

## QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIPS ON OCTANOL-AIR PARTITION COEFFICIENTS OF POLYCHLORINATED PERSISTENT ORGANIC POLLUTANTS

Jingwen Chen<sup>1,2</sup>, K. -W. Schramm<sup>2</sup>, Xie Quan<sup>1</sup> and A. Kettrup<sup>2,3</sup>

1. School of Environmental Science and Technology, Dalian University of Technology, Dalian 116012, P. R. China; 2. Institute of Ecological Chemistry, GSF-National Research Center for Environment and Health, D-85764, Neuherberg, Germany; 3. Lehrstuhl für Ökologische Chemie und Umweltanalytik, Technische Universität München, 85350 Freising-Weihenstephan, Germany

### Introduction

The octanol-air partition coefficient ( $K_{OA}$ ) is a key descriptor of chemicals partitioning between the atmosphere and organic phases<sup>1</sup>. Recently,  $K_{OA}$  based approaches have been successfully employed to model surface-air partitioning of persistent organic pollutants to aerosols<sup>2-7</sup>, soil<sup>8-9</sup>, vegetation<sup>10-11</sup>, and even indoor carpet<sup>12</sup>. However the experimental determination of  $K_{OA}$  needs special equipment and samples, is expensive and time and labor consuming. Thus it cannot meet the needs of fate assessment of persistent organic pollutants. The methodology of quantitative structure-property relationship (QSPR) is necessary to model and predict  $K_{OA}$ .

Chlorobenzenes (CBs), polychlorinated biphenyls (PCBs), polychlorinated naphthalenes (PCNs), and polychlorinated dibenzo-*p*-dioxin/dibenzofurans (PCDD/Fs) are typical persistent chlorinated organic pollutants. At first, Harner and Mackay determined  $K_{OA}$  values at 25 °C for selected CBs and PCBs using a generator column method<sup>13</sup>. Later, Harner and Bidleman determined  $K_{OA}$  values for selected PCBs (20 °C)<sup>14</sup> and PCNs (25 °C)<sup>15</sup> respectively. Recently, Harner *et al.* determined  $K_{OA}$  values at 25 °C for selected PCDD/Fs<sup>1</sup>. It is purpose of this study to develop QSPR models based on the determined  $K_{OA}$  values.

### Materials and Methods

PM3 Hamiltonian<sup>16</sup> contained in the quantum chemical computation software MOPAC (Ver. 6.0; Stewart, J. J. P., 1990, Frank J. Seiler Research Laboratory, U. S. Air Force Academy, Co 80840) was used to compute quantum chemical descriptors of the chlorinated compounds. A total of 12 MOPAC derived descriptors reflecting the overall character of the chlorinated compounds were selected, as listed in Table 1.

Partial least squares (PLS) analysis<sup>17</sup> was used in developing QSPR models. Simca (Simca-S Version 6.0, Umetri AB & Erisoft AB) software was used to perform the PLS analysis. The criterion used to determine the model dimensionality 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 a significance limit (0.097), the tested PLS component is considered significant. When the cumulative  $Q^2$  for the extracted components,  $Q^2_{cum}$ , is larger than

0.5, the model is considered to have a good prediction ability. Model adequacy was mainly measured as the number of PLS principal components ( $k$ ),  $Q^2_{cum}$ , the correlation coefficient between observed and fitted values ( $R$ ), and the significance level ( $p$ ). Besides these statistics, a standard error ( $SE$ ) was adopted to characterize the prediction precision of models.  $SE$  was defined like that in multiple regression analysis, i.e.,

$$SE = \sqrt{\frac{\sum_{i=1}^n [\log K_{OA}(\text{observed})_i - \log K_{OA}(\text{predicted})_i]^2}{n - k - 1}}$$

Where  $n$  stands for the number of compounds in the training set.

The chlorinated compounds under study were divided into three groups for the QSPR study, (1) CBs, PCNs, p,p'-DDT, (2) PCBs, and (3) PCDD/Fs. For each group, some compounds with reported  $K_{OA}$  values were selected to constitute a validation set. For the validation set of group (1),  $K_{OA}$  values were determined using the generator column method<sup>13,15</sup>. For the validation set of group (2),  $K_{OA}$  values were determined semi-empirically using a multi-column method by Zhang *et al.*<sup>18</sup>. For the validation set of group (3),  $K_{OA}$  values were determined semi-empirically from retention time indices using eq. 2 of Harner *et al.*<sup>1</sup> (Provided by Dr. Tom Harner).

### Results and Discussion

By PLS analysis, QSPR model (1), (2) and (3) were obtained respectively for the three groups of polychlorinated compounds.

$$\begin{aligned} \text{QSPR model (1): } \log K_{OA} (25 \text{ }^\circ\text{C}) = & 1.192 \times 10^{-2} \alpha - 8.034 \times 10^{-5} EE + 9.328 \times 10^{-5} CCR - \\ & 5.311 \times 10^{-4} TE + 4.700 \times 10^{-3} Mw - 5.466 \times 10^{-1} E_{lumo} + 6.388 \times 10^{-1} E_{homo} - \\ & 2.484 Q_C^- + 3.935 \times 10^{-1} Q_{Cl}^+ + 1.510 \times 10^{-2} HOF + 4.355 \times 10^{-1} \mu - 6.595 Q_H^+ \\ & + 7.307 \end{aligned}$$

$$n = 22, k = 2, Q^2_{cum} = 0.973, R = 0.995, p = 1.581 \times 10^{-21}, SE = 0.143$$

$$\text{QSPR model (2): } \log K_{OA} (20 \text{ }^\circ\text{C}) = 3.661 - 1.714 E_{lumo} + 1.762 \times 10^{-2} \alpha - 5.764 \times 10^{-2} HOF -$$

Table 1. List of the molecular structural descriptors

Descriptors	Meanings
$Mw$	Molecular weight
$\alpha$	Average molecular polarizability
$\mu$	Dipole moment
$HOF$	Final heat of formation
$TE$	Total energy
$EE$	Electronic energy
$CCR$	Core-core repulsion energy
$E_{homo}$	The energy of the highest occupied molecular orbital
$E_{lumo}$	The energy of the lowest unoccupied molecular orbital
$Q_C^-$	The largest negative net atomic charge on a carbon atom
$Q_H^+$	The most positive net atomic charges on a hydrogen atom
$Q_{Cl}^+$	The most positive net atomic charges on a chlorine atom

$$2.566 \times 10^{-4} TE + 2.240 \times 10^{-3} Mw - 7.047 Q_{Cl}^+ + 1.603 \times 10^{-5} EE - 3.242 \times 10^{-5} CCR - 2.241 \times 10 Q_C^-$$

$$n = 19, k = 4, Q_{cum}^2 = 0.962, R = 0.988, p = 2.716 \times 10^{-15}, SE = 0.181$$

QSPR model (3):  $\log K_{OA} (25 \text{ }^\circ\text{C}) = 1.068 \times 10^{-2} \alpha + 2.881 \times 10^{-3} Mw - 2.760 \times 10^{-4} TE - 3.353 \times 10^{-5} EE + 3.748 \times 10^{-5} CCR - 1.603 E_{lumo} + 3.904 Q_{Cl}^+ + 3.436$

$$n = 11, k = 2, Q_{cum}^2 = 0.981, R = 0.994, p = 7.602 \times 10^{-10}, SE = 0.147$$

As the cross-validated  $Q_{cum}^2$  values of models (1)~(3) is remarkably larger than 0.50, these three models are surely stable and have a good prediction ability. As shown by Figure 1, for the compounds in training set and validation set, the predicted values were consistent with the corresponding determined values. It has been validated that the three models can be used for predictions of other structural similar chlorinated compounds.

Based on the *VIP* (Variable Importance in the Projection) values and the pseudo-regression coefficients included in QSPR models (1)~(3), the following conclusions can be obtained. (I) The intermolecular dispersive interactions and thus the size of the molecules play a leading role in governing  $\log K_{OA}$ . (II) Increasing  $E_{lumo}$  values of the molecules leads to decreasing  $\log K_{OA}$  values, implying possible intermolecular covalent interactions between the molecules under study and octanol molecules. (III) The intermolecular electrostatic interactions between molecules of octanol and the chlorinated compounds play a secondary role in governing the  $\log K_{OA}$  values.

### Acknowledgements

The study was supported by the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of MOE

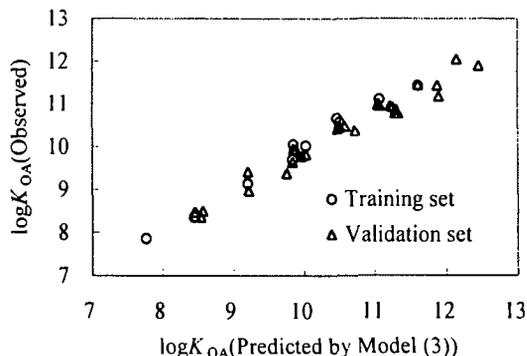
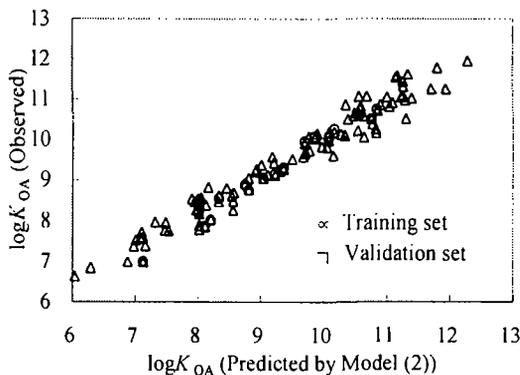
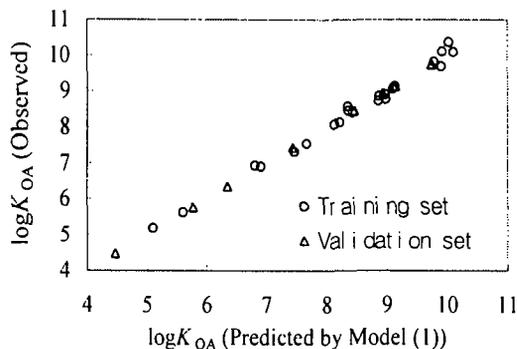


Figure 1. Plot of observed  $\log K_{OA}$  values versus those predicted by QSPR model (1), (2) and (3)

(TRAPOYT), P. R. China. The research results were attained with the assistance of the Alexander von Humboldt (AvH) Foundation and we thank the AvH foundation. We thank Dr. Tom Harner from Meteorological Service of Canada, Environment Canada, for providing us the data.

## References

1. Harner T., Green N. J. L. and Jones K. C. (2000) *Environ Sci Technol.* 34, 3109.
2. Finizio A., Mackay D., Bidleman T. and Harner T. (1997) *Atmos Environ.* 31, 2289.
3. Harner T. and Bidleman T. F. (1998) *Environ Sci Technol.* 32, 1494.
4. Kaupp H. and McLachlan M. S. (1999) *Chemosphere*, 38, 3411.
5. Lee R. G. M. and Jones K. C. (1999) *Environ Sci Technol.* 33, 3596.
6. Lohmann R., Lee R. G. M., Green N. J. L. and Jones K. C. (2000) *Atmos Environ.* 34, 2529.
7. Falconer R. L. and Harner T. (2000) *Atmos Environ.* 34, 4043.
8. Hippelein M. and McLachlan M. S. (1998) *Environ Sci Technol.* 32, 310.
9. Cousins I. T., Beck A. J. and Jones K. C. (1999) *Sci Total Environ.* 228, 5.
10. Thomas G., Sweetman A. J., Ockenden W. A., Mackay D. and Jones K. C. (1998) *Environ Sci Technol.* 32, 936.
11. Weiss P. (2000) *Environ Sci Technol.* 34, 1707.
12. Won D., Corsi R. L. and Kynes M. (2000) *Environ Sci Technol.* 34, 4193.
13. Harner T. and Mackay D. (1995) *Environ Sci Technol.* 29, 1599.
14. Harner T. and Bidleman T. F. (1996) *J Chem Eng Data*, 41, 895.
15. Harner T. and Bidleman T. F. (1998) *J Chem Eng Data*, 43, 40.
16. Stewart J. J. P. (1989) *J Comput Chem.* 10, 209.
17. Wold S., Wold H. and Dunn WJIII. (1984) Report UMINF-83, Department of Chemistry, University of Umeå, Sweden
18. Zhang X. M., Schramm K-W., Henkelman B., Klimm G., Kaune A., Kettrup A., and Lu P. Z. (1999) *Anal Chem.* 71, 3834.