Chemosphere 80 (2010) 665–670

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

QSPR/QSAR models for prediction of the physico-chemical properties and biological activity of polychlorinated diphenyl ethers (PCDEs)

Xu Hui-Ying [a,*], Zou Jian-Wei [b], Hu Gui-Xiang [b], Wang Wei [c]

[a] College of Biology & Environment Engineering, Zhejiang Shuren University, Hangzhou 310015, China
[b] Key Laboratory for Molecular Design and Nutrition Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315104, China
[c] Zhenhai Environmental Monitoring Station, Ningbo 315200, China

ARTICLE INFO

Article history:
Received 14 November 2009
Received in revised form 31 March 2010
Accepted 22 April 2010
Available online 21 May 2010

Keywords:
Polychlorinated diphenyl ethers
Molecular electrostatic potential
QSRR
QSPR
QSAR

ABSTRACT

Polychlorinated diphenyl ethers (PCDEs) are a group of important persistent organic pollutants. In the present study, geometrical optimization and electrostatic potential calculations have been performed for all 209 PCDE congeners at the HF/6-31G* level of theory. A number of statistically-based parameters have been obtained. Linear relationships between gas-chromatographic relative retention time (*RRT*), *n*-octanol/water partition coefficient (log K_{OW}), 298 K supercooled liquid vapour pressures (log p_L), aqueous solubilities (log $S_{w,L}$) and the immunotoxicity values (log ED_{50}) of PCDEs and the structural descriptors have been established by multiple linear regression method. The result shows that the quantities derived from electrostatic potential $V_{s,min}$, $\sum V_s^+$, $V_{s,av}^-$, Π, σ_{tot}^2, σ_+^2, ν, and N_v^+, together with the number of the chlorine atoms on the two phenyl rings (N_{Cl}) can be well used to express the quantitative structure–property (activity) relationships of PCDEs. Good predictive capabilities have also been demonstrated by leave-group(1/5)-out cross-validation and external test set. Based on these equations, the predicted values have been presented for those PCDE congeners whose experimentally determined physico-chemical properties are unavailable.

1. Introduction

Polychlorinated diphenyl ethers (PCDEs) are a group of halogenated aromatic compounds, which are structurally located between polychlorinated biphenyls (PCBs) and polychlorinated dibenzofuran (PCDFs). PCDEs are mainly arise as by-products of technical chlorophenols and chlorinated phenoxyacetic acids, incomplete combustion e.g. emissions from municipal waste incinerators (Nevalainen et al., 1994). However, their ubiquitous environmental appearance is basically the result of their presence as impurities in chlorophenol preparations, where they have been identified at levels of 100–1000 mg/kg (Kurz and Ballschmiter, 1995; de Boer and Denneman, 1998). They have been frequently detected in a wide range of environmental samples including sediments (Koistinen et al., 1995a), fishes (Koistinen et al., 1993), white-tailed sea eagle (Koistinen et al., 1995b), foodstuffs (vegetables, tubers, pulses, cereals, fruits, fish and shellfish, meat and meat products, eggs, milk and dairy products, and fats and oils) (Bocio et al., 2004) and human adipose tissue (Stanley et al., 1991; Williams et al., 1991). The studies have shown that PCDEs have similar toxic properties to PCBs (Becker et al., 1991), and therefore they are also regarded as a type of persistent indicator molecules other than PCBs for a global environmental pollution by organochlorine compounds (Kurz and Ballschmiter, 1999). Wide distribution, high lipophilicity, and persistence of PCDEs have raised concern about their bioaccumulation, their potential biomagnification in the food webs, and their adverse effects (Domingo, 2006). Hence PCDEs have attracted great attention recently as an important type of environmental pollutants.

Physico-chemical properties of an organic chemical compound play an important role in determining its distribution and fate in the global environment. There are 209 PCDE congeners depending on the number and positions of the chlorine atoms on the two phenyl rings. Due to not only the time consumption and high expense, but also the unavailability of chemical standards of many PCDEs, it is very difficult to determine experimentally the physico-chemical properties and biological activity for all the PCDE congeners. Therefore, alternative approaches are needed. Many previous studies showed that it was indeed feasible to predict the properties or activities with quantitative structure–property (activity) relationship (QSPR/QSAR) models for many organic compounds (Tuppurainen and Ruuskanen, 2000; Zou et al., 2005; Ghasemi and Saaidpour, 2007; Toropov et al., 2008; Puzyn et al., 2009). In fact, QSPR/QSAR studies with respect to PCDEs can also be found sporadically in recent publications (Nevalainen et al., 1994; öberg, 2002; Yang et al., 2003; Huang et al., 2004; Chen et al., 2007; Sun et al., 2007; Zeng et al., 2007), although most of

* Corresponding author. Tel.: +86 571 88297172.
E-mail address: xuhy65@163.com (X. Hui-Ying).

0045-6535/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chemosphere.2010.04.050

these studies aimed at only one or few properties (usually chromatographic relative retention time).

The choice of appropriate structural parameters plays a crucial role in the QSPR and QSAR studies. It has been known that most of the physico-chemical properties (also including chromatographic ones) and biological activities are associated with intermolecular non-covalent interactions, which are mostly electrostatic in nature, and the electrostatic potentials, especially those distributed on the molecular surface, can be well used to quantify the molecular interactions. The structural descriptors derived from the molecular surface electrostatic potentials not only have definite physical meaning but also represent good repeatability despite being statistically based (i.e. they are kept constant so long as the three-dimensional molecular structure provided remains the same). In fact, this type of descriptors has been successfully applied in research of the quantitative structure–property (activity) relationships, and their application range is being increasingly extended (Murray et al., 1994, 1999; Betsy et al., 2002; Zou et al., 2005; Xu et al., 2008a,b). In this paper, we report five quantitative structure–property (activity) relationships of PCDEs established by using the structural descriptors derived from the molecular surface electrostatic potentials. It serves on one hand to improve our understanding how the physico-chemical properties and biological activity change with the numbers and position of substituted Cl, and to prove further that the parameter set derived from molecular electrostatic potential has more general applicability for predicting physico-chemical properties or biological activities of the persistent organic pollutants on the other hand.

2. Methods

The initial geometries of each compound were optimized with MOPAC 6.0 program implemented in VEGA package using AM1 method (keywords: "PRECISE", "GEO-OK") (Pedretti et al., 2002). Then the molecular geometries were reoptimized at the HF/6-31G* level with Gaussian98 software package (Frisch et al., 1998). Based on these optimized geometries, calculations of electronic density and electrostatic potentials with grid method were performed. The grid control option was set to "cube = 100". As a result, for each molecule, there are about 100^3 points at which the values of electronic density and electrostatic potentials were computed. On the basis of the above calculations, structural descriptors for all 209 PCDEs along with diphenyl ether were extracted and those pertinent to the subsequent discussion are described as follows:

(a) $V_{s,min}$, the minima of electrostatic potentials on the molecular surface (an outer contour of the electronic density $\rho(r) = 0.001$ a.u.).

(b) $\sum V_s^+$, defined as the sum of the surface maxima values of the electrostatic potential. The sum started from the most positive potential on the molecular surface, and if two $V_{s,max}$ appeared within 2.1 Å of each other, only the maxima with the most positive potential was included in the sum.

(c) For comparison, $\sum V_s^-$, defined as the sum of the surface minima values of the electrostatic potential. $V_{s,av}^-$ is its average.

(d) Π, a measure of charge separation or local polarity with the following definition (Murray et al., 1994).

$$\Pi = \frac{1}{n}\sum_{i=1}^{n}|V(r_i) - \overline{V}_s| \quad \text{(i)}$$

where $V(r_i)$ is the value of the electrostatic potential at a point r_i on the molecular surface and $\overline{V}_s$ is the average of the potential on the surface.

(e) σ_+^2 and σ_{tot}^2 are the positive and total variances of $V(r_i)$, which reflect the dispersion tendency of electrostatic potential (Murray et al., 1994).

$$\sigma_+^2 = \frac{1}{m}\sum_{i=1}^{m}\left|V^+(r_i) - \overline{V}_s^+\right|^2 \quad \text{(ii)}$$

$$\sigma_{tot}^2 = \sigma_+^2 + \sigma_-^2 = \frac{1}{m}\sum_{i=1}^{m}\left|V^+(r_i) - \overline{V}_s^+\right|^2 + \frac{1}{n}\sum_{i=1}^{n}\left|V^-(r_i) - \overline{V}_s^-\right|^2 \quad \text{(iii)}$$

where $V^+(r_i)$ and $V^-(r_i)$ are the positive and negative values, respectively, of the electrostatic potential at a point r_i on the molecular surface, and $\overline{V}_s^+$ and $\overline{V}_s^-$ are their averages.

(f) The ν term is a parameter of electrostatic equilibrium on molecular surface, which reaches a maximum value of 0.25 when σ_+^2 and σ_-^2 are equal. ν was defined as (Murray et al., 1994).

$$\nu = \frac{\sigma_+^2 \cdot \sigma_-^2}{(\sigma_{tot}^2)^2} \quad \text{(iv)}$$

(g) N_v^+, the number of independent points of the positive electrostatic potentials on molecular surface.

Furthermore, because the physico-chemical properties have close relation to the molecular size, molecular surface area (A_s) and molecular volume (V_{mc}) were calculated with TSAR software package based on these molecular geometries reoptimized at the HF/6-31G* level. The number of the chlorine atoms on the two phenyl rings (N_{Cl}) was also considered as a structural parameter reflecting molecular size.

Finally, correlation between the structural descriptors of PCDEs and the physico-chemical properties and biological activity was established by stepwise linear regression analysis with SPSS 10.0 package. In the course of stepwise regression analysis, the probability values of F were set as $\leqslant 0.050$ and $\geqslant 0.100$, respectively, for entering and removing a variable. The collinearity diagnostics procedures between variables were performed so that only one descriptor was retained from a pair with similar contribution. The predictive power of the QSPR/QSAR model was validated by a leave-1/5-out cross-validated analysis. Furthermore, predictions for external test set were also made for K_{OW}, p_L and $S_{w,L}$.

3. Results and discussion

All experimentally determined physico-chemical properties and biological activities of PCDEs were taken from previous publications, including the relative retention time (*RRT*) (Nevalainen et al., 1994), *n*-octanol/water partition coefficient ($\log K_{OW}$), 298 K supercooled liquid vapour pressures ($\log p_L$), aqueous solubilities ($\log S_{w,L}$) (Kurz and Ballschmiter, 1999) and the immunotoxicity values ($\log ED_{50}$) (Nevalainen and Kolehmainen, 1994) of PCDEs. These data, together with the chemical names and the predicted properties from QSPR/QSAR models for all 209 PCDE congeners along with diphenyl ether, are collected in Table S1 of the Supplementary materials.

Table S2 (Supplementary materials) lists the computed structural descriptors for all 209 PCDE congeners and diphenyl ether. As can be seen from Table S2, some electrostatic potential derived quantities, e.g. $V_{s,min}, \Pi, \sigma_{tot}^2, V_{s,av}^-$, vary significantly with the ring substitution pattern. While the number of chlorine atoms lead to elusive variation of these quantities, their magnitudes seems sensitive to the position of substituted chlorine atoms. Especially, for PCDE congeners with some molecular formula, the isomer with *ortho*-chlorine substitution always have smaller $V_{s,min}$ and $V_{s,av}^-$