Partitioning behavior of environmental contaminants

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften

(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II – Chemie, Physik und Mathematik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt

von Frau M.Sc. Angelika Stenzel

geboren am 26. September 1985 in Köthen

Gutachter: 1. Prof. Dr. Kai-Uwe Goss

2. Prof. Dr. Beate Escher

Verteidigt am: 03.07.2014

Abstract

The ability to describe the partitioning behavior of environmental contaminants is essential to understand their environmental behavior and to assess potential risks. However, experimental partition coefficients (K) are often missing or questionable especially for environmentally relevant compounds. Hence, for quickly filling data gaps and for a plausibility check of existing experimental data prediction methods are necessary. Well-calibrated poly-parameter linear free energy relationships (pp-LFERs) are an accurate way to predict partition coefficients for neutral organic chemicals based on experimentally determined interaction descriptors. In this work, pp-LFER substance descriptors of 212 environmentally relevant substances, mainly pesticides and flame retardants, were determined experimentally using gas chromatographic (GC) retention times and liquid/liquid partition coefficients.

Substance descriptors were validated through a comparison between predicted and experimental log K for the systems octanol/water (K_{ow}), water/air (K_{wa}), and organic carbon/water (K_{oc}) revealing a high reliability of pp-LFER predictions based on the descriptors determined in this work. For instance, the root mean squared error (rmse) for log K_{ow} was found to be 0.41.

In addition to descriptor determination itself, general guidelines are discussed of how the descriptors for environmentally relevant chemicals should best be determined. Suggestions for the best combination of experimental systems suited for descriptor determination are presented, offering high sensitivity as well as experimental ease, robustness, and accuracy.

However, despite the large number of substance descriptor determined in this work, the availability of experimental substance descriptors especially for complex and multifunctional compounds is still a limiting factor for the application of pp-LFER equations. Therefore, additional prediction methods are required. Possibly suitable methods include COSMOthermX, ABSOLV, and SPARC exhibiting approaches that are mechanistically based. These three methods were validated based on the consistent and diverse experimental data measured in this work. Results indicate that COSMOthermX and ABSOLV are recommendable whereas SPARC performed substantially worse.

Zusammenfassung

Das Verteilungsverhalten von umweltbelastenden Chemikalien beschreiben zu können ist essentiell für das Verständnis ihres Umweltverhaltens und für die Abschätzung potentieller Risiken. Allerdings sind experimentelle Verteilungskoeffizienten (K) häufig nicht verfügbar oder fragwürdig, insbesondere für umweltrelevante Verbindungen. Daher ist der Einsatz von Vorhersagemethoden notwendig. Zuverlässige Vorhersagen von Verteilungskoeffizienten neutraler organischer Chemikalien ermöglicht die Poly-Parameter lineare freie Energie Beziehung (pp-LFER). In dieser Arbeit wurden die pp-LFER Substanz-Deskriptoren von 212 umweltrelevanten Substanzen, hauptsächlich Pestizide und Flammschutzmittel, experimentell bestimmt. Die Bestimmung basiert zum einen auf Retentionszeiten, gemessen an einem Gaschromatographen (GC), und zum anderen auf flüssig/flüssig Verteilungskoeffizienten.

Die Validierung der Substanz-Deskriptoren erfolgte durch einen Vergleich von vorhergesagten und experimentellen Verteilungskoeffizienten in den Systemen Oktanol/Wasser (K_{ow}), Wasser/Luft (K_{wa}) und organischer Kohlenstoff/Wasser (K_{oc}). Die Validierung bestätigt eine hohe Zuverlässigkeit der pp-LFER Vorhersagen mit den Deskriptoren dieser Arbeit, z.B. ist der root mean squared error (rmse) für log K_{ow} 0.41.

Neben der Bestimmung der Substanz-Deskriptoren, ist ein weiterer Schwerpunkt die Diskussion allgemeiner Richtlinien für die Deskriptor Bestimmung umweltrelevanter Substanzen. Diese Arbeit enthält Vorschläge für die optimale Kombination experimenteller Systeme, die sich zur Bestimmung von Substanz-Deskriptoren eignen. Diese Systeme sind gekennzeichnet durch hohe Empfindlichkeit, einfache experimentelle Handhabung, Robustheit und Genauigkeit.

Trotz der hohen Zahl an Substanz-Deskriptoren, die in dieser Arbeit bestimmt wurden, ist die Verfügbarkeit experimenteller Deskriptoren ein limitierender Faktor für die Anwendung der pp-LFER Gleichung. Dies gilt insbesondere für komplexe und multifunktionale Chemikalien. In diesen Fällen müssen zusätzliche Vorhersagemethoden genutzt werden. Mögliche Methoden sind COSMOthermX, ABSOLV und SPARC, da sie über mechanistisch basierte Ansätze verfügen. Die Validierung dieser drei Methoden, basierend auf den konsistenten und vielfältigen experimentellen Daten, die in dieser Arbeit gemessen wurden, zeigt, dass COSMOthermX und ABSOLV empfehlenswert sind. Dagegen ist die Leistungsfähigkeit von SPARC deutlich schlechter.

Preface

The present work was performed between December 2010 to January 2014 at the Helmholtz Centre for Environmental Research, Leipzig at the Department of Analytical Environmental Chemistry. The thesis was written in a cumulative form and is based on the following articles:

A. Stenzel, S. Endo, K. U. Goss: Measurements and prediction of hexadecane/air partition coefficients of 387 environmentally relevant compounds, J. Chromatogr. A 1220 (2012), 132-142
(SI available at:
http://www.sciencedirect.com/science/article/pii/S0021967311017547#MMCvFirst)

A. Stenzel, K. U. Goss, S. Endo: Determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for established and alternative flame retardants, Environ. Sci. Technol. 47 (2013), 1399-1406
(SI available at: http://pubs.acs.org/doi/suppl/10.1021/es304780a)

A. Stenzel, K. U. Goss, S. Endo: Experimental determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for pesticides and other contaminants: New measurements and recommendations, Environ. Sci. Technol. 47 (2013), 14204-14214
(SI available at: http://pubs.acs.org/doi/suppl/10.1021/es404150e)

A. Stenzel, K. U. Goss, S. Endo: Prediction of partition coefficients for complex environmental contaminants: Validation of COSMOtherm, ABSOLV, and SPARC, Environ. Toxicol. Chem. 33 (2014), 1537-1543
(SI available at: http://onlinelibrary.wiley.com/doi/10.1002/etc.2587/suppinfo)

Note that text passages and tables in the summary are partly taken from the original publication without further indication. Abstracts of the original publications were included at the end. Supporting information is either available at the specified internet address or upon request (angelika.stenzel@ufz.de).

Contents

1   Summary: Partitioning behavior of environmental contaminants

1.1   Introduction

The ability to describe the partitioning properties of chemicals is essential to understand their environmental behavior and to assess potential risks. However, the experimental determination of partition coefficients (K) can be challenging, especially for extremely large or small K values due to low solute concentrations in one phase. Such measurements are prone to elevated experimental artifacts due to, e.g., a loss to a third phase in the system, incomplete phase separation, and contaminations in the experimental procedure.[1] Furthermore, the experimental measurements are time consuming and not feasible for tens of thousands of chemicals that need to be screened for their environmental behavior. Often, it is also difficult to assess the reliability of experimental values which can disagree widely depending on the sources.[2] In these cases, prediction methods are a useful and necessary alternative.

There are various quantitative structure activity relationships (QSARs) and other models that predict partition coefficients based only on the molecular structure.[3-5] Most of the available models are highly empirical and can only account for chemical substructures, molecular interactions, and partitioning systems included in the respective calibration data sets. Hence, these QSAR are limited by the size of the available calibration data set and do not perform well as a predictive tool for new chemical structures.[6] Furthermore, it is practically not possible to construct meaningful QSAR for environmentally relevant partition coefficients such as lipid/water, protein/water, and aerosol/air partition coefficients because of limited data. Simple correlations with the log of the octanol/water partition coefficient (K_{ow}) do not provide accurate estimations when diverse substances are considered.[7, 8]

An alternative approach for estimating partition coefficients is provided by the poly-parameter linear free energy relationship (pp-LFER) model. The general form of a pp-LFER equation is shown below:[9]

$$\log K = s\,S + a\,A + b\,B + v\,V + l\,L + c \tag{1}$$

K is the partition coefficient between two arbitrary phases. The capital letters on the right hand side of the equation describe the properties of the distributed substance, and the small letters represent the complementary system properties. L is the hexadecane/air partition

coefficient, S characterizes the dipolarity/polarizability of the substance, A describes the H-bond donor properties, B the H-bond acceptor properties, and V is the molar volume. L and V represent the nonspecific interactions of the substance: van der Waals interactions and interactions related to cavity formation. S, A and B characterize the specific interactions. Note that other forms of pp-LFER equations also exist in which the E descriptor, the excess molar refraction, replaces the L or V descriptor of eq 1.[9, 10] Calibrated pp-LFER system parameters are available for a multitude of environmentally relevant systems, e.g., water/air, soil organic carbon/water, storage lipid/water, membrane lipid/ water, and protein/water.[7, 9, 11-17] If these calibrated pp-LFER system parameters are combined with experimental substance descriptors they typically allow predictions of log K with root mean squared errors (rmse) between 0.2 (for well defined homogeneous systems like water, solvents, lipids, and polymers) to 0.5 (for more heterogeneous systems like humic matter, aerosols, proteins, and others). Besides its accuracy, the pp-LFER approach is very effective because only one set of substance descriptors is needed for any partitioning system. Furthermore, the approach allows a flexible substance descriptor determination, i.e., experimentally difficult partitioning systems can be avoided for substance descriptor determination, because substance descriptors are in principle system-independent.[18]

The limiting factor to apply pp-LFER equations is usually the availability of experimental substance descriptors that are to date still more accurate than predicted ones, i.e., typical rmse using predicted descriptors are between 0.7 to 1.0 log units.[19] By now, experimental substance descriptors have been determined for many compounds (> 3,000). However, the majority of descriptors refer to simple, monofunctional structures. The number of substance descriptors being available for complex, multifunctional compounds is still small compared to the number of compounds used now and in the past. Hence, additional prediction methods are still required for substances without experimental pp-LFER substance descriptors.

Prediction methods used in addition to the pp-LFER approach should not suffer from the above-mentioned drawbacks of simple QSARs. Suitable methods include COSMOthermX, ABSOLV and SPARC, which are expected to have wide application ranges due to their mechanistically based approaches. All three also have the advantage not to be limited to a specific partitioning system and they performed well in previous validations.[19-24] However, a thorough validation of COSMOthermX, ABSOLV and SPARC for a large set of complex, multifunctional substances in differing systems has been missing so far.

1.2   Objective of this study

A main goal of this work is to expand the applicability of the pp-LFER approach to complex substances, i.e., this work seeks to extend the pp-LFER substance descriptor set and to improve the prediction accuracy of already available descriptors. Therefore, pp-LFER substance descriptors for a large number of complex, multifunctional, and environmentally relevant substances such as pesticides and flame retardants are to be determined experimentally. Accuracy of determined descriptors is to be assured by internal and external validation.

In addition, general schemes are to be discussed of how the descriptors for environmentally relevant chemicals should best be determined. This work wants to provide general recommendations for experimental determination of pp-LFER substance descriptors and present suggestions for the best combination of experimental systems suited for descriptor determination.

Finally, experimental data measured in this work shall be used to validate the three prediction methods COSMOthermX, ABSOLV and SPARC. The consistent, diverse, and quality controlled experimental data set should allow a rigorous validation that can outline the possibilities and constraints of COSMOthermX, ABSOLV, and SPARC.

1.3   Materials

Chemicals used as calibration substances and analytes were purchased from ABCR GmbH & Co. KG (Karlsruhe, Germany), Sigma-Aldrich Chemie GmbH (Steinheim, Germany), TCI Europe N.V. (Zwijndrecht, Belgium), Accu Standard Inc. (New Haven, USA), Dr. Ehrenstorfer GmbH (Augsburg, Germany) and Alfa Aesar GmbH & Co. KG (Karlsruhe, Germany). Solvents (purity at least 99%) were purchased from Merck KGaA (Darmstadt, Germany), Carl Roth GmbH & Co. KG (Karlsruhe, Germany), Fluorochem Ltd. (Hadfield, United Kingdom) and Sigma-Aldrich. Water was purified with a Milli-Q A10 Ultrapure Water Purification System (Millipore GmbH, Schwalbach, Germany) before use. Poly(dimethyl siloxane) (PDMS) sheets (Specialty Silicone Products Inc., Ballston Spa, USA; density 1.17 g mL^{-1}) as well as PDMS-coated glass fibers (Polymicro Technologies Inc., Phoenix, USA) were used for experiments.

1.4   Experimental determination of pp-LFER system parameters and substance descriptors

General approach

Five pp-LFER substance descriptors need to be determined according to eq 1. The molar volume V does not need to be determined experimentally because it can be calculated through the McGowan approach[25] from the structure of the molecule. The remaining four descriptors S, A, B and L need to be calibrated based on experimental partition coefficients or retention times. In this work, three steps were taken to accomplish this task. First, pp-LFER system parameters were (re-)calibrated for several partitioning systems including gas chromatographic (GC) columns using probe compounds for which pp-LFER substance descriptors are available. Second, pp-LFER substance descriptors of the target compounds were determined from experimental partition coefficients or retention times measured in this work by using the calibrated pp-LFER equations. Third, to validate the determined pp-LFER substance descriptors, they were used to predict several partition coefficients (e.g., K_{ow}, water/air partition coefficients (K_{wa})), which were compared with experimental data from the literature (the third step is explained in section 1.5).

Selection of partitioning systems

To determine four descriptors for each chemical, at least four calibrated pp-LFER equations are necessary. More than four equations were prepared in this work, providing an over-determined equation system and the opportunity to check the consistency of the obtained substance descriptors. Regarding the selection of partitioning systems, it is important to have diverse partitioning systems that have varying sensitivity to the various substance descriptors. Also, the experimental methods chosen for partition coefficient/retention time determination of calibration as well as target compounds are a crucial point because the range of partition coefficients/retention times can be very large, and the goal was to analyze a large number of substances within a reasonable time. In this work, two complementary methods were chosen: the measurement of GC net retention times (t_{net}) and the determination of liquid/liquid and PDMS/water partition coefficients.

GC retention time measurements are simple, reproducible and offer the possibility to measure a wide range of compounds by temperature variation. The columns used in this work were SPB Octyl (Supelco, Taufkirchen, Germany) and DB-200 (Agilent Technologies Deutschland GmbH, Böblingen, Germany). The retention on SPB Octyl is only influenced by nonspecific

interactions (*l* and *v*; see also refs 16 and 26). DB-200 has a significant dipolarity/polarizability (*s*) in addition to nonspecific interaction properties.[16] Measurements on two other polar columns (HP-INNOWax, Agilent and SP-2340, Supelco) were done but could not be used for descriptor determination because either the retention times were not reproducible (INNOWax) or the calibration was not consistent (SP-2340).

Organic liquid/organic liquid partition coefficients are relatively easy to measure even with hydrophobic compounds because concentration difference between the two phases is relatively small and the solubilities in organic liquids are high. The systems ethylene glycol/1,2-dichloroethane and heptane/propylene carbonate were chosen because the former possesses high *a* and the latter high *s* and *a* values.[17] In addition, measurements in the heptane/2,2,2-trifluoroethanol and heptane/1,1,1,3,3,3-hexafluoroisopropanol system were conducted, because they are characterized by a high *b* value.[27] However, in this work both fluoroalcohols were found to have a *B* sensitivity that differs from water, i.e., for some compounds *B* values determined in these systems may differ substantially from *B* values determined in water-containing systems. Hence, if one wants to predict environmentally relevant partition coefficients, *B* values should be calibrated using water-containing systems, because water is the most important and strongest H-bond donor solvent in the environment. The water-containing systems chosen in this work were PDMS/water and cyclohexane/water.

Experimental methods

Retention time measurements were performed on a gas chromatograph (7890A GC System) coupled to a mass spectrometer (5975C inert MSD, both Agilent). Concentration measurements necessary for partition coefficient determination were carried out on the mentioned GC-MS for the majority of the compounds (column: HP-5MS UI, Agilent).

Liquid/liquid *K* for completely organic systems as well as cyclohexane/water were measured with a shake-flask method based on the method by Ahmed and Poole.[28] $K_{PDMS/water}$ values were determined with three established methods depending on the expected log $K_{PDMS/water}$ value.[29-31] A shake-flask method and a depletion method with PDMS coated glass fibers were used for log $K_{PDMS/water} < 4$ and a passive-dosing method with preloaded PDMS sheets for log $K_{PDMS/water}$ between 4 and 6. Substances having an expected log $K_{PDMS/water} > 6$ were not measured because the reliability of log $K_{PDMS/water}$ is comparably low.

In total, 452 GC retention times and 517 partition coefficients (liquid/liquid and PDMS/water) were measured and used to calibrate the pp-LFER substance descriptors (not including measurements to calibrate system parameters). GC retention times were highly reproducible

with a typical standard deviation (sd) < 0.003 log units. Partition coefficients exhibit sd typically < 0.1 log units, with 0.30 (liquid systems) and 0.27 log units (PDMS/water) in the worst cases.

Calibration of system parameters

The pp-LFER system parameters were determined by multiple linear regression analysis with substance descriptors from the literature and experimental retention times or partition coefficients. Retention times were all determined in this work. Partition coefficients of calibration compounds for liquid/liquid systems and PDMS/water were mostly taken from the literature[17, 27, 32-34] and partly complemented with new data from this work. System parameters for the systems used in substance descriptor determination are provided in Table 1.

Table 1. pp-LFER system parameters and statistics for the systems (GC-columns and liquid systems) used for pp-LFER substance descriptor determination.[a]

system	T [°C]	l	v	b	a	s	c	r²	sd	n
SPB Octyl	180	0.50 (±0.01)	-0.44 (±0.02)	-	-	-	-2.29 (±0.03)	0.995	0.043	56
	250	0.38 (±0.01)	-0.42 (±0.01)	-	-	-	-2.25 (±0.03)	0.994	0.051	59
DB-200	160	0.28 (±0.01)	0.16 (±0.04)	-	0.27 (±0.04)	0.67 (±0.04)	-2.52 (±0.06)	0.981	0.078	59
	180	0.23 (±0.01)	0.21 (±0.05)	-	0.20 (±0.04)	0.66 (±0.04)	-2.49 (±0.05)	0.985	0.071	60
	250	0.21 (±0.01)	-0.02 (±0.05)	-	0.15 (±0.04)	0.49 (±0.04)	-2.61 (±0.07)	0.985	0.063	37
heptane/ propylene carbonate	25	0.25 (±0.04)	0.09 (±0.14)	-0.61 (±0.13)	-2.85 (±0.10)	-2.12 (±0.13)	0.54 (±0.12)	0.971	0.223	83
ethylene glycol/ 1,2-dichloro-ethane	25	-0.03 (±0.04)	-0.76 (±0.16)	0.74 (±0.17)	2.27 (±0.12)	0.13 (±0.14)	-0.48 (±0.13)	0.950	0.227	54
PDMS/water[b]	25	0.37 (±0.04)	2.37 (±0.19)	-3.84 (±0.11)	-2.85 (±0.13)	-1.55 (±0.12)	0.46 (±0.08)	0.986	0.243	157
cyclohexane/ water[c]	25	0.40 (±0.04)	3.40 (±0.16)	-5.25 (±0.09)	-3.99 (±0.07)	-1.83 (±0.09)	0.49 (±0.07)	0.984	0.200	159

[a]values in parentheses: standard error of coefficients
[b]recalibration using partition coefficients from ref 33
[c]recalibration using partition coefficients from ref 34

Determination of substance descriptors

Substance descriptors were derived by using the Excel Solver to minimize the rmse by simultaneously adjusting the substance descriptors. The rmse is defined as:

$$rmse = \sqrt{\frac{\sum \left(\log K_{exp} - \log K_{pred}\right)^2}{n}} \qquad (2)$$

K_{exp} is the experimentally determined partition coefficient (or t_{net}), K_{pred} the predicted partition coefficient (or t_{net}), and n the number of available log K/log t_{net}. Here and for substance descriptor validation (section 1.5), predicted values refer to calculations with pp-LFER equations, i.e., equations in Table 1 for descriptor determination as well as internal validation and equations of chosen external validation systems such as octanol/water or water/air for external descriptor validation. In the COSMOthermX, ABSOLV, and SPARC validation (section 1.7), predicted values refer to calculations with the respective method.

Final pp-LFER substance descriptors are displayed in Table 2 (for comfortable reading Table 2 is provided separately in section 1.10). In total, substance descriptors were determined for 212 substances (isomers are counted separately) comprising 136 pesticides, 40 flame retardants, 14 polychlorinated biphenyls (PCBs), 8 nitroaromatic compounds, 6 heterocyclic aromatic compounds containing either N or O, 5 phthalate esters, 2 pharmaceuticals, and 1 hormone (pesticides, flame retardants, and PCBs include transformation products).

1.5 Validation of determined pp-LFER substance descriptors

Internal validation

The rmse (eq 2) resulting from the adjustment of the descriptors made it possible to check the internal consistency of the determined substance descriptors. Moreover, this "internal rmse" can serve as a quantitative indicator for substance descriptor reliability. The smaller the internal rmse the more reliable are the descriptors. Internal rmse values are displayed in Table 2 as well. The mean internal rmse is 0.08 (range 0.00-0.35) indicating an overall high internal consistency of the determined descriptors.

External validation

External consistency was checked through the comparison of pp-LFER predicted and experimental log K collected from the literature for K_{ow}, K_{wa}, and organic carbon/water (K_{oc}) (system parameters are from the literature).[7, 9] These validation systems were used because they are virtually the only partitioning systems for which enough literature data are available to perform a thorough validation.

Experimental log K to be used in external validation were mainly collected from three databases (the experimental databases stored in the EPISuite™, the Pesticide Properties Database, and LOGKOW© - A databank of evaluated octanol-water partition coefficients) complemented by additional literature data in some cases. If more than one literature value was available, the mean of the log values was used.

The comparison between predicted and experimental literature log K_{ow} is displayed in Figure 1.

Figure 1. Comparison between experimental (x-axis) and pp-LFER predicted (y-axis) logarithmic partition coefficients for the system octanol/water (K_{ow}). The continuous line is the 1:1 line; the two dashed lines mark a difference of ± 1 log unit from the 1:1 line.

The rmse resulting from the differences between experimental and predicted log K_{ow} is 0.41. This agreement is regarded as high, considering the fact that the descriptors have been calibrated independently from K_{ow}, and the experimental K_{ow} data are from diverse sources with a varying degree of uncertainty. The predicted log K_{ow} for only two compounds show a

deviation of slightly more than 1 log unit from the experimental value: hexabromocyclododecane (deviation 1.03 log units) and tetrabromobisphenol A (deviation 1.08 log units).

The results of the comparison between experimental and pp-LFER predicted log K_{wa} displayed in Figure 2 scatter more, with an rmse of 1.32.

Figure 2. Comparison between experimental (x-axis) and pp-LFER predicted (y-axis) logarithmic partition coefficients for the system water/air (K_{wa}). The continuous line is the 1:1 line; the two dashed lines mark a difference of ± 1 log unit from the 1:1 line, and the two dotted lines mark a difference of ± 3 log units.

As was also mentioned by Tülp et al.,[35] the strong scattering is likely due to a high uncertainty in experimental data of K_{wa}. Often, the log K_{wa} are derived from water solubility and saturation vapor pressure, and the measurement of a small value of either of the properties is prone to large errors.[2, 36] In this study five compounds exhibit a deviation of more than 3 log units from the experimental log K_{wa}: tris(2-chloroethyl)phosphate, hexabromocyclododecane, fensulfothion, fenarimol, and tecnazene possessing the largest deviation of 4.76 log units. Such large errors are unlikely caused only by the descriptor values that at the same time provide reasonable predictions for log K_{ow}. An exception is hexabromocyclododecane which has different isomers not considered in this study (i.e., a technical mixture was used). The literature experimental partitioning data are also not for specific isomers. There is an indication that the partitioning properties differ across hexabromocyclododecane isomers.[37, 38]

The rmse obtained from comparing experimental and pp-LFER predicted log K_{oc} is 0.74 (comparison in Figure 3).

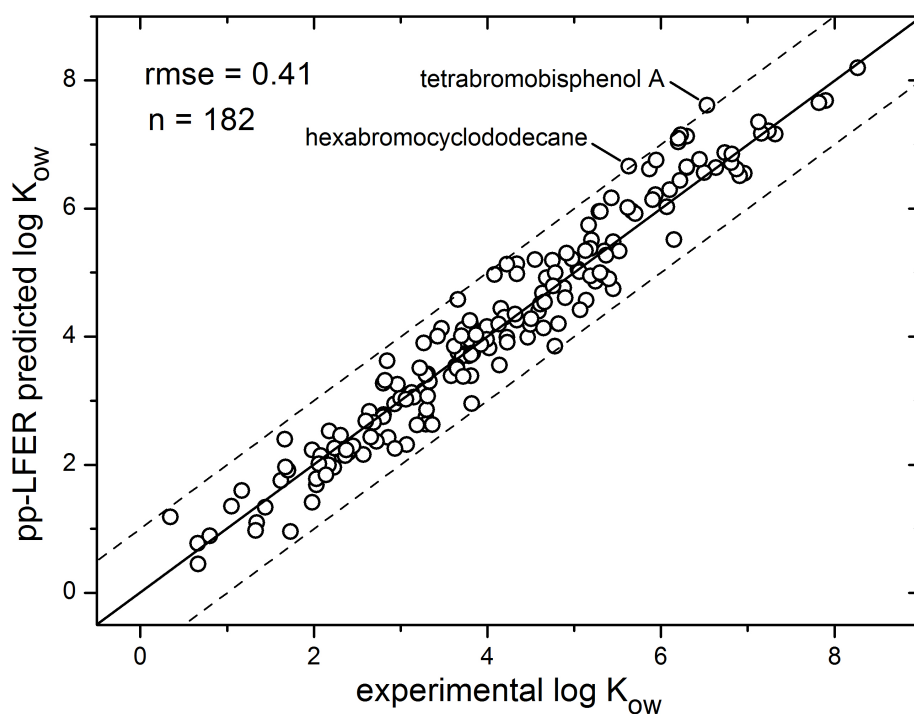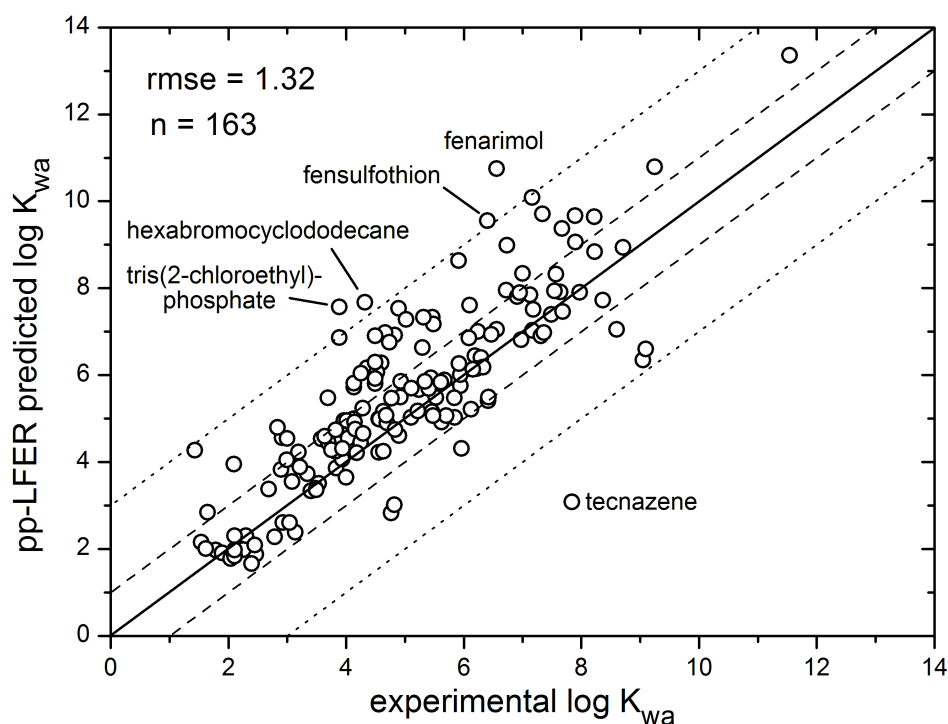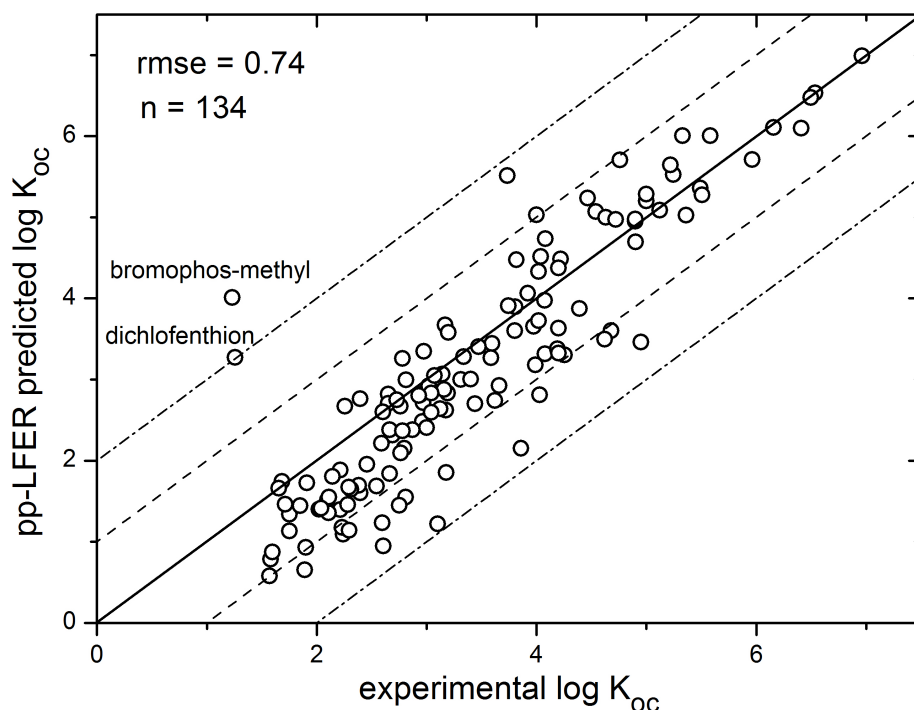**Figure 3**. Comparison between experimental (x-axis) and pp-LFER predicted (y-axis) logarithmic partition coefficients for the system organic carbon/water (K_{oc}). The continuous line is the 1:1 line; the two dashed lines mark a difference of ± 1 log unit from the 1:1 line, and the two dotted-dashed lines mark a difference of ± 2 log units.

The rmse is higher than that for the log K_{ow} comparison but is still considered reasonable, because organic carbon is not a uniquely defined phase like octanol or water. That is, there are different forms of organic matter contributing to the organic carbon content. Natural variability of log K_{oc} for different soils and sediments is about ± 0.3 log units sd.[39] In addition, pp-LFER predictions appear to be less accurate for K_{oc} than for well-defined organic phases because of heterogeneous sorption sites on organic matter molecules.[7, 40] The two compounds possessing the highest deviations are bromophos-methyl (deviation 2.78 log units) and dichlofenthion (deviation 2.02 log units).

Comparison to previously published substance descriptors

Substance descriptors were available in the literature only for about one-third of the target compounds (78 out of 212). Thus, the descriptors presented in this work significantly extend the set of available pp-LFER substance descriptors especially for complex and multifunctional compounds. To compare the descriptors from this work with previously

measured ones, experimental log K_{ow}, log K_{wa}, and log K_{oc} were compared to the pp-LFER predicted values. Predictions using the descriptors from this work were compared to predictions from descriptors determined by Tülp et al.[35], Bronner et al.[26] or other sources (mainly from Abraham or Poole and coworkers[1, 41-49]). The resulting rmse values are displayed in Table 3. Rmse values were calculated for subsets of the substance list as well to allow a direct comparison of rmse values based on the descriptors of this study and the literature. Note that Tülp et al.[35] did not report L but E. Therefore, partition coefficients were predicted using S, A, B, V, and E in the case of the Tülp descriptors.

Table 3. Rmse values resulting from the comparison between experimental and predicted log K using substance descriptors from this study as well as previous literature sources.

substance descriptors from	log K_{ow}		log K_{wa}		log K_{oc}	
	rmse	n	rmse	n	rmse	n
this work (complete data set)	0.41	182	1.32	163	0.74	134
Tülp et al.[35]	0.55	34	1.82	30	0.72	34
this work (restricted data set)[a]	0.35	34	0.93	30	0.63	34
Bronner et al.[26]	0.60	33	1.03	28	0.61	33
this work (restricted data set)[b]	0.32	33	0.89	28	0.57	33
other sources[c,1, 41-49]	0.42	44	0.54	40	0.67	32
this work (restricted data set)[d]	0.37	44	0.66	40	0.62	32

[a]restriction to compound set of Tülp et al.[35]
[b]restriction to compound set of Bronner et al.[26]
[c]collected using UFZ-LSER database[50]; descriptors from Tülp et al.[35] as well as Bronner et al.[26] were excluded
[d]restriction to compound set from other sources[1, 41-49]

In comparison to the descriptor sets from Tülp et al.[35] as well as Bronner et al.,[26] the descriptors from the present study resulted in lower rmse values, i.e., strong improvement in prediction accuracy for K_{ow} and K_{wa} and slight improvement for K_{oc}. This improvement can be explained by the different experimental methods used to determine the descriptors and in case of the Tülp descriptors partly by the different pp-LFER equation type.

The comparison of the descriptors from this study and those from sources other than Tülp et al. and Bronner et al. (i.e., refs 1, 41-49) reveals comparable prediction accuracy for all of

K_{ow}, K_{wa}, and K_{oc}. As the calibration methods differ in different sources and are sometimes unknown, it is difficult to interpret these results. The good overall agreement between predictions with descriptors from this work and literature descriptor predictions is considered as a support for the approach to reduce the set of calibration systems to the necessary minimum (i.e., five systems to determine four descriptors).

1.6   General recommendations for the experimental determination of pp-LFER substance descriptors

Criteria

The recommendations for determining pp-LFER descriptors for environmentally relevant compounds are based on experience from this work as well as experience from others (Poole et al.[1, 40] and Abraham et al.[10]). The suitability of the partitioning systems to be used for descriptor calibration can be evaluated based on three criteria: a) sensitivity to the descriptors that can be quantitatively assessed with the magnitude of the system parameters, b) accuracy of the model fit, and c) experimental applicability.

Methods

The commonly used methods for descriptor determination, namely, GC, high performance liquid chromatography (HPLC), and liquid/liquid partitioning are evaluated.

GC retention times can generally be measured accurately and precisely with a high through-put, and pp-LFER models for many GC columns are of excellent quality (typical sd 0.01-0.04).[1] Of all systems, GC offers the highest sensitivity to L especially on completely nonpolar columns such as SPB Octyl. Sensitivities to S and A are moderate compared to other systems. Slightly polar columns such as DB-200 (suitable to determine S) can be recommended. More polar columns such as HP-INNOWax and SP-2340 have high s and a values according to the system parameters published by Poole and Poole.[16] However, in this work these columns could not be used for substance descriptor determination (compare "Selection of partitioning systems", section 1.4). The largest shortcoming of the GC system is that no conventional GC columns exhibit H-bond donor properties (i.e., b is zero).

HPLC retention times can also be measured reproducibly and accurately. The through-put is also high, though the fitting of pp-LFER models are less precise compared to GC systems (typical sd 0.02-0.07).[1] HPLC has not been used in this work due to its drawbacks, i.e., small

system parameters in combination with comparably high sd, leading to high uncertainties especially for S and A.

Liquid/liquid partitioning systems are generally characterized by large system parameters for s, a, and b. Compared to chromatographic retention time measurements, liquid/liquid partitioning experiments generally are more labor intensive and somewhat less reproducible and accurate (typical model sd 0.1-0.5).[1] Water-containing systems always exhibit high b values.[1] Therefore, any aqueous system is well suited for improved determination of B. However, low solubility or instability of the target chemical in water can make accurate determination of the partition coefficient difficult. Partitioning systems with two immiscible organic solvents, which in contrast to water offer greater solubility and more stability, are useful alternatives.[17, 32, 51] Here, we decided on heptane/propylene carbonate (high s and a) as well as ethylene glycol/1,2-dichloroethane (high a). Heptane/fluoroalcohol systems should be suitable to determine B. The experiments in this work confirm this in general, but for some substances, B values calibrated on such fluoroalcohol systems do not accurately predict partition coefficients of water-containing systems such as K_{ow}. Hence, water-containing systems are considered necessary for determining the B descriptors for environmental systems, because many important environmental partitioning systems contain the water phase and because it is not possible to predict a priori which compounds would exhibit system-dependent B values (compare "Selection of partitioning systems", section 1.4).

Recommendations

To summarize, the recommendations for a minimal set of systems to determine pp-LFER substance descriptors are: GC-SPB Octyl (for L), GC-DB-200 (for S), heptane/propylene carbonate (for S and A), ethylene glycol/dichloroethane (for A), and a water-containing system such as PDMS/water or cyclohexane/water (for B). Any additional GC, LC, or partitioning data would improve the accuracy of determined descriptors and it would help to check internal consistency, though at the cost of more experimental work.

1.7   Validation of COSMOthermX, ABSOLV, and SPARC

Prediction models

The prediction models to be evaluated exploit different approaches. COSMOthermX predicts a large variety of properties based on the COSMO-RS theory, which uses quantum-chemical calculations and statistical thermodynamics.[52, 53] ABSOLV predicts the pp-LFER substance

descriptors by summing up molecular fragment contributions.[54] SPARC relies on a combination of linear free energy theory for thermodynamic properties and perturbation molecular orbital methods for quantum mechanic contributions such as delocalization energies.[23, 24, 55] ABSOLV and SPARC both use molecular increments and require calibration with experimental data. However, the molecular fragments are not calibrated for their contribution to a specific partition process but for their fundamental intermolecular interaction properties. By this mechanistic approach the calibration with experimental data becomes more meaningful than in the case of simple QSARs and more robust because virtually any type of partitioning data can be used for calibration. Still, it is important to realize that both methods have an application domain that is limited by their calibration. In contrast, the COSMOthermX approach contains only a few empirical parameters which do not infer any limitation in the application domain. Initially, our validated models also included connectivity indices. They were not considered any further due to their low accuracy in predicting log $K_{hexadecane/air}$.

Validation results

The validation is based on experimental partition coefficients and GC net retention times measured in this work. Prediction of GC net retention times turned out to be no decisive part in this validation. Hence, it will not be discussed in this summary. The rmse values and mean errors for the partition coefficients, i.e., the liquid/liquid systems, PDMS/water, and hexadecane/air are provided in Table 4.

Table 4. Rmse and mean error resulting from the comparison between experimental and predicted log K.[a]

	COSMOthermX[b]			ABSOLV			SPARC		
	rmse	mean error	n	rmse	mean error	n	rmse	mean error	n
PDMS/water	0.90	0.29	154	0.95	0.34	159	1.43	0.68	157
heptane/ 2,2,2-trifluoroethanol	0.58	-0.12	45	0.79	-0.31	49	2.85	2.45	49
heptane/ propylene carbonate	0.58	0.17	255	0.83	-0.05	270	2.63	1.85	267
ethylene glycol/ 1,2-dichloroethane	0.79	-0.39	175	0.64	-0.25	185	1.55	-1.17	182
hexadecane/air	0.83	0.20	364	1.01	0.63	387	1.00	0.42	381

[a]temperature always 25°C
[b]version C30_1401, parameterization BP-TZVPD-FINE

In the liquid/liquid systems and PDMS/water, COSMOthermX (rmse: 0.58-0.90) and ABSOLV (rmse: 0.64-0.95) gave more accurate predictions than SPARC (rmse: 1.43-2.85). In predicting log $K_{hexadecane/air}$, COSMOthermX was the most accurate (rmse 0.83) and results for ABSOLV (rmse 1.01) and SPARC (rmse 1.00) were comparable. The relatively low performance of SPARC stands in contrast to earlier experiences where SPARC had performed similarly well compared to COSMOthermX for systems like octanol/water and air/water.[13, 56, 57] The results here suggest that SPARC's prediction accuracy for complex chemicals is substantially lower than that for simple chemicals. It should also be noted, that even the rmse for the most successful case is considerably higher than typical errors in good experimental values or predictions based on the pp-LFER approach with experimental pp-LFER descriptors. It is thus desirable for many environmental applications that all methods are further improved. In addition, different COSMOthermX versions and parameterizations were compared revealing that the newest version C30_1401 in combination with the BP-TZVPD-FINE parameterization consistently provides comparable or more accurate predictions.

Summing up the validation results, it can be concluded that SPARC performed substantially worse than ABSOLV and COSMOthermX when complex compounds are dealt with. The overall performance of COSMOthermX and ABSOLV is comparable. Hence, both methods can be recommended. The approach of COSMOthermX is more general and more flexible. However, COSMOthermX requires specific input files whose calculation can be time consuming for large molecules and the handling of the COSMOthermX software cannot be mastered within a few minutes. ABSOLV, on the other hand, is only useful if the pp-LFER system parameters are available for the considered system but it is very quick and easy to handle.

1.8   Conclusions

In this work, pp-LFER substance descriptors were determined experimentally for 212 complex, multifunctional, and environmentally relevant substances including 136 pesticides and 40 flame retardants. Determined substance descriptors were subject to internal and external validation. Internal validation was possible due to an over-determined equation system and external validation was done by comparing predicted and experimental literature K_{ow}, K_{wa}, and K_{oc}. Validations indicated a high reliability of pp-LFER predictions based on the determined descriptors (e.g., rmse of 0.41 for log K_{ow}). The complete set of descriptors for 134 compounds was measured for the first time in this work. Hence, the descriptors presented

in this work in combination with existing pp-LFER system equations substantially extend (and in some cases correct) the knowledge on partitioning properties of the 212 target substances. This knowledge is important to understand the environmental behavior of these substances and to assess potential risks such as the bioaccumulation in organisms.

In addition, the results of this work provide insight into some general guidelines with respect to the method combination best suited for deriving descriptors for environmentally relevant compounds. These guidelines can facilitate future determinations of pp-LFER substance descriptors possibly as a routine method.

Finally, the prediction methods COSMOthermX, ABSOLV, and SPARC, being possible alternatives when no experimental pp-LFER substance descriptors are available, were validated. The thorough validation revealed a significant lack of accuracy for SPARC in predicting complex substances. Prediction accuracy of COSMOthermX and ABSOLV is comparable and both methods can be recommended though the achieved rmse of 0.6-0.9 for log K might be higher than, e.g., experimental errors. Nevertheless, this rmse may be sufficient for a first screening of diverse chemicals in terms of their partitioning properties making COSMOthermX and ABSOLV useful for regulatory purposes where very diverse and new chemical structures have to be assessed for their partitioning behavior.

1.9   References

(1) Poole, C.F.; Atapattu, S.N.; Poole, S.K.; Bell, A.K. Determination of solute descriptors by chromatographic methods. *Anal. Chim. Acta.* **2009**, *652*, 32-53.

(2) Pontolillo, J.; Eganhouse, R.P., *The search for reliable aqueous solubility (S_w) and octanol-water partition coefficient (K_ow) data for hydrophobic organic compounds: DDT and DDE as a case study.* 2001, U.S. Department of the Interior, U.S. Geological Survey: Water-Resources Investigations Report 01-4201. Reston, Virginia.

(3) Katritzky, A.R.; Kuanar, M.; Slavov, S.; Hall, C.D.; Karelson, M.; Kahn, I.; Dobchev, D.A. Quantitative correlation of physical and chemical properties with chemical structure: Utility for prediction. *Chem. Rev.* **2010**, *110*, 5714-5789.

(4) Dearden, J.C.; Rotureau, P.; Fayet, G. QSPR prediction of physico-chemical properties for REACH. *SAR QSAR Environ. Res.* **2013**, *24*, 279-318.

(5) Boethling, R.S.; Howard, P.H.; Meylan, W.M. Finding and estimating chemical property data for environmental assessment. *Environ. Toxicol. Chem.* **2004**, *23*, 2290-2308.

(6) Avdeef, A.; Bendels, S.; Di, L.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA-Critical factors for better predictions of absorption. *J. Pharm. Sci.* **2007**, *96*, 2893-2909.

(7) Bronner, G.; Goss, K.U. Predicting sorption of pesticides and other multifunctional organic chemicals to soil organic carbon. *Environ. Sci. Technol.* **2011**, *45*, 1313-1319.

(8) Harris, A.; Dangerfield, N.; Woudneh, M.; Brown, T.; Verrin, S.; Ross, P. Partitioning of current-use and legacy pesticides in salmon habitat in British Columbia, Canada. *Environ. Toxicol. Chem.* **2008**, *27*, 2253-2262.

(9) Goss, K.-U. Predicting the equilibrium partitioning of organic compounds using just one linear solvation energy relationship (LSER). *Fluid Phase Equilib.* **2005**, *233*, 19-22.

(10) Abraham, M.H.; Ibrahim, A.; Zissimos, A.M. Determination of sets of solute descriptors from chromatographic measurements. *J. Chromatogr., A.* **2004**, *1037*, 29-47.

(11) Geisler, A.; Endo, S.; Goss, K.U. Partitioning of organic chemicals to storage lipids: elucidating the dependence on fatty acid composition and temperature. *Environ. Sci. Technol.* **2012**, *46*, 9519-9524.

(12) Endo, S.; Bauerfeind, J.; Goss, K.U. Partitioning of neutral organic compounds to structural proteins. *Environ. Sci. Technol.* **2012**, *46*, 12697-12703.

(13) Endo, S.; Escher, B.; Goss, K.U. Capacities of membrane lipids to accumulate neutral organic chemicals. *Environ. Sci. Technol.* **2011**, *45*, 5912-5921.

(14) Arp, H.P.H.; Schwarzenbach, R.P.; Goss, K.-U. Ambient gas/particle partitioning II: The influence of particle source and temperature on sorption to dry terrestrial aerosols *Environ. Sci. Technol.* **2008**, *42*, 5951-5957.

(15) Abraham, M.H.; Ibrahim, A.; Acree, W.E. Air to blood distribution of volatile organic compounds: A linear free energy analysis. *Chem. Res. Toxicol.* **2005**, *18*, 904-911.

(16) Poole, C.F.; Poole, S.K. Separation characteristics of wall-coated open-tubular columns for gas chromatography. *J. Chromatogr., A.* **2008**, *1184*, 254-280.

(17) Karunasekara, T.; Poole, C.F. Models for liquid-liquid partition in the system ethylene glycol-organic solvent and their use for estimating descriptors for organic compounds. *Chromatographia.* **2011**, *73*, 941-951.

(18) Vitha, M.; Carr, P.W. The chemical interpretation and practice of linear solvation energy relationships in chromatography. *J. Chromatogr., A.* **2006**, *1126*, 143-194.

(19) Platts, J.A.; Abraham, M.H.; Butina, D.; Hersey, A. Estimation of molecular linear free energy relationship descriptors by a group contribution approach. 2. Prediction of partition coefficients. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 71-80.

(20) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J.W. Refinement and parametrization of COSMO-RS. *J. Phys. Chem.* **1998**, *102*, 5074-5085.

(21) Klamt, A. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* **1995**, *99*, 2224-2235.

(22) Goss, K.U. Predicting equilibrium sorption of neutral organic chemicals into various polymeric sorbents with COSMO-RS. *Anal. Chem.* **2011**, *83*, 5304-5308.

(23) Hilal, S.H.; Karickhoff, S.W.; Carreira, L.A. Prediction of the solubility, activity coefficient and liquid/liquid partition coefficient of organic compounds. *QSAR Comb. Sci.* **2004**, *23*, 709-720.

(24) Hilal, S.H.; Ayyampalayam, S.N.; Carreira, L.A. Air-liquid partition coefficient for a diverse set of organic compounds: Henry`s law constant in water and hexadecane. *Environ. Sci. Technol.* **2008**, *42*, 9231-9236.

(25) Abraham, M.H.; McGowan, J.C. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. *Chromatographia.* **1987**, *23*, 243-246.

(26) Bronner, G.; Fenner, K.; Goss, K.U. Hexadecane/air partitioning coefficients of multifunctional compounds: experimental data and modeling. *Fluid Phase Equilib.* **2010**, *299*, 207-215.

(27) Qian, J.; Poole, C.F. Distribution of neutral organic compounds between n-heptane and fluorine-containing alcohols. *J. Chromatogr., A.* **2007**, *1143*, 276-283.

(28) Ahmed, H.; Poole, C.F. Distribution of neutral organic compounds between n-heptane and methanol or N,N-dimethylformamide. *J. Sep. Sci.* **2006**, *29*, 2158-2165.

(29) Endo, S.; Droge, S.T.J.; Goss, K.U. Polyparameter linear free energy models for polyacrylate fiber–water partition coefficients to evaluate efficiency of solid-phase microextraction. *Anal. Chem.* **2011**, *83*, 1394-1400.

(30) Endo, S.; Mewburn, B.; Escher, B. Liposome and protein-water partitioning of polybrominated diphenyl ethers (PBDEs). *Chemosphere.* **2013**, *90*, 505-511.

(31) Ter Laak, T.L.; Durjava, M.; Struijs, J.; Hermens, J. Solid phase dosing and sampling technique to determine partition coefficients of hydrophobic chemicals in complex matrixes. *Environ. Sci. Technol.* **2005**, *39*, 3736-3742.

(32) Karunasekara, T.; Poole, C.F. Models for liquid-liquid partition in the system propylene carbonate-organic solvent and their use for estimating descriptors for organic compounds. *J. Chromatogr., A.* **2011**, *1218*, 809-816.

(33) Sprunger, L.; Proctor, A.; Acree, W.E.; Abraham, M.H. Characterization of the sorption of gaseous and organic solutes onto polydimethyl siloxane solid-phase microextraction surfaces using the Abraham model. *J. Chromatogr., A.* **2007**, *1175*, 162-173.

(34) Abraham, M.H.; Chadha, H.S.; Whiting, G.S.; Mitchell, R.C. Hydrogen bonding. 32. An analysis of water-octanol and water-alkane partitioning and the Δlog P parameter of Seiler. *J. Pharm. Sci.* **1994**, *83*, 1085-1100.

(35) Tülp, H.C.; Goss, K.-U.; Schwarzenbach, R.P.; Fenner, K. Experimental determination of LSER parameters for a set of 76 diverse pesticides and pharmaceuticals. *Environ. Sci. Technol.* **2008**, *42*, 2034-2040.

(36) Delle Site, A. The vapor pressure of environmentally significant organic chemicals: A review of methods and data at ambient temperatures. *J. Phys. Chem. Ref. Data.* **1997**, *26*, 157-193.

(37) Marvin, C.H.; Tomy, G.T.; Armitage, J.; Arnot, J.; McCarty, L.; Covaci, A.; Palace, V. Hexabromocyclododecane: Current understanding of chemistry, environmental fate and toxicology and implications for global management. *Environ. Sci. Technol.* **2011**, *45*, 8613-8623.

(38) Goss, K.-U.; Arp, H.P.; Bronner, G.; Niederer, C. Partition behavior of hexachlorocyclohexane isomers. *J. Chem. Eng. Data.* **2008**, *53*, 750-754.

(39) Schwarzenbach, R.P.; Gschwend, P.M.; Imboden, D.M., *Environmental Organic Chemistry*. 2. ed. 2003, Hoboken: John Wiley & Sons.

(40) Poole, C.F.; Ariyasena, T.C.; Lenca, N. Estimation of the environmental properties of compounds from chromatographic measurements and the solvation parameter model. *J. Chromatogr., A.* **2013**, *1317*, 85-104.

(41) Abraham, M.H.; Al-Hussaini, J.M. Solvation parameters for the 209 PCBs: calculation of physicochemical properties. *J. Environ. Monit.* **2005**, *7*, 295-301.

(42) Abraham, M.H.; Enomoto, K.; Clarke, E.D.; Roses, M.; Rafols, C.; Fuguet, E. Henry`s law constants or air to water coefficients for 1,3,5-triazines by an LFER method. *J. Environ. Monit.* **2007**, *9*, 234-239.

(43) Sprunger, L.M.; Achi, S.S.; Acree, W.E.; Abraham, M.H. Development of correlations for describing solute transfer into acyclic alcohol solvents based on the Abraham model and fragment-specific equation coefficients. *Fluid Phase Equilib.* **2010**, *288*, 139-144.

(44) Abraham, M.H.; Ibrahim, A.; Acree, W.E. Partition of compounds from gas to water and from gas to physiological saline at 310 K: Linear free energy relationships. *Fluid Phase Equilib.* **2007**, *251*, 93-109.

(45) Karunasekara, T.; Poole, C.F. Compounds for expanding the descriptor space for characterizing separation systems. *J. Chromatogr., A.* **2012**, *1266*, 124-130.

(46) Abraham, M.H. Hydrogen bonding XXVII. Solvation parameters for functionally substituted aromatic compounds and heterocyclic compounds, from gas-liquid chromatographic data. *J. Chromatogr.* **1993**, *644*, 95-139.

(47) Abraham, M.H.; Acree, W.E. Characterisation of the water-isopropyl myristate system. *Int. J. Pharm.* **2005**, *294*, 121-128.

(48) Abraham, M.H.; Sanchez-Moreno, R.; Gil-Lostes, J.; Acree, W.E.; Cometto-Muniz, J.; Cain, W.S. The biological and toxicological activity of gases and vapors. *Toxicol. Vitro.* **2010**, *24*, 357-362.

(49) Houser, E.J.; Mlsna, T.E.; Nguyen, V.K.; Chung, R.; Mowery, R.L.; McGill, R.A. Rational materials design of sorbent coatings for explosives: applications with chemical sensors. *Talanta.* **2001**, *54*, 469-485.

(50) Endo, S.; Watanabe, N.; Ulrich, N.; Bronner, G.; Goss, K.-U. UFZ-LSER database v 2.0 [Internet], Leipzig, Deutschland, Helmholtzzentrum für Umweltforschung – UFZ. 2013 [letzter Zugriff 17.01.2014] Erreichbar unter: http://www.ufz.de/index.php?de= 31698&contentonly=1&lserd_data[mvc]=Public/start

(51) Karunasekara, T.; Poole, C.F. Model for the partition of neutral compounds between n-heptane and formamide. *J. Sep. Sci.* **2010**, *33*, 1167-1173.

(52) Klamt, A. The COSMO and COSMO-RS solvation models. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 699-709.

(53) Klamt, A.; Eckert, F.; Arlt, W. COSMO-RS: An alternative to simulation for calculating thermodynamic properties of liquid mixtures. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 101-122.

(54) Platts, J.A.; Butina, D.; Abraham, M.H.; Hersey, A. Estimation of molecular linear free energy relation descriptors using a group contribution approach. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 835-845.

(55) Karickhoff, S.W.; V.K., M. Predicting chemical reactivity by computer. *Environ. Tox. Chem.* **1991**, *10*, 1405-1416.

(56) Goss, K.U.; Arp, H.P.H.; Bronner, G.; Niederer, C. Nonadditive effects in the partitioning behavior of various aliphatic and aromatic molecules. *Environ. Toxicol. Chem.* **2009**, *28*, 52-60.

(57) Arp, H.P.; Niederer, C.; Goss, K.-U. Predicting the Partitioning Behavior of various Highly Fluorinated Compounds. *Environ. Sci. Technol.* **2006**, *40*, 7298-7304.

1.10 Table with final pp-LFER substance descriptors

This table comprises

- the 40 descriptors from "Determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for established and alternative flame retardants",

- the 111 descriptors from "Experimental determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for pesticides and other contaminants: New measurements and recommendations",

- and 61 new descriptors.

Table 2. pp-LFER substance descriptors and internal rmse values for the 212 target compounds.

CAS	compound	S	A^a	B	V^b	L	rmse
3424-82-6	2,4'-DDE	1.38	0.00	0.21	2.05	9.09	0.06
72-55-9	4,4'-DDE	1.36	0.00	0.16	2.05	9.42	0.06
789-02-6	2,4'-DDT	1.34	0.00	0.45	2.22	9.95	0.15
50-29-3	4,4'-DDT	1.78	0.00	0.24	2.22	9.98	0.11
15972-60-8	alachlor	1.60	0.00	1.07	2.14	8.24	0.06
309-00-2	aldrin	1.03	0.00	0.26	2.01	8.69	0.01
35575-96-3	azamethiphos	1.59	0.00	1.70	1.97	10.12	0.26
2642-71-9	azinphos-ethyl	2.19	0.00	1.43	2.37	10.92	0.04
86-50-0	azinphos-methyl	2.31	0.00	1.21	2.09	10.37	0.04
131860-33-8	azoxystrobin	2.34	0.00	2.14	2.92	12.73	0.01
71626-11-4	benalaxyl	1.88	0.00	1.18	2.58	10.05	0.07
82657-04-3	bifenthrin	1.36	0.00	0.95	3.03	10.85	0.08
4824-78-6	bromophos-ethyl	1.26	0.00	0.71	2.24	9.30	0.00
2104-96-3	bromophos-methyl	1.51	0.00	0.46	1.96	8.65	0.16
69327-76-0	buprofezin	1.44	0.00	0.85	2.45	9.56	0.05
786-19-6	carbophenothion	1.62	0.00	0.94	2.36	10.12	0.23
5103-71-9	cis-chlordane	1.49	0.00	0.29	2.13	9.10	0.10
470-90-6	chlorfenvinfos 1[c]	1.56	0.00	0.98	2.33	8.91	0.10
470-90-6	chlorfenvinfos 2[c]	1.71	0.00	0.97	2.33	8.90	0.12
1897-45-6	chlorothalonil	2.08	0.00	0.31	1.52	7.58	0.08
2921-88-2	chlorpyrifos-ethyl	1.36	0.00	0.61	2.15	8.64	0.04
5598-13-0	chlorpyrifos-methyl	1.50	0.00	0.50	1.87	8.03	0.05
1861-32-1	chlorthal-dimethyl	1.48	0.00	0.67	1.92	8.58	0.06
60238-56-4	chlorthiophos 1[c]	1.91	0.00	0.68	2.37	9.59	0.11
60238-56-4	chlorthiophos 2[c]	1.65	0.00	0.87	2.37	9.85	0.13
56-72-4	coumaphos	2.44	0.00	1.01	2.43	11.36	0.01
1194-65-6	dichlobenil	1.40	0.00	0.32	1.12	5.63	0.07
97-17-6	dichlofenthion	1.25	0.00	0.74	2.07	8.25	0.14
1085-98-9	dichlofluanid	2.03	0.00	0.65	2.07	8.22	0.14

CAS	compound	S	A^a	B	V^b	L	rmse
62-73-7	dichlorvos	1.61	0.00	0.27	1.31	4.84	0.10
60-57-1	dieldrin	1.59	0.00	0.33	2.01	9.23	0.03
50563-36-5	dimethachlor	1.77	0.00	1.17	2.00	8.00	0.06
87674-68-8	dimethenamid	1.35	0.00	1.29	2.06	8.27	0.04
298-04-4	disulfoton	1.29	0.00	0.84	2.05	7.84	0.13
33213-66-0	α-endosulfan	1.39	0.00	0.65	2.08	9.22	0.01
1031-07-8	endosulfan sulfate	1.91	0.00	0.98	2.14	10.10	0.04
72-20-8	endrin	1.53	0.00	0.46	2.01	9.48	0.00
133855-98-8	epoxiconazole	1.97	0.00	1.30	2.24	10.18	0.02
66230-04-4	esfenvalerate	2.48	0.00	0.81	3.19	12.60	0.16
563-12-2	ethion	1.66	0.00	1.25	2.67	9.99	0.05
13194-48-4	ethoprophos	1.18	0.00	1.07	1.88	7.17	0.01
299-84-3	fenchlorphos	1.48	0.00	0.47	1.91	8.20	0.05
122-14-5	fenitrothion	2.17	0.00	0.51	1.86	7.98	0.10
39515-41-8	fenpropathrin	1.91	0.00	0.68	2.77	10.80	0.07
67306-00-7	fenpropidin	0.47	0.00	0.90	2.54	9.05	0.02
67306-03-0	fenpropimorph	0.57	0.00	1.14	2.74	9.31	0.06
115-90-2	fensulfothion	2.08	0.00	1.40	2.19	9.56	0.06
55-38-9	fenthion	1.73	0.00	0.67	1.99	8.35	0.23
70124-77-5	flucythrinate 1^c	2.52	0.00	0.88	3.34	12.20	0.14
70124-77-5	flucythrinate 2^c	2.58	0.00	0.85	3.34	12.27	0.15
62924-70-3	flumetralin	1.65	0.00	0.86	2.61	9.65	0.17
85509-19-9	flusilazole	2.22	0.00	0.88	2.31	9.19	0.13
944-22-9	fonofos	1.35	0.00	0.69	1.87	7.75	0.21
1024-57-3	heptachlor epoxide	1.49	0.00	0.38	1.96	8.70	0.08
23560-59-0	heptenophos	1.53	0.00	0.79	1.64	6.58	0.00
118-74-1	hexachloro-benzene	0.85	0.00	0.17	1.45	7.64	0.04
35554-44-0	imazalil	1.83	0.00	1.12	2.11	9.26	0.00
18181-70-9	iodfenphos	1.41	0.00	0.67	2.05	9.27	0.01
1634-78-2	malaoxon	1.38	0.00	1.80	2.21	8.31	0.14
121-75-5	malathion	1.84	0.00	1.18	2.32	8.32	0.09
57837-19-1	metalaxyl	1.86	0.00	1.41	2.23	8.14	0.15
67129-08-2	metazachlor	1.73	0.00	1.23	2.09	8.81	0.02
950-37-8	methidathion	2.03	0.00	1.03	1.89	8.63	0.05
72-43-5	methoxychlor	1.64	0.00	0.67	2.37	10.48	0.07
51218-45-2	metolachlor	1.46	0.00	1.27	2.28	8.75	0.00
7786-34-7	mevinphos	1.51	0.00	1.17	1.57	5.84	0.08
15299-99-7	napropamide	1.71	0.00	1.17	2.25	9.31	0.02
54-11-5	nicotin	0.86	0.00	0.98	1.37	5.95	0.03
1836-75-5	nitrofen	1.73	0.00	0.71	1.80	9.40	0.07
26530-20-1	octhilinone	1.53	0.00	1.04	1.79	7.87	0.07
34622-58-7	orbencarb	1.29	0.00	0.85	1.96	8.42	0.10
311-45-5	paraoxon-ethyl	1.33	0.00	1.29	1.89	8.53	0.26
950-35-6	paraoxon-methyl	1.48	0.00	1.31	1.61	7.83	0.25
56-38-2	parathion-ethyl	1.49	0.00	0.88	2.00	8.67	0.11
298-00-0	parathion-methyl	1.91	0.00	0.65	1.72	7.90	0.22

CAS	compound	S	A[a]	B	V[b]	L	rmse
66246-88-6	penconazole	1.82	0.00	0.69	2.05	8.76	0.07
1825-21-4	pentachloroanisole	0.96	0.00	0.35	1.53	7.63	0.05
52645-53-1	permethrin 1[c]	1.51	0.00	0.98	2.82	11.67	0.05
52645-53-1	permethrin 2[c]	1.53	0.00	0.98	2.82	11.76	0.04
298-02-2	phorate	1.19	0.00	0.80	1.91	7.41	0.18
2310-17-0	phosalone	2.32	0.00	0.91	2.40	10.64	0.06
51-03-6	piperonyl butoxide	1.42	0.00	1.22	2.73	10.43	0.09
23103-98-2	pirimicarb	1.46	0.00	1.25	1.89	7.86	0.01
32809-16-8	procymidone	1.71	0.00	0.91	1.86	9.00	0.12
41198-08-7	profenofos	1.61	0.00	0.77	2.26	9.39	0.03
2312-35-8	propargite	1.53	0.00	1.16	2.75	10.54	0.05
34643-46-4	prothiofos	1.39	0.00	0.29	2.21	9.44	0.03
13457-18-6	pyrazophos	2.01	0.00	1.25	2.61	11.10	0.02
119-12-0	pyridaphenthion	1.45	0.00	1.54	2.41	10.75	0.13
82-68-8	quintozene	1.29	0.00	0.07	1.50	7.70	0.16
3689-24-5	sulfotep	1.36	0.00	0.89	2.27	7.30	0.10
35256-85-0	tebutam	1.23	0.00	1.13	2.10	7.56	0.04
117-18-0	tecnazene	1.21	0.00	0.28	1.38	6.95	0.05
13071-79-9	terbufos	1.11	0.00	0.95	2.19	7.92	0.04
2227-13-6	tetrasul	1.25	0.00	0.31	1.98	10.08	0.22
117718-60-2	thiazopyr	1.70	0.00	0.67	2.60	8.69	0.02
640-15-3	thiometon	1.42	0.00	0.78	1.77	7.32	0.24
731-27-1	tolylfluanid	1.77	0.00	0.90	2.21	8.86	0.02
43121-43-3	triadimefon	1.75	0.00	1.28	2.15	8.57	0.05
1582-09-8	trifluralin	1.45	0.00	0.29	2.26	7.48	0.06
50471-44-8	vinclozolin	1.78	0.00	0.80	1.84	8.11	0.01
78-38-6	diethyl ethyl phosphonate	0.79	0.00	1.07	1.33	4.71	0.14
78-40-0	triethyl phosphate	0.97	0.00	1.00	1.39	4.72	0.09
1330-78-5	tricresyl phosphate 1[c]	1.63	0.00	0.94	2.79	11.59	0.08
1330-78-5	tricresyl phosphate 2[c]	2.00	0.00	0.82	2.79	11.50	0.15
1330-78-5	tricresyl phosphate 3[c]	1.71	0.00	0.99	2.79	11.80	0.11
1330-78-5	tricresyl phosphate 4[c]	2.01	0.00	1.00	2.79	11.69	0.08
115-86-6	triphenyl phosphate	1.87	0.00	0.98	2.37	10.12	0.11
636-28-2	1,2,4,5-tetrabromo-benzene	0.96	0.00	0.11	1.42	7.68	0.04
615-54-3	1,2,4-tribromobenzene	0.90	0.00	0.14	1.24	6.47	0.01
23488-38-2	2,3,5,6-tetrabromo-p-xylene	0.88	0.00	0.23	1.70	9.13	0.06
2039-82-9	4-bromostyrene	0.77	0.00	0.35	1.13	5.05	0.02
3278-89-5	allyl-2,4,6-	1.01	0.00	0.39	1.68	8.03	0.14

CAS	compound	S	A[a]	B	V[b]	L	rmse
	tribromo-phenylether						
41318-75-6	BDE 28	1.38	0.00	0.27	1.91	9.68	0.03
5436-43-1	BDE 47	1.45	0.00	0.34	2.08	10.66	0.05
60348-60-9	BDE 99	1.51	0.00	0.44	2.26	11.71	0.05
189084-64-8	BDE 100	1.48	0.00	0.41	2.26	11.48	0.06
68631-49-2	BDE 153	1.54	0.00	0.52	2.43	12.65	0.10
207122-15-4	BDE 154	1.50	0.00	0.53	2.43	12.58	0.01
207122-16-5	BDE 183	1.65	0.00	0.57	2.61	13.62	0.07
87-82-1	hexabromo-benzene	1.23	0.00	0.40	1.77	10.55	0.09
59447-55-1	pentabromobenzyl acrylate	1.80	0.00	0.49	2.19	11.16	0.08
85-22-3	pentabromo-ethylbenzene	1.05	0.00	0.27	1.87	10.27	0.11
87-83-2	pentabromo-toluene	1.13	0.00	0.31	1.73	9.99	0.17
2051-60-7	PCB 1	1.07	0.00	0.23	1.45	6.43	0.07
2051-61-8	PCB 2	1.10	0.00	0.15	1.45	6.81	0.04
2051-62-9	PCB 3	1.08	0.00	0.33	1.45	6.87	0.08
13029-08-8	PCB 4	1.19	0.00	0.12	1.57	6.90	0.14
2050-68-2	PCB 15	1.24	0.00	0.13	1.57	7.75	0.15
7012-37-5	PCB 28	1.17	0.00	0.09	1.69	8.14	0.02
15862-07-4	PCB 29	1.08	0.00	0.13	1.69	8.07	0.02
35693-99-3	PCB 52	1.23	0.00	0.13	1.81	8.38	0.04
37680-73-2	PCB 101	1.19	0.00	0.15	1.94	9.21	0.01
35065-27-1	PCB 153	1.17	0.00	0.24	2.06	10.11	0.01
35065-29-3	PCB 180	1.18	0.00	0.37	2.18	10.92	0.01
486-25-9	9-fluorenone	1.55	0.00	0.37	1.37	7.41	0.07
119-65-3	isoquinoline	1.25	0.00	0.36	1.04	5.31	0.03
91-22-5	quinoline	1.11	0.00	0.54	1.04	5.27	0.00
92-83-1	xanthene	1.17	0.00	0.37	1.42	7.16	0.10
84-74-2	di-n-butyl phthalate	1.36	0.00	0.90	2.27	8.60	0.01
131-18-0	di-n-pentyl phthalate	1.42	0.00	0.74	2.56	9.49	0.02
84-75-3	di-n-hexyl phthalate	1.50	0.00	0.78	2.84	10.40	0.04
85-69-8	butyl 2-ethylhexyl phthalate	1.25	0.00	0.89	2.84	10.22	0.00
84-61-7	dicyclohexyl phthalate	1.23	0.00	1.03	2.62	11.20	0.07
86-57-7	1-nitronaphthalene	1.49	0.00	0.46	1.26	6.79	0.12
121-14-2	2,4-dinitrotoluene	1.84	0.00	0.56	1.21	6.27	0.09
91-23-6	2-nitroanisole	1.57	0.00	0.61	1.09	5.41	0.04
555-03-3	3-nitroanisole	1.22	0.00	0.48	1.09	5.59	0.13
100-17-4	4-nitroanisole	1.47	0.00	0.46	1.09	5.74	0.13
602-60-8	9-nitroanthracene	1.73	0.00	0.53	1.63	8.98	0.15
5522-43-0	1-nitropyrene	1.91	0.00	0.45	1.76	10.85	0.07

CAS	compound	S	Aᵃ	B	Vᵇ	L	rmse
81-15-2	musk xylene	1.47	0.00	0.65	2.08	8.25	0.07
57-83-0	progesterone	1.97	0.00	1.51	2.62	12.47	0.12
90-15-3	1-naphthol	1.13	0.71	0.11	1.14	6.13	0.02
53-19-0	2,4'-DDD	1.61	0.14	0.19	2.10	9.35	0.08
115-32-2	4,4'-dicofol	1.00	0.29	1.09	2.28	10.91	0.05
834-12-8	ametryn	1.02	0.33	1.16	1.80	8.36	0.07
1912-24-9	atrazine	0.97	0.49	1.05	1.62	7.60	0.08
314-40-9	bromacil	1.41	0.40	0.92	1.63	8.23	0.15
18181-80-1	bromopropylate	1.83	0.00	1.00	2.51	10.58	0.11
1689-84-5	bromoxynil	1.04	0.57	0.49	1.28	7.24	0.12
63-25-2	carbaryl	1.54	0.00	0.93	1.54	8.03	0.22
1563-66-2	carbofuran	1.37	0.11	1.04	1.69	7.29	0.04
510-15-6	chlorobenzilate	1.58	0.17	0.92	2.27	9.67	0.04
101-21-3	chlorpropham	1.25	0.29	0.52	1.58	6.97	0.00
99-30-9	dicloran	1.66	0.27	0.36	1.24	7.10	0.02
122-39-4	diphenylamine	0.96	0.61	0.20	1.42	6.89	0.01
22224-92-6	fenamiphos	1.99	0.14	0.93	2.35	9.15	0.09
60168-88-9	fenarimol	1.57	0.50	1.39	2.29	11.09	0.01
319-84-6	α-HCH	1.31	0.15	0.50	1.58	7.35	0.05
319-85-7	β-HCH	1.52	0.40	0.51	1.58	7.47	0.07
58-89-9	γ-HCH (lindane)	1.35	0.17	0.47	1.58	7.59	0.04
319-86-8	δ-HCH	1.61	0.27	0.44	1.58	7.56	0.07
25311-71-1	isofenphos	1.67	0.00	0.95	2.64	9.02	0.15
31120-85-1	isofenphos-oxon	1.68	0.13	1.39	2.54	8.76	0.06
16752-77-5	methomyl	1.28	0.00	0.91	1.21	6.67	0.25
841-06-5	methoprotryne	0.92	0.40	1.45	2.14	9.75	0.14
21087-64-9	metribuzin	0.71	0.65	1.35	1.62	8.14	0.16
63284-71-9	nuarimol	1.92	0.18	1.32	2.21	10.01	0.05
40487-42-1	pendimethalin	1.39	0.16	0.71	2.15	9.00	0.05
527-20-8	pentachloroaniline	1.12	0.23	0.40	1.43	7.95	0.05
7287-19-6	prometryn	1.11	0.33	0.87	1.94	8.36	0.01
139-40-2	propazine	1.07	0.38	1.17	1.76	7.64	0.03
122-42-9	propham	1.01	0.27	0.68	1.45	6.17	0.01
23950-58-5	propyzamide	1.61	0.24	0.89	1.84	7.57	0.11
7286-69-3	sebuthylazine	1.14	0.40	0.94	1.76	7.92	0.03
122-34-9	simazine	0.87	0.49	1.04	1.48	7.57	0.11
886-50-0	terbutryn	1.20	0.30	0.83	1.94	8.41	0.00
13674-87-8	tris(1,3-dichloro-2-propyl) phosphate	2.10	0.03	1.24	2.55	9.93	0.08
115-96-8	tris(2-chloroethyl) phosphate	2.09	0.03	0.98	1.76	7.18	0.03
3194-55-6	1,2,5,6,9,10-hexabromo-cyclododecane	2.27	0.17	0.76	2.74	12.98	0.11
3296-90-0	2,2-bis(bromo-methyl)-1,3-propanediol	1.10	0.42	0.55	1.28	6.37	0.03
35109-60-5	2,3-dibromo-	1.67	0.00	0.58	2.07	10.35	0.08

CAS	compound	S	A[a]	B	V[b]	L	rmse
	propyl-2,4,6-tribromophenyl ether						
118-79-6	2,4,6-tribromophenol	0.92	0.52	0.27	1.30	7.06	0.02
615-58-7	2,4-dibromophenol	0.46	0.71	0.37	1.13	5.87	0.05
79-94-7	3,3',5,5'-tetrabromo-bisphenol A	1.31	1.01	0.70	2.56	12.85	0.01
626-41-5	3,5-dibromophenol	0.93	0.83	0.18	1.13	6.21	0.02
-	3-OH BDE47	1.15	0.40	1.20	2.14	11.62	0.05
79755-43-4	6-OH BDE47	1.08	0.49	1.23	2.14	11.27	0.02
115-27-5	chlorendic anhydride	0.83	0.65	0.80	1.83	8.91	0.35
38521-51-6	pentabromobenzyl bromide	1.43	0.11	0.25	1.91	10.93	0.03
608-71-9	pentabromophenol	1.02	0.64	0.59	1.65	9.94	0.02
3322-93-8	tetrabromoethyl-cyclohexane	1.53	0.00	0.36	1.83	8.89	0.10
79-95-8	tetrachloro-bisphenol A	1.73	0.76	0.32	2.35	11.09	0.06
1522-92-5	tribromoneopentyl alcohol	1.07	0.50	0.67	1.40	6.73	0.03
158076-63-2	4-hydroxy-2',3,3',5,5',6'-hexachloro-biphenyl	1.23	0.56	0.20	2.14	10.78	0.01
56558-18-0-4	4-hydroxy-2',3,4',5,6'-pentachloro-biphenyl	1.10	0.50	0.45	2.02	9.98	0.01
189578-00-5	4-hydroxy-2',3,4',6'-tetrachloro-biphenyl	0.96	0.57	0.60	1.89	9.19	0.01
86-74-8	carbazole	1.91	0.44	0.15	1.32	7.33	0.10
120-72-9	indole	1.26	0.40	0.32	0.95	5.08	0.03
298-46-4	carbamazepine	1.68	0.00	1.26	1.81	9.91	0.21
3380-34-5	triclosan	1.11	0.42	0.63	1.81	9.06	0.06

[a]Due to the non-existence of any H-bond donor functional group, A has been set to zero for the first half of the chemicals ranging from 2,4'-DDE up to progesterone.
[b]molar volume in $(cm^3/mol)/100$
[c]isomers in the order of peak appearance in the GC

2 Abstracts of original publications

2.1 Measurements and prediction of hexadecane/air partition coefficients of 387 environmentally relevant compounds

The hexadecane/air partition coefficient L is a common descriptor for non-specific interaction properties of solutes and is used in poly-parameter linear free energy relationships (pp-LFERs) to predict other partition coefficients. However, the L value data set available for complex and multifunctional substances is rather small. This limits the applicability of the pp-LFER equation. Hence, we experimentally determined L values for 387 complex compounds using GC-retention time measurements on a non-polar column (SPB™ Octyl). The target substances include environmentally relevant compounds such as pesticides, flame retardants and hormones. We determined L values that span a large range of 4.28 to 15.92. In addition to these experimental measurements several prediction tools (connectivity indices, SPARC, ABSOLV, COSMOthermX) for the L value were evaluated. The root mean squared errors (rmse) were 1.55 (connectivity indices), 1.28 (SPARC), 0.99 (ABSOLV) and 0.94 (COSMOthermX). The number of outliers (prediction error > 3) was 18 (connectivity indices), 12 (SPARC), 2 (ABSOLV) and 0 (COSMOthermX). Based on these results the best prediction accuracy in this evaluation is reached by ABSOLV and COSMOthermX, whose results are comparable.

2.2 Determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for established and alternative flame retardants

Poly-parameter linear free energy relationships (pp-LFERs) can predict partition coefficients for a multitude of environmental and biological phases with high accuracy. In this work, the pp-LFER substance descriptors of 40 established and alternative flame retardants (e.g., polybrominated diphenyl ethers, hexabromocyclododecane, bromobenzenes, trialkyl phosphates) were determined experimentally. In total, 251 data for gas-chromatographic (GC) retention times and liquid/liquid partition coefficients (K) were measured and used to calibrate the pp-LFER substance descriptors. Substance descriptors were validated through a comparison between predicted and experimental log K for the systems octanol/water (K_{ow}), water/air (K_{wa}), organic carbon/water (K_{oc}) and liposome/water (K_{lipw}), revealing a high reliability of pp-LFER predictions based on our descriptors. For instance, the difference between predicted and experimental log K_{ow} was < 0.3 log units for 17 out of 21 compounds for which experimental values were available. Moreover, we found an indication that the H-bond acceptor value (B) depends on the solvent for some compounds. Thus, for predicting environmentally relevant partition coefficients it is important to determine B values using measurements in aqueous systems. The pp-LFER descriptors calibrated in this study can be used to predict partition coefficients for which experimental data are unavailable, and the predicted values can serve as references for further experimental measurements.

2.3 Experimental determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for pesticides and other contaminants: New measurements and recommendations

Well-calibrated poly-parameter linear free energy relationships (pp-LFERs) are an accurate way to predict partition coefficients (K) for neutral organic chemicals. In this work, pp-LFER substance descriptors of 111 environmentally relevant substances, mainly pesticides, were determined experimentally using gas chromatographic (GC) retention times and liquid/liquid partition coefficients. The complete set of descriptors for 50 compounds are being reported here for the first time. Validation of the measured substance descriptors was done by comparing predicted and experimental log K for the systems octanol/water (K_{ow}), water/air (K_{wa}), and organic carbon/water (K_{oc}), all of which indicated a high reliability of pp-LFER predictions based on the determined descriptors (e.g., a root mean squared error of 0.39 for log K_{ow}). The descriptors presented in this work in combination with existing pp-LFER system equations substantially extend (and in some cases correct) our knowledge on partition properties of these 111 chemicals. In addition, the results of this work provide insight on some general guidelines with respect to the method combination best suited for deriving descriptors for environmentally relevant compounds.

2.4 Prediction of partition coefficients for complex environmental contaminants: Validation of COSMOtherm, ABSOLV, and SPARC

Prediction of partition coefficients is essential for screening of environmentally relevant compounds. Prediction methods using only the molecular structure as input are especially useful for this purpose. In the present study, we validated three prediction methods, COSMOtherm, ABSOLV and SPARC, which are based on more mechanistic approaches than most other quantitative structure activity relationships. Validation is based on a consistent experimental data set of up to 270 compounds, mostly pesticides and flame retardants. The validation systems include three gas chromatographic (GC) columns and four liquid/liquid systems that represent all relevant types of intermolecular interactions. Results revealed that the overall prediction accuracy of COSMOtherm and ABSOLV is comparable, whereas SPARC performance is substantially lower than the other two. For instance, the root mean squared error (rmse) for the four liquid/liquid partition coefficients was 0.65 to 0.93 log units for COSMOtherm, 0.64 to 0.95 log units for ABSOLV, and 1.43 to 2.85 log units for SPARC. In addition, version and parameterization influences of COSMOtherm on the prediction accuracy were determined.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Weiterhin erkläre ich, dass ich noch keine vergeblichen Promotionsversuche unternommen habe und die Dissertation weder in der gegenwärtigen noch in einer anderen Fassung bereits einer anderen Fakultät vorgelegen hat.

Leipzig, 24. Januar 2014

Angelika Stenzel

Angaben zur Person und zum Bildungsgang

Name	Angelika Stenzel
Geburtsdatum	26.09.1985
Geburtsort	Köthen

Seit 12/2010 **Helmholtzzentrum für Umweltforschung GmbH – UFZ, Leipzig**

Department Analytische Umweltchemie
Doktorand

10/2008 – 09/2010 **Masterstudiengang Chemie, Universität Leipzig**

- Abschluss: Master of Science
- Thema der Masterarbeit: Identifizierung von Phytochelatin-Komplexen in Pflanzen mittels HPLC/ICP-MS und HPLC/ESI-MS

10/2005 – 09/2008 **Bachelorstudiengang Chemie, Universität Leipzig**

- Abschluss: Bachelor of Science
- Thema der Bachelorarbeit: Einrichten einer Profilanalytik für HPLC-MS

09/1996 – 07/2005 **Ludwigsgymnasium, Köthen**

- Abschluss: Abitur

Publikationsliste

Veröffentlichungen

- A. Stenzel, K. U. Goss, S. Endo: Prediction of partition coefficients for complex environmental contaminants: Validation of COSMOtherm, ABSOLV, and SPARC, Environ. Toxicol. Chem. 33 (2014), 1537-1543
- A. Stenzel, K. U. Goss, S. Endo: Experimental determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for pesticides and other contaminants: New measurements and recommendations, Environ. Sci. Technol. 47 (2013), 14204-14214
- A. Stenzel, K. U. Goss, S. Endo: Determination of poly-parameter linear free energy relationship (pp-LFER) substance descriptors for established and alternative flame retardants, Environ. Sci. Technol. 47 (2013), 1399-1406
- A. Stenzel, S. Endo, K. U. Goss: Measurements and prediction of hexadecane/air partition coefficients of 387 environmentally relevant compounds, J. Chromatogr. A 1220 (2012), 132-142

Konferenzbeiträge

12. - 16.05.2013   SETAC Europe, Glasgow

- A. Stenzel, S. Endo, K. U. Goss: Partition coefficients of flame retardants: Applications of poly-parameter linear free energy relationships (pp-LFER), Vortrag

10. - 13.09.2012   SETAC GLB, Leipzig

- A. Stenzel, S. Endo, K. U. Goss: Gleichgewichtsverteilungen von Flammschutzmitteln: Anwendungen der Poly-Parameter linearen freien Energie Beziehung (pp-LFER), Vortrag

20. - 24.05.2012   SETAC Europe, Berlin

- A. Stenzel, S. Endo, K. U. Goss: Evaluation of ABSOLV for predicting pp-LFER descriptors of complex and multifunctional organic compounds, Poster

Danksagung

Mein Dank gilt meinem Betreuer Prof. Kai-Uwe Goss für die stets hervorragende Betreuung, für sein Engagement und seine Unterstützung bei der Erstellung dieser Arbeit.

Außerdem danke ich meinen Kollegen: Satoshi Endo für seinen Rat, Anleitung, Kommentare und Korrekturen, Andrea Pfennigsdorff für ihre Hilfe bei den Messungen im Labor, sowie Martina Schneider, Anett Krause, Trevor Brown, Luise Oemisch, Kai Bittermann, Wolfgang Larisch und Lukas Linden für ihre Hilfsbereitschaft, Motivation und die gute Zusammenarbeit.

Auch meinen Eltern möchte ich dafür danken, dass sie mich während dieser Arbeit immer unterstützt haben.