Available online at www.sciencedirect.com

ScienceDirect

Progress in Natural Science 18 (2008) 867–872

Progress in Natural Science

www.elsevier.com/locate/pnsc

Prediction of soot–water partition coefficients for selected persistent organic pollutants from theoretical molecular descriptors

Qing Zhang, Jun Huang, Gang Yu *

POPs Research Centre, Department of Environmental Science and Engineering, Tsinghua University, Beijing 100084, China

Received 11 December 2007; received in revised form 2 February 2008; accepted 2 February 2008

Abstract

Quantitative structure–property relationship (QSPR) models were developed for soot–water partition coefficient (K_{SC}) values of selected persistent organic pollutants (POPs), i.e. 10 polychlorinated dibenzo-*p*-dioxins and dibenzofurans, nine polychlorinated biphenyls, four polycyclic aromatic hydrocarbons and two polybrominated diphenyl ethers, using partial least squares (PLS) regression. Quantum chemical descriptors computed by parameterized model revision 3 Hamiltonian method were used as predictor variables. The cross-validated Q^2_{cum} value for the optimal QSPR model is 0.844, indicating a good predictive capability for the $\log K_{SC}$ values of these chemicals. The QSPR results showed that average molecular polarizability (α), standard heat of formation (ΔH_f) and energy of the lowest unoccupied molecular orbital (E_{LUMO}) have dominant effects on K_{SC} of POPs. The results suggested that $\log K_{SC}$ values of POPs increase with the increase of α. Contrarily, $\log K_{SC}$ values decrease with the increase of E_{LUMO} and ΔH_f of POPs.

Keywords: Quantitative structure–property relationships; Black carbon; Persistent organic pollutants; Quantum chemical descriptors; Partial least squares

---

1. Introduction

Persistent organic pollutants (POPs) are ubiquitous hydrophobic organic compounds (HOCs) and some congeners (such as 2,3,7,8-polychlorinated dibenzo-*p*-dioxins and dibenzofurans) are strongly toxic. In recent decades, POPs are considered as typical endocrine disrupting chemicals with great potential risk to human health [1,2]. Human exposure to POPs has drawn great attention because of their wide occurrence and their adverse impacts on ecosystem and human health [3,4]. Recently, extensive research has been dedicated to evaluate the environmental behaviors of POPs for the purpose of their environmental risk assessment [5–7].

Sorption of POPs to soils and sediments is a key process controlling the environmental fate of such compounds [7,8]. For instance, sorption can limit both biological uptake and microbial degradation of POPs. Hence, a quantitative understanding of POP sorption is a necessary prerequisite for estimating various hazards that organic pollutants may pose in the environment. Generally, total particulate organic matter was considered as experimental sorbents to investigate the sorption behaviors of HOCs in the environment [9]. However, recent reports indicated that soot carbon (or black carbon) produced mainly from biomass burning and fossil fuel combustion is significantly better sorbent than total particulate organic matter for HOCs [5,10–13]. The enhanced sorption capacity may be attributed to the differences in physicochemical properties between soot and particulate organic matter in terms of surface area, elemental composition, functional groups, and so on. Furthermore, the experimental results in the field showed that sorption behaviors for POPs cannot be explained by organic matter absorption alone. Recent soot sorption experiments revealed that the environmental fate of POPs may also be affected by interaction with soot carbon [14]. Investigations on interaction between soot carbon

---

* Corresponding author. Tel.: +86 10 6278 7137; fax: +86 10 6279 4006.
E-mail address: yg-den@tsinghua.edu.cn (G. Yu).

1002-0071/$ - see front matter
doi:10.1016/j.pnsc.2008.02.006

and HOCs are of great importance to fully understand the partition and exposure of POP molecules in environmental systems [9,15–18].

Soot–water partition coefficient (K_{SC}) is one of the most important parameters characterizing sorption behaviors of HOCs and this parameter is indispensable for environmental risk assessment of these chemicals. So far, few studies have been performed on soot–water partition of POPs. For example, some researchers [15,18,19] determined K_{SC} values of a series of POPs using the soot cosolvency-column method, specially designed to study interaction of highly hydrophobic compounds with strongly sorbing matrices such as soot. They found that the affinity of POPs to soot was considerably stronger than the predicted by bulk organic matter partitioning models. Unfortunately, experimentally determined K_{SC} values of POPs are available for only a limited number of POPs because of large expenditures of money and time. Furthermore, because of the limited number of standard POPs, it seems impossible to measure K_{SC} values for all the other POPs. Thus, in order to quickly estimate the environmental behavior of other known POPs, quantitative structure–property relationship (QSPR) models relating soot–water partition coefficients to theoretical molecular descriptors of POPs may be applied to evaluate and predict $\log K_{SC}$ values. Reliable and stable QSPR models can predict $\log K_{SC}$ values efficiently and give some insight into the interactional mechanisms between POPs and soot particles. The purpose of this study is to develop reliable QSPR models to estimate K_{SC} values and deduce the probable interactional mechanisms for soot–water partition of POPs.

Quantum chemical descriptors, such as standard heat of formation (ΔH_f), electronic energy (EE), total energy (TE) and dipole moment (μ) can clearly describe some defined molecular properties. Since they can be easily obtained by computation and are not restricted to closely related compounds, the development of QSPR models in which quantum chemical descriptors are used is of great importance. Partial least squares (PLS) algorithm can analyze data with strongly collinear, noisy and numerous variables [20]. It cannot only search the relationship between dependent variables and predictor variables, but also reduce the dimension of the matrices while concurrently maximizing the relationship between the descriptors. In the present study, PLS algorithm was applied to analyze the interactions between POPs and black carbon affected by quantum chemical descriptors.

2. Materials and methods

2.1. Data set

The $\log K_{SC}$ data determined with the soot cosolvency-column method for ten polychlorinated dibenzo-*p*-dioxins and dibenzofurans (PCDD/Fs), nine polychlorinated biphenyls (PCBs), four polycyclic aromatic hydrocarbons (PAHs) and two polybrominated diphenyl ethers (PBDEs) [15,18,19] were used to develop QSPR models. In the experiments of soot–water partition, diesel particulate soot matter (standard reference material SRM-1650) from the US National Institute of Standards and Technology (NIST, Gaithersburg, MD, USA) was selected as it represents one of the environmentally most relevant anthropogenic types of soot [21,22]. Moreover, 48% of its solid phase consists of soot carbon, and its surface properties are well characterized [21]. The $\log K_{SC}$ data were listed in Table 1. The results demonstrated that a nearly twofold difference exists between the known $\log K_{SC}$ value for naphthalene ($\log K_{SC} = 4.55–4.72$) and that for 2,4,5,2′,4′-pentabromodiphenylether ($\log K_{SC} = 8.11 \pm 0.06$). In present study, these reported soot–water partition coefficients were employed as dependent variables (see Table 1).

2.2. Theoretical molecular structural descriptors

Quantum chemical descriptors of POPs were used to develop QSPR models in present study. These descriptors were obtained from MOPAC 2000 in the CS Chem3D Ultra (Ver. 6.0) and the quantum parameters were computed using semi-empirical parameterized model revision 3 (PM3) Hamiltonian method. The molecular structures were optimized using eigenvector following [23], a geometry optimization procedure, in which the geometry optimization criteria of Gradient Norm was set at 0.1, within MOPAC 2000. A total of 16 derived descriptors reflecting the overall characters of the POP molecules were used in this study. A full list is given in Table 2. The values of the selected molecular descriptors are summarized in Table 3 and the others for the studied compounds are available on request. The compound numbers in Table 3 correspond to those in Table 1. The molecular orbital energies of a given molecule are related to chemical reactivity. Inductive effects and resonance effects exerted by the presence of different substituents and substructural groups within the molecule affect the electron partition and stability of the molecular orbitals. The combinations of frontier molecular orbital energies, $E_{LUMO} - E_{HOMO}$, $(E_{LUMO} - E_{HOMO})^2$ and $E_{LUMO} + E_{HOMO}$, which were proven to be significant in previous QSPR studies of POPs [24,25], were also selected as predictor variables. The $E_{LUMO} - E_{HOMO}$ and $E_{LUMO} + E_{HOMO}$ can be related to absolute hardness and electronegativity, respectively [26,27].

2.3. Statistical analysis

QSPR models were developed using PLS regression, as implemented in the Simca (Simca-S Version 6.0, *Umetri AB and Erisoft AB*) software. The conditions for the computation were based on the default values of the software. The criterion used to determine the model dimensionality – the number of significant PLS components – is cross validation (CV). With CV, when the fraction of the total variation of the dependent variables that can be predicted by a component, Q^2, for the whole data set is larger than