu AY, Lu T, Chen YD (2015) Accurate activity predictions of B-Raf Type II Inhibitors via molecular docking and QSAR methods. Acta Phys-Chim Sin 31(11):2191–2206. https://doi.org/10.3866/pku.Whxb201510134

Lunghini F, Marcou G, Azam P, Enrici MH, Van Miert E, Varnek A (2020) Consensus QSAR models estimating acute toxicity to aquatic organisms from different trophic levels: algae, *Daphnia* and fish. SAR QSAR Environ Res 31(9):655–675. https://doi.org/10.1080/1062936X.2020.1797872

Mhin BJ, Lee JE, Choi W (2002) Understanding the congener-specific toxicity in polychlorinated dibenzo-p-dioxins: chlorination pattern and molecular quadrupole moment. J Am Chem Soc 124(1):144–148. https://doi.org/10.1021/ja016913q

Nassar AMK, AbdelHalim KY, Abbassy MA (2021) Mitochondrial biochemical and histopathological defects induced by the herbicide pendimethalin in tilapia fish (*Oreochromis niloticus*) comp biochem physiol. C Toxicol Pharmacol 242:108949. https://doi.org/10.1016/j.cbpc.2020.108949

Netzeva TI, Worth AP, Aldenberg T, Benigni R, Cronin MTD, Gramatica P, Jaworska JS, Kahn S, Klopman G, Marchant CA, Myatt G, Nikolova-Jeliazkova N, Patlewicz GY, Perkins R, Roberts DW, Schultz TW, Stanton DT, Van de Sandt JJM, Tong W, Veith G, Yang C (2005) Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships: the report and recommendations of ECVAM workshop 52. Altern Lab Anim 33(2):155–173. https://doi.org/10.1177/026119290503300209

Niu JF, Yu G (2004) Molecular structural characteristics governing biocatalytic oxidation of PAHs with hemoglobin. Environ Toxicol Pharmacol 18(1):39–45. https://doi.org/10.1016/j.etap.2004.05.002

Nohair M, Mallouk N, Benmarzouk M, Mohssine EM (2009) Statistical approaches to estimating the relative contribution of intermolecular interactions in aliphatic alcohols: application to QSPR/QSAR modeling of their boiling points. Chem Prod Process Model 4(1):1–22. https://doi.org/10.2202/1934-2659.1274

Pandey SK, Ojha PK, Roy K (2020) Exploring QSAR models for assessment of acute fish toxicity of environmental transformation products of pesticides (ETPPs). Chemosphere 252:126508. https://doi.org/10.1016/j.chemosphere.2020.126508

Pavan M, Worth AP (2008) Review of estimation models for biodegradation. QSAR Comb Sci 27(1):32–40. https://doi.org/10.1002/qsar.200710117

Qin YM, Hu CY, Pang Y, Kastaniotis AJ, Hiltunen JK, Zhu YX (2007) Saturated very-long-chain fatty acids promote cotton fiber and *Arabidopsis* cell elongation by activating ethylene biosynthesis. Plant Cell 19(11):3692–3704. https://doi.org/10.1105/tpc.107.054437

Qiu J, Dai Y, Zhang XS, Chen GS (2013) QSAR modeling of toxicity of acyclic quaternary ammonium compounds on *Scenedesmus quadricauda* using 2D and 3D descriptors. Bull Eviron Contam Toxicol 91(1):83–88. https://doi.org/10.1007/s00128-013-1006-7

Robin M, George R, Gilles-Eric S, Malcolm W, Michael A (2017) Multiomics reveal non-alcoholic fatty liver disease in rats following chronic exposure to an ultra-low dose of Roundup herbicide. Sci Rep 7(1):39328. https://doi.org/10.1038/srep39328

Sizochenko N, Leszczynski J (2016) Review of current and emerging approaches for quantitative nanostructure-activity relationship modeling: the case of inorganic nanoparticles. J Nanotoxicol Nanomed 1(1):1–16. https://doi.org/10.4018/JNN.2016010101

Su LM, Zhao YH, Yuan X, Mu CF, Wang N, Yan JC (2010) Evaluation of combined toxicity of phenols and lead to *Photobacterium phosphoreum* and quantitative structure-activity relationships. Bull Environ Contam Toxicol 84(3):311–314. https://doi.org/10.1007/s00128-009-9665-0

Sun P, Gao SM, Liu H, Chen JT (2013) QSAR analyzes for the predictive toxicity of substituted phenols and anilines to fish. Appl Mech Mater 295–298:109–112. https://doi.org/10.4028/www.scientific.net/AMM.295-298.109

Toropova MA, Veselinović AM, Veselinović JB, Stojanović DB, Toropov AA (2015) QSAR modeling of the antimicrobial activity of peptides as a mathematical function of a sequence of amino acids. Comput Biol Chem 59:126–130. https://doi.org/10.1016/j.compbiolchem.2015.09.009

Tropsha A (2010) Best practices for QSAR Model development, validation, and exploitation. Mol Inform 29(6–7):476–488. https://doi.org/10.1002/minf.201000061

Valerio LGJ, Arvidson KB, Chanderbhan RF, Contrera JF (2007) Prediction of rodent carcinogenic potential of naturally occurring chemicals in the human diet using high-throughput QSAR predictive modeling. Toxicol Appl Pharmacol 222(1):1–16. https://doi.org/10.1016/j.taap.2007.03.012

Walker JD, Carlsen L, Hulzebos E, Simon-Hettich B (2002) Global Government applications of analogues, SAR s and QSAR s to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. SAR QSAR Environ Res 13(6):607–616. https://doi.org/10.1080/1062936021000020062

Wang ZY, Han XY, Wang LS (2005) Quantitative correlation of chromatographic retention and acute toxicity for Alkyl (1-phenylsulfonyl) cycloalkane carboxylates and their structural parameters by DFT. Chinese J Struct Chem 24(7):851–857. https://doi.org/10.14102/j.cnki.0254-5861.2005.07.023

Wu SQ, Wang L, Xia ZH (2020) QSAR modelling for predicting comprehensive toxicity of aromatic substances to anaerobic microflora in petrochemical wastewater. Asian J Ecotoxicol 15(6):167–174. https://doi.org/10.7524/AJE.1673-5897.20191216001

Xi Z, Yu ZH, Niu CW, Ban SR, Yang GY (2006) Development of a general quantum-chemical descriptor for steric effects: density functional theory based QSAR study of herbicidal sulfonylurea analogues. J Comput Chem 27(13):1571–1576. https://doi.org/10.1002/jcc.20464

Yang L, Wang YH, Hao WY, Chang J, Pan YF , Li JZ, Wang HL (2020) Modeling pesticides toxicity to sheepshead minnow using QSAR. Ecotoxicol Environ Safe 193:110352. https://doi.org/10.1016/j.ecoenv.2020.110352

Yang L, Sang CH, Wang YH, Liu WT, Hao WY, Chang J, Li JZ (2021) Development of QSAR models for evaluating pesticide toxicity against *Skeletonema costatum*. Chemosphere 285:131456. https://doi.org/10.1016/j.chemosphere.2021.131456

Zakarya D, Larfaoui EM, Boulaamail A, Lakhlifi T (1996) Analysis of structure-toxicity relationships for a series of amide herbicides using statistical methods and neural network. SAR QSAR Environ Res 5(4):269–279. https://doi.org/10.1080/10629369608031716

Zhang HJ, Zhang JY, Zhu YM (2008) In vitro investigations for the QSAR mechanism of lymphocytes apoptosis induced by substituted aromatic toxicants. Fish Shellfish Immunol 25(6):710–717. https://doi.org/10.1016/j.fsi.2008.02.008

Zhang X, Xu J, He J (2011) Assessing non-inferiority with time-to-event data via the method of non-parametric covariance. Stat Methods Med Res 22(3):346–360. https://doi.org/10.1177/0962280211402261

Zhang SS, Li TT, Wang J, Hu YJ, Zhang HX, Zhao SX, Zhao YH, Li C (2019) QSAR models for predicting the aqueous reaction rate constants of aromatic compounds with hydrated electrons. Environ Chem 38(5):1005–1013. https://doi.org/10.7524/j.issn.0254-6108.2018062001

Zhao FF, Xiang QQ, Zhou Y, Xu X, Qiu XY, Yu Y, Ahmadd F (2017) Evaluation of the toxicity of herbicide topramezone to *Chlorella vulgaris*: oxidative stress, cell morphology and photosynthetic activity. Ecotoxicol Environ Saf 143:129–135. https://doi.org/10.1016/j.ecoenv.2017.05.022

Zhu M, Fei G, Zhu R, Wang X, Zheng X (2010) A DFT-based QSAR study of the toxicity of quaternary ammonium compounds on *Chlorella vulgaris*. Chemosphere 80(1):46–52. https://doi.org/10.1016/j.chemosphere.2010.03.044

Zuriaga E, Giner B, Valero MS, Gomez M, Garcia CB, Lomba L (2019) QSAR modelling for predicting the toxic effects of traditional and derived biomass solvents on a *Danio rerio* biomodel. Chemosphere 227:480–488. https://doi.org/10.1016/j.chemosphere.2019.04.054

Zvinavashe E, Du TT, Griff T, Van den Berg HHJ, Soffers AEMF, Vervoort J, Murk AJ, Rietjens IMCM (2009) Quantitative structure-activity relationship modeling of the toxicity of organothiophosphate pesticides to Daphnia magna and Cyprinus carpio. Chemosphere 75(11):1531–1538. https://doi.org/10.1016/j.chemosphere.2009.01.081