Molecular Diversity (2006) 10: 133–145
DOI: 10.1007/s11030-005-9003-3

Full-length paper

Modeling the bioconcentration factors and bioaccumulation factors of polychlorinated biphenyls with posetic quantitative super-structure/activity relationships (QSSAR)

Teodora Ivanciuc[1,*], Ovidiu Ivanciuc[2] & Douglas J. Klein[1,*]
[1]*Department of Marine Sciences, Texas A&M University, Galveston, Texas 77551, USA;* [2]*Department of Human Biology Chemistry and Genetics, University of Texas Medical Branch, Galveston, Texas 77555-0857, USA*
*(*Author for correspondence, E-mails: ivanciuc@netscape.net, kleind@tamug.edu, Tel: 409-740-4512)*

Received 1 September 2005; Accepted 19 October 2005

Key words: quantitative structure-activity relationships, QSAR, polychlorinated biphenyls, PCBs, reaction network, Hasse diagram, partially ordered set, poset, bioconcentration factor, BCF, bioaccumulation factor, BAF, cluster expansion, splinoid poset, quantitative super-structure-activity relationships, QSSAR

Summary

During bioconcentration, chemical pollutants from water are absorbed by aquatic animals *via* the skin or a respiratory surface, while the entry routes of chemicals during bioaccumulation are both directly from the environment (skin or a respiratory surface) and indirectly from food. The bioconcentration factor (BCF) and the bioaccumulation factor (BAF) for a particular chemical compound are defined as the ratio of the concentration of a chemical inside an organism to the concentration in the surrounding environment. Because the experimental determination of BAF and BCF is time-consuming and expensive, it is efficacious to develop models to provide reliable activity predictions for a large number of chemical compounds. Polychlorinated biphenyls (PCBs) released from industrial activities are persistent pollutants of the environment thereby producing widespread contamination of water and soil. PCBs can bioaccumulate in the food chain, constituting a potential source of exposure for the general population. To predict the bioconcentration and bioaccumulation factors for PCBs we make use of the biphenyl substitution-reaction network for the sequential substitution of H-atoms by Cl-atoms. Each PCB structure then occurs as a node of this reaction network, which is some sort of super-structure, turning out mathematically to be a partially ordered set (poset). Rather than dealing with the molecular structure via ordinary QSAR we use only this poset, making different quantitative super-structure/activity relationships (QSSAR). Thence we developed cluster expansion and splinoid QSSAR for PCB bioconcentration and bioaccumulation factors. The predictive ability of the BAF and BCF models generated for 20 data sets (representing different conditions and fish species) was evaluated with the leave-one-out cross-validation, which shows that the splinoid QSSAR (r between 0.903 and 0.935) are better than models computed with the cluster expansion (r between 0.745 and 0.887). The splinoid QSSAR models for BAF and BCF yield predictions for the missing PCBs in the investigated data sets.

Introduction

Many environmental pollutants, such as dichlorodiphenyltrichloroethane (DDT), hexachlorobenzene (HCB), dieldrin (HEOD), and polychlorinated biphenyls (PCBs), move through food chains and accumulate at sizeable levels in the tissues of animals and man [1–6]. Chemicals that have both a high lipophilicity and a high environmental persistence should be thoroughly investigated for their potential toxicity through bioconcentration and bioaccumulation, both measured for long periods of exposure. The bioconcentration and bioaccumulation of chemical compounds in aquatic and terrestrial organisms represent important criteria for ecotoxicological evaluation and hazard assessment [7–15].

In order to determine the environmental fate of chemicals released from industrial, agricultural, or residential sources it is essential to determine their bioconcentration in aquatic species. The bioconcentration factor (BCF) of a chemical compound is defined as the ratio between the concentration of that chemical in an organism (or in the fat, or in a certain tissue of the organism) and the concentration of the chemical in the aqueous environment [16–23]:

$$\mathrm{BCF} = \frac{C_{\mathrm{org}}}{C_{\mathrm{m}}}$$

where C_{org} is the concentration of the chemical in the organism (or tissue), C_m is the concentration of the chemical in the aqueous environment in which the respective organism lives, and both concentrations are measured after long-term exposure until steady state is reached. The calculation of BCFs can be based on the wet weight, BCF_w, or on the lipid content, BCF_L, of the aquatic organism or its tissue. BCFs are usually determined for various species of fish, but other aquatic organisms can be used, such as algae or mussels.

Aquatic organisms can accumulate chemical compounds both directly from the environment (*via* skin or respiratory surface) and indirectly (by collecting and concentrating a chemical compound from food). This process is called bioaccumulation, and is measured with the bioaccumulation factor (BAF) which is defined as the ratio of the concentration of a chemical accumulated inside an organism (from food and direct exposure) to the concentration in the surrounding environment [2, 3, 13, 16]:

$$\mathrm{BAF} = \frac{C_{org}}{C_m}$$

The bioaccumulation level depends on the nature of the chemical compound, species, duration of exposure, concentration in water and its accumulation level in food. BAF can be particularly high for lipophilic compounds highly soluble in the lipid fraction of the organism, and with a low or negligible metabolism rate in the organism (which makes them very persistent). Even when the water concentration of such chemicals is too low to cause health problems from drinking the water, their high bioaccumulation may pose risks for those eating fish or shellfish.

BCF values depend not only on the chemical structure but also on the level of environmental exposure, on the species, and on various characteristics of the aquatic organism (age, fat content, duration of exposure to the chemical). Because the experimental determination of BCF is time-consuming and expensive, various quantitative structure-activity relationships (QSAR) models have been investigated for the BCF prediction from octanol–water partition coefficients or from the structural characteristics of molecules [24–47].

Posets (or partially ordered sets) have been proposed [48–50] as of fundamental chemical utility, and reaction-networks were [51] recognized as posets, which might be used in identifying (super-structural) regularities in the properties of the compounds appearing in the network. The consequent reaction network poset diagrams capture the structural dependences of molecular properties and can be used to develop quantitative super-structure-activity relationship (QSSAR) models specifically tailored for networks of chemicals derived by substitution from a parent skeleton. Brüggemann and co-workers [52–57] have advocated the use of poset relationships in the form of Hasse diagrams as an attractive way of handling complex information within the environmental area. Sørensen, Carlsen and co-workers used the relationships induced by Hasse diagrams to rank chemical compounds according to their environmental effects [58–61]. QSSAR formulations based on substitution reaction networks [62–68] give reliable models for various physico-chemical and biological properties [63, 64]. We have used the poset reaction diagram of chlorobenzenes to model various toxicity indices for guppy, fathead minnow, brine shrimp, *Daphnia magna*, algae, and tadpoles [64]. In the same study we found also that fish bioconcentration factors can be accurately predicted with poset QSSAR models. Based on these encouraging results, we intend to extend the application of QSSAR models to other classes of chemicals.

The objective of this study is to establish BAF and BCF predicting models for PCBs based on the poset (partially ordered set) reaction diagrams using our original QSSAR methods [56–58]. A BCF data set was aggregated from various literature reports [39–46] while the 19 BAF data sets were taken from the comprehensive study of Burkhard and co-workers [47] who measured bioaccumulation factors data for various fish species (carp, alewife, shad, walleye, smelt, and yellow perch) in the Green Bay area and in the Hudson River.

Data and procedures

Experimental data

The release of polychlorinated biphenyls (PCBs) from industrial sources and their persistence in the environment have resulted in widespread contamination of water and soil, with subsequent potential exposure of the general population. Due to their lipophilicity, PCBs from food and from the aqueous environment accumulate in the fatty tissues of fish and shellfish [6–9]. PCBs have a low metabolism rate in the aquatic species, which makes them very persistent. Even low levels of PCBs in water can result in significant bioaccumulation in the food chain, due to their very slow degradation and lipophilicity. Thus, even small concentrations in PCBs in rivers, lakes, seas and oceans can result in significant bioaccumulation in fish and shellfish, which may pose human health risks from their consumption. Due to their importance as environmental pollutants that can produce serious risks for human health, we developed QSSAR models for BAFs and BCFs of PCBs. All experimental data were collected from the literature.

The BCF data set, for 58 PCBs, was aggregated from various literature reports [39–46]. We consolidated experimental log BCF values for several fish species (guppies, fathead minnow, rainbow trout, and bluegill sunfish) because there are too few data for each individual species. Our assumption is that the PCB accumulation and metabolism mechanisms in all these species are similar, and the bioaccumulation factors depend mainly on the PCBs molecular structure and more particularly on the placement of each PCB in the biphenyl-chlorination reaction network described in the next subsections.

Table 1. Origin of experimental data and leave-one-out cross-validation statistics (correlation coefficient *r* and standard deviation *s*) for cluster-expansion and splinoid poset QSSAR for polychlorinated biphenyls BCF and BAF.

			Cluster-expansion		Splinoid poset		
No.	Origin	Property	*r*	*s*	*r*	*s*	# PCBs
1		BCF to several species (guppies, fathead minnow, rainbow trout, bluegill sunfish)	0.908	0.372	0.958	0.245	58
2	Zone 1 Green Bay	BAF to alewife YOY	0.828	0.396	0.921	0.237	45
3		BAF to alewife adult	0.884	0.422	0.922	0.331	45
4		BAF to carp 8 yrs	0.840	0.470	0.908	0.336	50
5		BAF to smelt YOY	0.816	0.420	0.924	0.241	45
6		BAF to smelt adult	0.860	0.396	0.924	0.259	46
7	Zone 2 Green Bay	BAF to alewife YOY	0.826	0.387	0.903	0.258	47
8		BAF to alewife adult	0.866	0.453	0.932	0.301	49
9		BAF to carp 3 yrs	0.745	0.495	0.911	0.275	49
10		BAF to carp 10 yrs	0.814	0.479	0.921	0.296	51
11		BAF to carp 12 yrs	0.840	0.471	0.926	0.305	49
12		BAF to shad YOY	0.762	0.454	0.904	0.263	48
13		BAF to smelt YOY	0.862	0.382	0.916	0.269	46
14		BAF to smelt adult	0.868	0.400	0.935	0.258	46
15		BAF to walleye 1 yrs	0.823	0.467	0.914	0.309	53
16		BAF to walleye 3 yrs	0.857	0.428	0.923	0.292	51
17		BAF to walleye 4 yrs	0.844	0.462	0.927	0.297	52
18	Zone 3 Green Bay	BAF to alewife adult	0.887	0.421	0.903	0.341	44
19		BAF to smelt YOY	0.879	0.398	0.905	0.301	40
20	Zone 4 Hudson R.	BAF to yellow perch	0.867	0.401	0.911	0.261	31

Dataset 1 from ref. [39–46] and datasets 2–20 from ref. [47].

We collected 19 BAF data sets from the comprehensive study of Burkhard and co-workers [47]. These data sets represent bioaccumulation factors for various fish species (carp, alewife, shad, walleye, smelt, and yellow perch) in the Green Bay area and in the Hudson River (see Table 1 for conditions and fish origin). The bioaccumulation factors reported by Burkhard and co-workers [47] represent log BAF_L^{fd} values, which are based upon concentrations of freely dissolved chemical in the ambient water and concentrations of the chemical in the lipid fraction of the organism.

Partially ordered sets

Formally, a *partially ordered set* (*poset*) consists of a set P with a (partial ordering) relation $\succ$ which satisfies two conditions: first, for $\alpha, \beta \in P$, $\alpha \succ \beta \Rightarrow \beta \not\succ \alpha$; and second for $\alpha, \beta, \gamma \in P$, $\alpha \succ \beta$ and $\beta \succ \gamma \Rightarrow \alpha \succ \gamma$. Here our set P consists of chemical compounds and the ordering $\alpha \succ \beta$ is to mean that β is obtainable from α after some (non-zero) number of chlorinations. The relation which allows either $\alpha \succ \beta$ or $\alpha = \beta$ is denoted $\alpha \succcurlyeq \beta$. The relation where $\alpha \succ \beta$ without any intervening members of P is denoted $\alpha \rightarrow \beta$, and in mathematical language one says α *covers* β. The poset P may be represented diagrammatically with each member of P represented by a node and each covering relation represented by an arrow, all organized so that the arrows all have a component in the "downward" direction. For our current case this (Hasse) diagram is simply the reaction network.

It has been emphasized that poset diagrams could be used to represent a range of chemical reaction networks that take place by a progressive chain of substitutions on a fixed molecular skeleton [51]. These hierarchical reaction diagrams encode structural information that can be used to predict various physico-chemical and biological properties of the chemicals that form the diagram.

Poset diagrams can also be used in other ways to organize information related to various environmental systems, as shown by Brüggemann and co-workers [52–57] in studies that investigate their use in the evaluation of toxicological

fish tests, in ordering environmental pollutants, and in ecosystem comparison. Poset diagrams based on molecular structure properties were used by Sørensen, Carlsen and co-workers to rank chemical compounds according to their environmental effects [58–61].

The super-structural poset diagrams based on the hierarchical network of substitution reactions have been used [48–51] to develop quantitative super-structure-activity relationship (QSSAR) models specifically tailored for chemical structures derived by substitution from a parent skeleton. Various mathematical fitting schemes can be developed from these poset diagrams – such including cluster expansion, average poset, or splinoid poset methods. The simple "average-poset" model was applied [63] for a wide range of properties of chloro- and methyl-benzenes, and more recently we applied all 3 of these QSSAR models for the poset diagram of chlorobenzenes to model various toxicity indices for several aquatic species [64].

The poset diagram for polychlorinated biphenyls

In this section we describe the poset diagram of polychlorinated biphenyls, which represents the basis for the QSSAR model. The QSSAR procedure is a predictive scheme based on similarity comparisons to the corresponding activities of related structures, in which the property predictions are made based on the posetic (reaction-network) "super-structure". This super-structure considered here is neatly represented by the diagrammatic substitution–reaction network, hierarchically organized following the substitution level: it starts with the unsubstituted skeleton, and ends with the completely substituted compound, having intermediate lavels of compounds with the same substitution degree.

The first three and last three layers of compounds from the poset diagram of PCBs are shown in Figure 1. The PCB poset diagram starts with biphenyl at the top and ends with decachlorobiphenyl at the bottom, while all the different patterns of substitution occur in between. The arrows indicate the hierarchic generation of the different patterns of more substituted compounds from the different patterns of less substituted ones. From this diagram it is easy to recognize that from biphenyl one may obtain three monochloro-biphenyl congeners, namely 2-chlorobiphenyl, 3-chlorobiphenyl and 4-chlorobiphenyl, respectively. Similarly, 2,3-dichlorobiphenyl has incoming connections from 2-chlorobiphenyl and 3-chlorobiphenyl, but not from 4-chlorobiphenyl, thus indicating the two distinct pathways for its generation from less substituted PCBs. The complete poset reaction network for PCBs has 210 vertices and 840 edges.

In Figure 2 we show a condensed notation for a group of PCBs that have the same number of substitutents for each benzene ring. The notation from the top line shows the substitution pattern (one Cl atom in the right ring and none in the left ring) and the number of PCBs having this substitution pattern, while the bottom line shows the corresponding three monosubstituted PCBs. Using the notation exemplified in Figure 2, we show in Figure 3 the condensed biphenyl substitution reaction poset.

QSSAR models

In this section we present the mathematical basis for the QSSAR models used in this study, namely the splinoid poset and cluster-expansion models. These procedures are general and can be applied for any network of chemical compounds that can be included in a formal hierarchical reaction network. An earlier "average-poset" model [63] was not used, since it requires a denser network of known activity values than currently available for the PCBs.

Splinoid poset model

The chloro-substitution network of biphenyl is represented here as a Hasse diagram $H(P)$ (Figures 1 and 3) which mathematically represents a finite poset P. An oriented edge in the Hasse diagram here represents the transition between a chemical compound α with n chlorine atoms to one β with $n+1$ chlorine atoms, and is denoted by $\alpha \to \beta$, and we attach a real variable $x_{\alpha\to\beta}$ ranging from 0 to 1, that represents the transformation of α into β. When formulating the splinoid QSSAR model [65] for a property X, we consider a cubic spline polynomial on the oriented edges $\alpha \to \beta$ of the Hasse diagram $H(P)$:

$$f_{\alpha\to\beta}(x_{\alpha\to\beta}) = a_{\alpha\to\beta}x^3_{\alpha\to\beta} + b_{\alpha\to\beta}x^2_{\alpha\to\beta} + c_{\alpha\to\beta}x_{\alpha\to\beta} + d_{\alpha\to\beta}$$

with $a_{\alpha\to\beta}$, $b_{\alpha\to\beta}$, $c_{\alpha\to\beta}$, and $d_{\alpha\to\beta}$ are constants. Each vertex α of $H(P)$ or P is identified by a value a_α and a slope b_α. The splinoid poset QSSAR model is generated based on known values of the property X for a subset of the chemical compounds, namely for vertices $\alpha \in K \subseteq P$. The splinoid fit is such that: first, the cubic splines match values a_α at the nodes $\alpha \in K$ to the known property values; second, the incoming and outgoing slopes through each node match the corresponding b_α value; and third, a relevant total "curvature" of the overall spline is minimized (subject to the constraints of the first two conditions). With the splinoid QSSAR determined for the vertices from K, one can predict the property values for the remaining chemical compounds that do not have an experimental value for the property X, compounds that form the set U of vertices $\alpha \notin K$. An algorithm results for predicting the values of X for the set U of chemical compounds. Let $\mathbf{A}$ denote the adjacency matrix of the Hasse diagram $H(P)$, and let $\mathbf{S}$ denote the oriented adjacency matrix of $H(P)$, where:

$$S_{\alpha\beta} = \begin{cases} 1 & \text{if } \beta \to \alpha \\ -1 & \text{if } \alpha \to \beta \\ 0 & \text{otherwise} \end{cases}$$

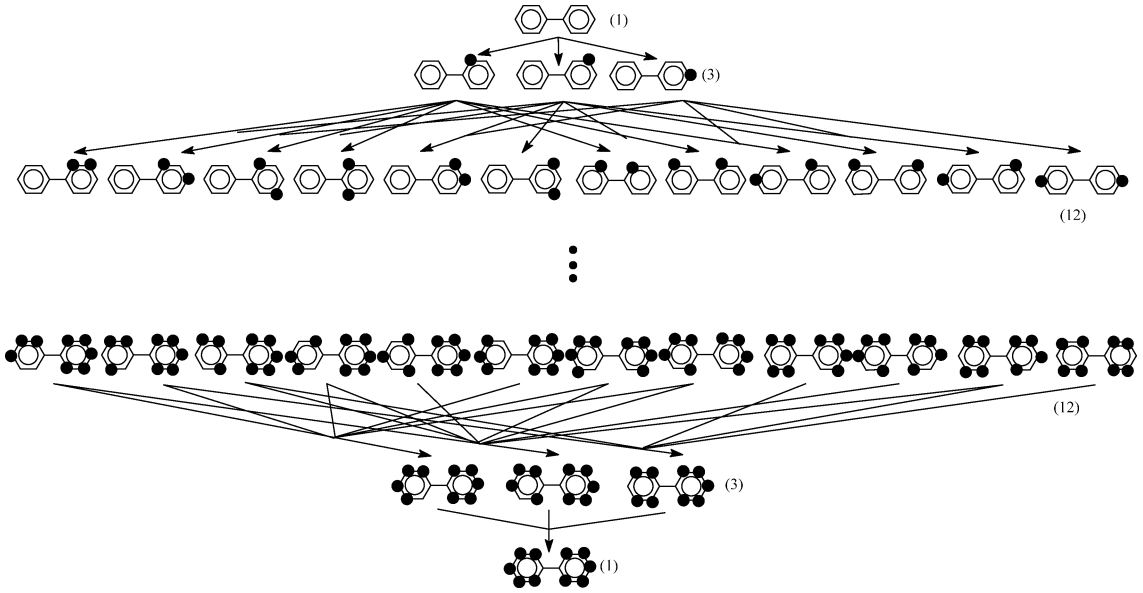


Figure 1. The biphenyl substitution reaction poset. The black enlarged dots indicate the sites on which an H atom of biphenyl has been replaced by a Cl atom. The number of each isomer on each substitution layer is given in parentheses.

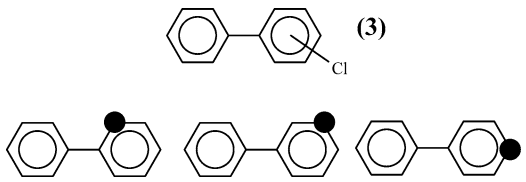


Figure 2. The condensed notation for the PCB congeners exemplified for monochlorobiphenyls.

The in-degree on vertex $\alpha \in P$ is denoted by $d_{\rightarrow\alpha}$, and the out-degree on vertex $\alpha \in P$ is denoted by $d_{\alpha\rightarrow}$. Based on this notation, we introduce the following two diagonal matrices:

$$\mathbf{D} = \mathrm{diag}[d_{\alpha\rightarrow} - d_{\rightarrow\alpha}]$$
$$\mathbf{\Delta} = \mathrm{diag}[d_{\alpha\rightarrow} + d_{\rightarrow\alpha}]$$

We define the matrices U (the $|\ U\ | \times |P|$ submatrix of the unity matrix $\mathbf{I}$, with rows indexed by the elements of U), and $\mathbf{K}$ (the $|K| \times |P|$ submatrix of the unity matrix $\mathbf{I}$, with rows indexed by the elements of K), and the derived matrix:

$$\mathbf{M} = 2(\mathbf{\Delta} - \mathbf{A}) - 3(\mathbf{D} - \mathbf{S})(\mathbf{A} + 2\mathbf{\Delta})^{-1}(\mathbf{D} + \mathbf{S})$$

The (column) vector of known property values is denoted by $\vec{k}$. Then, the vector $\vec{u}$ that contains the unknown property values a_α is computed from:

$$\vec{u} = -(\mathbf{UMU}^t)^{-1}(\mathbf{UMK}^t)\vec{k}$$

For a few different reaction networks we have studied the matrix $\mathbf{UMU}^t$ which appears in practice to be invertible regardless of how sparse the "known" data are in the network up to the point that very few ($\leq$2) known data are available. The a, b, c, d coefficients appearing in the spline polynomials f do not explicitly appear in our splinoid formula for $\vec{u}$, but they are complicit in the derivation of this formula for $\vec{u}$. The present formula gives $\vec{u}$ in terms of the poset, and thence completes the splinoid QSSAR algorithm, which turns out to give a robust model in accommodating a diversity of missing values for several compounds (which may possibly even be adjacent). This is a significant advantage of the splinoid model, which uses the topology of the Hasse diagram to generate a response web for the investigated property.

Cluster expansion model

Formal cluster expansions in general re-express a scalar function (or property) for the different members of a poset in terms of related (transformed) functions focusing more strongly on earlier members of the poset. Chemical application in the case that the partial ordering is based on a poset of molecular graphs is described in [62, 66–68]. Generally for a scalar property X defined on the members of a poset P (with partial ordering $\succ$) one may expand X for $\alpha \in P$, as

$$X(\alpha) = \sum_{\beta}^{\succ\alpha \text{ or } =\alpha} f(\beta, \alpha) X_f(\beta)$$

where the sum goes over all $\beta \succcurlyeq \alpha$, $f(\beta,\alpha)$ is a *cluster* function that maps pairs of members of P onto real numbers with $f(\beta, \alpha) = 0$ whenever $\beta \not\succcurlyeq \alpha$, and is such that $f(\alpha,\alpha) \neq 0$. Further $X_f(\beta)$ is an *f transform* property which is obtained by

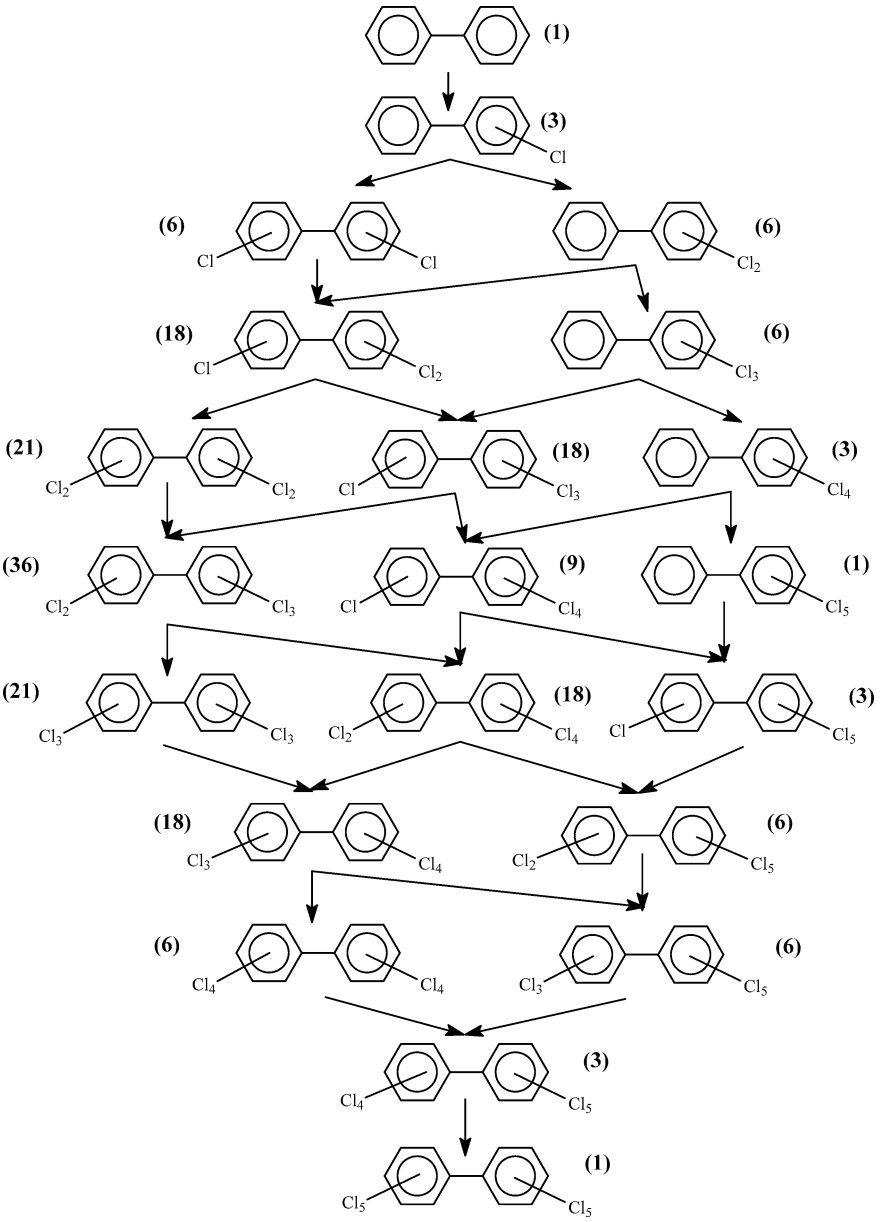


Figure 3. The condensed schematic biphenyl substitution reaction poset.

some sort of fitting procedure and which depends on X and the cluster function f. Conveniently, this cluster- expansion may be truncated to a limited sequence of non-zero cluster terms X_f, and so applied when the earlier terms alone offer a good approximation for the property X.

For our reaction-network posets we choose [64, 68] that $f(\beta,\alpha)$ be the number of ways in which the substitution pattern α occurs as a subset of substitution pattern β. For the poset diagram of polychlorinated biphenyls, we have truncated the cluster-expansion model to X_f contributions from the chlorine atoms situated through the second row of the poset (in Figure 1) as well as some of the terms in the third row. To keep the number of parameters (*i.e.*, the $X_f(\beta)$) down we choose to retain only 6 of the 12 $X_f(\beta)$ from the third row as non-zero, and further we assume some equalities, so that there remain just 3 independent parameters from this row:

$$\begin{aligned} X_f(2,3\text{-}Cl_2\phi\phi) &= X_f(3,4\text{-}Cl_2\phi\phi) = X_f(4,5\text{-}Cl_2\phi\phi) \\ &= X_f(5,6\text{-}Cl_2\phi\phi) \equiv d \\ X_f(2,4\text{-}Cl_2\phi\phi) &= X_f(3,5\text{-}Cl_2\phi\phi) = X_f(4,6\text{-}Cl_2\phi\phi) \equiv e \\ X_f(2,5\text{-}Cl_2\phi\phi) &= X_f(3,6\text{-}Cl_2\phi\phi) \equiv f \end{aligned}$$

(where ϕ indicates phenyl and $\phi\phi$ biphenyl). These equalities are rationalized in that the effects of two Cl ligands on the same benzene ring are reasonably imagined to have similar effects on the electronic structure (of the ring and of the

ligands) whenever the (topological) distance between the two Cl atoms is the same. For a single isolated benzene ring this is certainly true, and has been so quite successfully used [64], and a rather similar treatment of another ring is found in [68]. The parameters associated to the second row of the poset are abbreviated to

$$X_f(2\text{-Cl}\phi\phi) \equiv a,\ X_f(3\text{-Cl}\phi\phi) \equiv b,\ X_f(4\text{-Cl}\phi\phi) \equiv c.$$

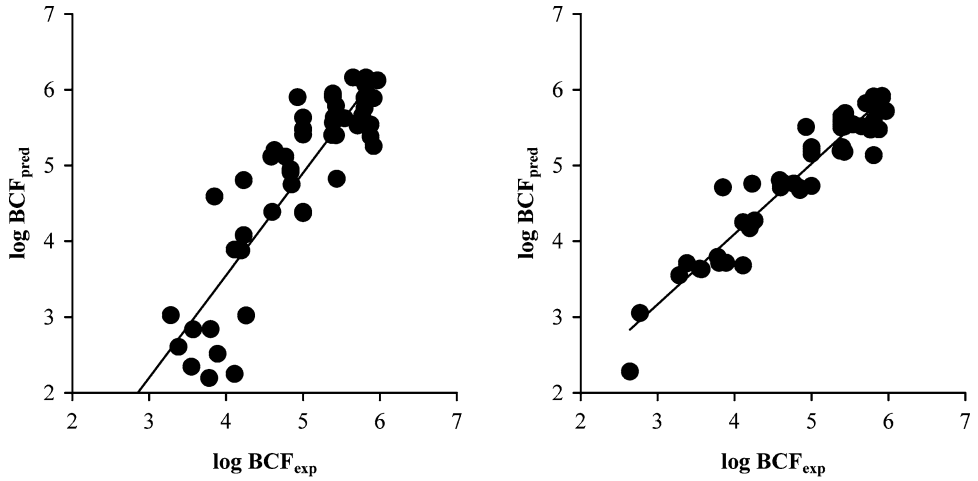
For a single isolated ring these would be the same, but here we retain their distinction, hoping that this will ameliorate the presumption of the equalities for the dichloro species. The X_f for dichloro species with one Cl in each benzene ring are taken $= 0$, so presuming that the effects of substitution in the two rings are independent, and accounted for by the parameters already retained. In each series of QSSAR models, biphenyl was considered as a reference structure, namely the property values are shifted so that $X(\phi\phi) = 0$ in which case $X_f(\phi\phi) \equiv 0$ (or alternatively one need not so shift the various $X(\alpha)$ values, but simply take $X_f(\phi\phi) = X(\phi\phi)$. The set of $X_f(\beta)$ parameters (a, b, c, d, e, f) can be computed [68] by a least-squares procedure based on a subset of molecules, or by "inversion" from small systems - and here we use the former choice.

Results and discussion

We applied the splinoid poset and cluster-expansion QSSAR for the modeling of bioconcentration factors and bioaccumulation factors of PCBs for various fish species. Particularly the splinoid-poset scheme is not much like a statistical fit, but rather is like an exact numerical interpolation – and out of a whole (infinite) ensemble of conceivable exact piecewise polynomial fits, the splinoid criterion selects that which is least "curved". To generate some sort of statistics for the splinoid scheme we have made successive fits leaving each one of the known values out and comparing its predicted value with its actually known value. The cluster-expansion procedure is more conventional in generating its own statistical indicators. But to make comparisons between the splinoid and cluster-expansion schemes, we treat (and report) both model's statistics generated by the same leave-one-out cross-validation procedure. The prediction statistics for all 20 data sets are collected in Table 1. In the last column of Table 1 we report the number of PCBs that were used to develop each QSSAR model.

The first model from Table 1 was obtained for bioconcentration factors of 58 PCBs collected from the literature [39–46]. This data set uses BCFs for several fish species, namely guppies, fathead minnow, rainbow trout, and bluegill sunfish. This aggregation of experimental BCF from different fish species is hopefully justified since the PCB accumulation and metabolism in all these species are similar, and the BCFs depend mainly on the PCBs' molecular structure. Also, the BCFs are determined in the laboratory, in a controlled environment and standardized conditions, which can further justify the aggregation of data for different fish species. Both cluster-expansion ($r_{cv} = 0.908$) and splinoid poset ($r_{cv} = 0.958$) predictions are good, with better results for the splinoid QSSAR. The plots of these two predictions are presented in Figure 4, showing that there are no significant outliers.

The remaining 19 QSSAR models reported in Table 1 were obtained with the PCB bioaccumulation factors determined for various fish species (carp, alewife, shad, walleye, smelt, and yellow perch) in the Green Bay area and in the Hudson River [47]. BAFs were measured in the natural ecosystems, which can add noise and other errors in the measurements. As a consequence, we can expect that the BAF models will be of somewhat lower quality than the QSSAR model obtained for BCF. There seems to be more variation of BAFs between species and even between age groups within a species, so that we retain 19 distinct data sets each of which



a) cluster-expansion: $r = 0.908$, $s = 0.372$ b) splinoid poset: $r = 0.959$, $s = 0.245$

Figure 4. Plot of experimental *vs.* predicted PCB log BCF for fish, with the posetic QSSAR models (dataset 1 from Table 1).

is separately fit. Indeed, the results reported in Table 1 show this trend. Overall, the BAF predictions are of good quality, showing again that the splinoid poset gives better predictions than the cluster expansion. The predicted correlation coefficients for the cluster-expansion QSSAR range between 0.745 and 0.887. The splinoid QSSARs give notably better predictions, with correlation coefficients ranging between 0.903 and 0.935.

The BCF values predicted for PCBs with the splinoid QSSAR method for different fish species (data set 1 from Table 1) and BAF values predicted for carp 10 yrs and to smelt adult (data sets 10 and 14 from Table 1) are presented in Table 2. The BCF and BAF experimental values used to compute the QSSAR models are presented in bold-face in Table 2.

For this table the PCBs are numbered according to the Ballschmiter list of congeners, and biphenyl is labeled with the number 0. The Ballschmiter labeling of PCBs is detailed at www.epa.gov/toxteam/pcbid.

The results of a cluster-expansion QSSAR model are presentable in a more conventional manner, and though presently of lower quality than the splinoid fits, they are not unreasonable. The fitted cluster-expansion parameters for the same cases are reported in Table 3 for the same cases as reported for the splinoid fits (of Table 2).

To summarily illustrate the BAF QSSAR, we present in Figures 5 and 6 the prediction plots for alewife young-of-the-year (YOY) from zone 2 Green Bay (experiment 8) and from zone 3 Green Bay (experiment 18). Again, the predictions do not show significant outliers or increased errors for low or high BAF values.

Conclusions

Chlorinated compounds are produced in large quantities and have properties that explain their accumulation in various ecosystems (namely, lipophilicity and low or negligible metabolism rate in the organism). Due to their lipophilicity, chlorinated compounds from food and from the aqueous environment accumulate in the fatty tissues of fish and shellfish. Although the aqueous concentration of these pollutants might be low, their bioaccumulation in various species of fish and shellfish can cause serious risks for human health. Polychlorinated biphenyls released from industrial activities are persistent pollutants of the environment that produce a widespread contamination of water and soil. PCBs can bioaccumulate in food chain, constituting a potential source of exposure for the general population.

In the present study we have investigated the application of quantitative super-structure-activity relationships for the prediction of the bioconcentration factors and bioaccumulation factors of PCBs. The experimental determination of BCFs and BAFs is time-consuming and expensive, so that theoretical models for their prediction are particularly appealing. To predict the bioconcentration and bioaccumulation

Table 2. PCBs experimental and predicted values with splinoid QSSAR for BCF (different fish species, data set 1 from Table 1) and BAF (carp 10 years and smelt adult, data sets 10 and 14 from Table 1). The experimental values are presented in bold.

# PCB	log BCF for fish	log BAF for smelt adult	log BAF for carp 10yrs
0	**2.64**	4.09	5.04
1	3.15	4.79	5.60
2	3.24	4.78	5.61
3	**2.77**	4.77	5.58
4	**3.38**	5.40	6.00
5	**4.11**	5.47	6.15
6	**3.80**	5.44	**6.24**
7	**3.55**	5.46	6.11
8	**3.57**	5.46	6.12
9	**3.89**	5.46	6.16
10	3.69	5.44	6.02
11	3.79	5.50	6.22
12	3.66	5.49	6.16
13	3.66	5.49	6.16
14	**3.78**	5.50	6.20
15	**3.28**	5.50	6.16
16	4.18	5.99	6.57
17	4.21	**6.05**	**6.61**
18	**4.11**	**5.90**	**6.57**
19	4.03	5.86	**6.09**
20	4.25	6.15	6.74
21	4.26	6.18	6.74
22	4.25	**6.23**	**6.77**
23	4.30	6.17	6.78
24	4.26	6.08	6.66
25	4.23	**6.13**	**6.61**
26	4.26	**6.24**	**6.89**
27	4.21	5.98	6.55
28	**4.20**	6.20	6.76
29	**4.26**	6.19	6.78
30	4.19	6.08	6.61
31	**4.23**	6.22	6.81
32	4.18	6.05	6.63
33	4.25	**6.13**	**6.68**
34	4.26	6.15	6.75
35	4.23	6.25	6.82
36	4.27	6.25	6.84
37	4.16	6.34	6.89
38	4.26	6.35	6.89
39	4.25	6.37	6.92
40	**4.23**	**6.61**	**7.13**
41	4.74	6.69	7.21
42	4.71	6.76	7.25
43	4.69	**6.55**	**7.14**
44	**4.84**	6.73	7.26
45	4.64	**6.37**	**6.96**
46	4.61	**6.20**	**6.71**

(*Continued on next page*)

Table 2. (*Continued*)

# PCB	log BCF for fish	log BAF for smelt adult	log BAF for carp 10yrs
47	**4.85**	6.79	7.21
48	**5.00**	6.75	7.26
49	**4.84**	**6.80**	**7.31**
50	4.64	6.53	6.97
51	4.64	**6.33**	**7.02**
52	**4.63**	**6.87**	**7.35**
53	4.63	**6.26**	**6.83**
54	**3.85**	6.37	6.81
55	4.71	6.86	7.33
56	4.73	6.90	7.39
57	4.73	6.92	7.42
58	4.73	6.89	7.38
59	4.68	6.72	7.23
60	4.69	6.96	7.43
61	4.75	6.90	7.40
62	4.69	6.76	7.26
63	4.70	**7.10**	**7.64**
64	**4.60**	6.80	7.30
65	4.70	6.77	7.29
66	4.69	6.96	7.42
67	4.76	6.92	7.41
68	4.72	6.91	7.38
69	4.67	6.67	7.15
70	**4.77**	6.96	7.45
71	4.70	6.59	7.15
72	4.75	6.88	7.38
73	4.70	6.67	7.18
74	4.68	**7.10**	**7.65**
75	4.68	6.69	7.13
76	4.77	6.89	7.38
77	**4.59**	7.10	7.57
78	4.84	7.00	7.47
79	4.83	7.05	7.52
80	4.79	6.955	7.44
81	4.84	**7.67**	**8.00**
82	5.15	**7.31**	**7.74**
83	5.15	**7.65**	**8.07**
84	5.10	7.07	7.53
85	5.18	**7.52**	**7.90**
86	5.26	7.29	7.75
87	**5.38**	**7.32**	**7.72**
88	5.16	7.09	7.53
89	5.13	**6.55**	**7.32**
90	**5.00**	7.38	7.80
91	5.13	**7.26**	**7.66**
92	5.22	7.32	7.74
93	5.13	7.14	7.60
94	5.11	7.06	7.55
95	5.15	7.13	7.57
96	5.00	6.94	7.43
97	**5.43**	**7.25**	**7.73**
98	5.09	7.02	7.46
99	**5.00**	**7.55**	**7.92**
100	5.05	**6.92**	**7.01**
101	**5.40**	**7.50**	**7.90**
102	5.16	7.04	7.53
103	5.11	7.13	7.57
104	4.90	6.95	7.40
105	**5.00**	7.50	7.92
106	5.22	7.45	7.88
107	5.20	**7.87**	**8.22**
108	5.18	7.48	7.89
109	**5.00**	7.24	7.68
110	5.15	7.31	7.75
111	5.21	7.47	7.90
112	5.15	7.30	7.73
113	5.16	7.33	7.78
114	5.20	7.60	8.02
115	5.10	7.37	7.79
116	5.15	7.30	7.74
117	5.12	7.40	7.83
118	**5.00**	**7.87**	**8.32**
119	5.10	7.25	**7.70**
120	5.23	7.53	7.96
121	5.11	7.29	7.74
122	5.21	7.47	7.90
123	5.23	7.62	8.04
124	5.27	7.53	7.97
125	5.17	7.25	7.73
126	**5.81**	7.73	8.15
127	5.35	7.56	7.99
128	**5.77**	7.74	8.14
129	5.59	7.71	8.12
130	5.55	**7.83**	**8.18**
131	5.53	**6.89**	**7.32**
132	5.55	7.54	7.96
133	5.57	7.79	8.20
134	5.52	7.62	8.03
135	5.53	7.61	8.08
136	**5.43**	7.46	7.89
137	**5.88**	7.80	8.21
138	**5.39**	7.80	8.20
139	5.49	7.63	7.99
140	5.48	7.51	7.92
141	**5.81**	7.71	8.11
142	5.56	7.52	7.94
143	5.54	7.52	7.98
144	5.55	7.53	8.00

(*Continued*)

(*Continued on next page*)

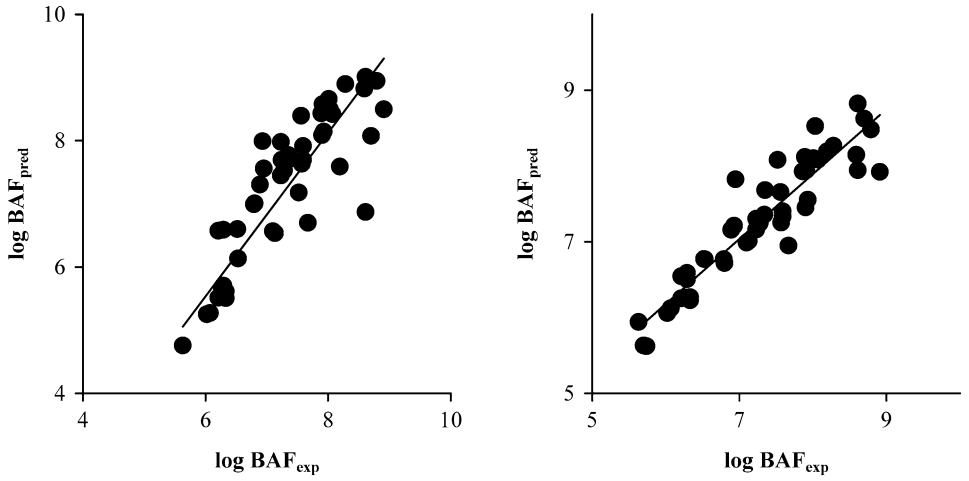
Table 2. (*Continued*)

# PCB	log BCF for fish	log BAF for smelt adult	log BAF for carp 10yrs
145	5.46	7.42	7.87
146	5.59	**7.79**	**8.16**
147	5.51	7.66	8.05
148	**5.39**	7.59	8.10
149	5.57	**7.48**	**7.81**
150	5.44	7.50	7.92
151	**5.54**	**7.32**	**7.65**
152	5.48	7.51	7.95
153	**5.65**	7.90	8.29
154	5.46	7.60	7.97
155	**4.93**	7.45	7.76
156	**5.39**	7.95	8.37
157	**5.39**	7.92	8.35
158	5.48	7.76	8.17
159	5.58	7.87	8.30
160	5.52	7.6	8.04
161	5.51	7.75	8.27
162	5.59	7.98	8.40
163	5.53	7.81	8.21
164	5.56	7.82	8.25
165	5.54	7.78	8.22
166	5.54	7.74	8.14
167	5.62	**8.46**	**8.95**
168	5.53	7.76	8.21
169	**5.97**	8.03	8.46
170	5.80	8.09	8.54
171	5.80	7.89	8.34
172	5.81	8.08	8.55
173	5.79	7.81	8.28
174	**5.80**	**7.84**	**8.18**
175	5.78	7.86	**9.02**
176	5.77	7.77	8.22
177	5.79	**8.03**	**8.42**
178	5.78	7.92	8.39
179	5.76	7.81	8.27
180	**5.80**	**8.18**	**8.65**
181	5.84	7.94	8.38
182	**5.80**	7.96	8.44
183	**5.84**	**7.93**	**8.33**
184	5.71	7.89	8.32
185	5.82	**7.50**	**8.00**
186	5.78	7.85	8.29
187	**5.80**	7.91	8.33
188	5.75	7.87	8.35
189	5.78	8.30	8.71
190	5.77	8.04	8.44
191	**5.84**	8.18	8.61
192	5.79	8.05	8.50
193	5.80	**8.50**	**8.94**
194	**5.81**	**8.68**	**9.27**
195	**5.92**	8.29	8.76
196	**5.92**	8.29	8.83
197	5.93	8.21	8.68
198	**5.88**	8.23	8.81
199	5.88	8.33	8.90
200	5.89	**8.18**	**8.64**
201	5.89	**8.05**	**8.50**
202	**5.82**	8.19	8.69
203	5.91	8.24	8.72
204	5.92	8.22	8.70
205	5.92	8.46	**8.50**
206	5.81	**8.66**	**9.26**
207	5.84	8.52	9.04
208	**5.71**	8.50	9.02
209	**5.44**	**8.66**	**9.28**

Table 3. Fitted cluster-expansion parameters for log BCF and log BAF.

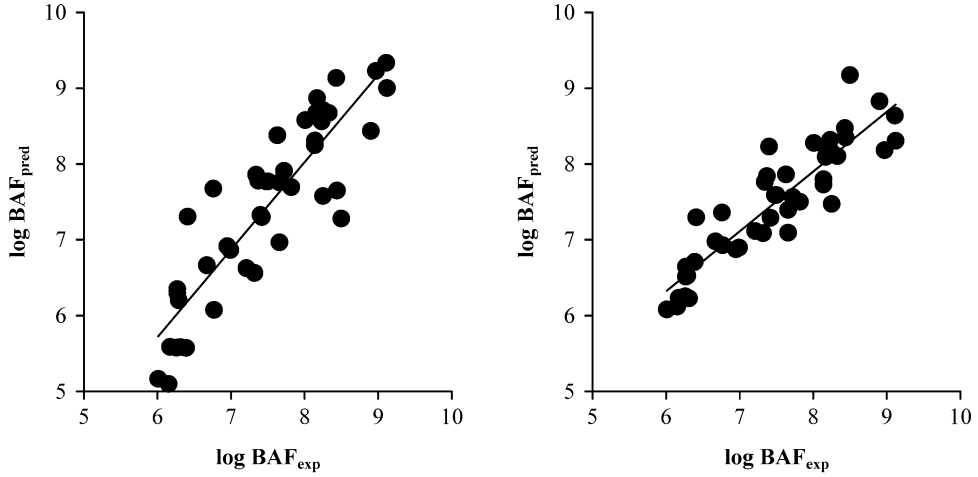
Case	*a*	*b*	*c*	*d*	*e*	*f*
BCF for fish	1.343	1.547	1.529	−0.544	−0.492	−0.328
BAF to carp 10 yrs	2.020	2.279	2.299	−0.703	−0.736	−0.624
BAF to smelt adult	1.788	2.114	2.251	−0.664	−0.655	−0.497

factors for PCBs we applied two poset QSSAR models (namely the splinoid poset and cluster-expansion) specifically developed for a network of chemical compounds that can be derived by a substitution reaction from a parent skeleton (here biphenyl). These QSSAR models based on the poset reaction diagram reflect in distinct ways the topology of the reaction network (or "super-structure", being "beyond" ordinary molecular structure) that describes the interconversion of the chemical species in the network. The cluster-expansion is a parametric method, which bears some relation [68] to QSAR models based on standard subgraphic cluster expansions. On the other hand the splinoid poset method is a global interpolation method. The two may be reasonably compared using the leave-one-out cross validation procedure. Problems [69,70] found in some QSAR schemes (*i.e.*, those where the set of descriptors chosen to make a fitting is optimally selected from a much larger super-set of descriptors) are avoided with the present cluster-expansion QSSAR scheme in that we entertained no larger super-set of descriptors – with the poset, the cluster-expansion descriptors are themselves partially ordered, and one naturally takes the earlier ones. For the splinoid scheme even speaking of descriptors (and thence sets or super-sets of descriptors) is simply somewhat "foreign". Our two (robust) QSSAR models considered here were purposely developed for modeling properties in series of compounds that can be formally derived by substitution from a parent skeleton, such as benzene, but also naphthalene, or



a) cluster-expansion: r = 0.866, s = 0.453 b) splinoid poset: r = 0.932, s = 0.301

Figure 5. Plot of experimental *vs.* predicted PCB log BAFs for alewife YOY from Green Bay zone 2, with the posetic QSSAR models (dataset 8 from Table 1).



a) cluster-expansion: $r = 0.887$, $s = 0.421$ b) splinoid poset: $r = 0.903$, $s = 0.341$

Figure 6. Plot of experimental vs. predicted PCB log BAFs for alewife YOY from Green Bay zone 3, with the posetic QSSAR models (dataset 18 from Table 1).

biphenyl, or other parent frameworks. Previously the splinoid poset QSSAR has been used to model various toxicity indices of chlorobenzenes for guppy, fathead minnow, brine shrimp, *Daphnia magna*, algae, and tadpoles [64].

Using the cluster-expansion and splinoid QSSAR we have here modeled the bioconcentration factors of 58 PCBs measured for several fish species, namely guppies, fathead minnow, rainbow trout, and bluegill sunfish. The best predictions were obtained with the splinoid poset ($r_{cv} = 0.958$), which gives a QSSAR model that can be used to predict the BCF for the remaining PCBs.

The BAF models were obtained for 19 series of PCB bioaccumulation factors determined for various fish species (carp, alewife, shad, walleye, smelt, and yellow perch) in the Green Bay area and in the Hudson River. Because BAFs were measured in natural ecosystems, the experimental values seem to be affected by larger errors than BCF values, which explains the somewhat lower prediction statistics obtained in the BAF QSSAR models. Nonetheless, the BAF predictions are of good quality, showing again that the splinoid poset gives better predictions than the present cluster-expansion.

Thus we find compelling evidence showing that QSSAR models based on poset reaction diagrams can be successfully used to model the BCFs and BAFs of polychlorinated biphenyls. Presumably the splinoid QSSAR models for BAFs and BCFs could be used to obtain reliable predictions for other reaction networks of interest.

Acknowledgments

The authors acknowledge the support (through grant BD-0894) of the Welch Foundation of Houston, Texas.

References

1. Schüürmann, G. and Klein, W., *Advances in bioconcentration prediction*, Chemosphere, 17 (1988) 1551–1574.
2. Feijtel, T., Kloepper-Sams, P., Den Haan, K., Van Egmond, R., Comber, M., Heusel, R., Wierich, P., Ten Berge, W., Gard, A., Wolf, W. and Niessen, H., *Integration of bioaccumulation in an environmental risk assessment*, Chemosphere, 34 (1997) 2337–2350.
3. Franke, C., Studinger, G., Berger, G., Böhling, S., Bruckmann, U., Cohors-Fresenborg, D. and Jöhncke, U., *The assessment of bioaccumulation*, Chemosphere, 29 (1994) 1501–1514.
4. Turne, L., Choplin, F., Dugard, P., Hermens, J., Jaeckh, R., Marsmann, M. and Roberts, D., *Structure-activity relationships in toxicology and ecotoxicology: An assessment*, Toxicol. *in vitro*, 1 (1987) 143–171.
5. Carlsen, L. and Walker, J.D., *QSARs for prioritizing PBT substances to promote pollution prevention*, QSAR Comb. Sci., 22 (2003) 49–57.
6. Swanson, G.M., Ratcliffe, H.E. and Fischer, L.J., *Human exposure to policlorinated biphenyls (PCBs): A critical assessment of the evidence for adverse health effects*, Regul. Toxicol. Pharmacol., 21 (1995) 136–150.
7. Faroon, O., Jones, D. and De Rosa, C., *Effects of polychlorinated biphenyls on the nervous system*, Toxicol. Ind. Health, 16 (2000) 305–333.
8. Kamrin, M.A. and Fischer, L.J., *Current status of sport fish consumption advisories for PCBs in the Great Lakes*, Regul. Toxicol. Pharmacol., 29 (1999) 175–181.
9. Judd, N., Griffith, W.C. and Faustman, E.M., *Contribution of PCB exposure from fish consumption to total dioxin-like dietary exposure*, Regul. Toxicol. Pharmacol., 40 (2004) 125–135.
10. Ivanciuc, O., *Support vector machines prediction of the mechanism of toxic action from hydrophobicity and experimental toxicity against Pimephales promelas and Tetrahymena pyriformis*, Internet Electron. J. Mol. Des., 3 (2004) 800–821, http://www.biochempress.com.
11. Carlsen, L., *A QSAR approach to physico–chemical data for organophosphates with special focus on known and potential nerve agents*, Internet Electron. J. Mol. Des., 4 (2005) 355–366, http://www.biochempress.com.
12. Seegal, R.F. and Bowers, W.J., *Consequences and mechanisms of action of fish-borne toxicants: what we do not know and why*, Environ. Toxicol. Pharmacol., 12 (2002) 63–68.
13. Mackay, D. and Fraser, A., *Bioaccumulation of persistent organic chemicals: mechanisms and models*, Environ. Pollut., 110 (2000) 375–391.
14. Hope, B., Scatolini, S. and Titus, E., *Bioconcentration of chlorinated biphenyls in biota from the North Pacific Ocean*, Chemosphere, 36 (1998) 1247–1261.
15. Voutsas, E., Magoulas, K. and Tassios, D., *Prediction of the bioaccumulation of persistent organic pollutants in aquatic food webs*, Chemosphere, 48 (2002) 645–651.
16. Meylan, W.M., Howard, P.H., Boethling, R.S., Aronson, D. Printup, H. and Gouchie, S., *Improved method for estimating bioconcentration/bioaccumulation factor from octanol/water partition coefficient*, Environ. Toxicol. Chem., 18 (1999) 664–672.
17. Ivanciuc, O., *Aquatic toxicity prediction for polar and nonpolar narcotic pollutants with support vector machines*, Internet Electron. J. Mol. Des., 2 (2003) 195–208, http://www.biochempress.com.
18. Hansch, C. and Leo, A., Exploring QSAR. Fundamentals and applications in chemistry and biology, American Chemical Society, Washington, DC, 1995, pp 580.
19. Gramatica, P. and Papa, E., *QSAR modeling of bioconcentration factor by theoretical molecular descriptors*, QSAR Comb. Sci., 22 (2003) 374–385.
20. Devillers, J., Domine, D., Bintein, S. and Karcher, W., *Fish bioconcentration modeling with log P*, Toxicol. Methods, 8 (1998) 1–10.
21. Dearden, J.C. and Shinnawei, N.M., *Improved prediction of fish bioconcentration factor of hydrophobic chemicals*, SAR QSAR Environ. Res., 15 (2004) 449–455.
22. Hu, H., Xu, F., Li, B., Cao, J., Dawson, R. and Tao, S., *Prediction of the bioconcentration factor of PCBs in fish using the molecular connectivity index and fragment constant models*, Water Environ. Res., 77 (2005) 136–150.
23. Davies, R.P. and Donas, A.J., *The prediction of bioconcentration in fish*, Water Res., 18 (1984) 1253–1262.
24. Fatemi, M.H., Jalali-Heravi, M. and Konuze, E., *Prediction of bioconcentration factor using genetic algorithm and artificial neural network*, Anal. Chim. Acta, 486 (2003) 101–108.
25. Boon, J.P. and Duinker, J.C., *Kinetics of polychlorinated biphenyl (PCB) components in juvenile sole (Solea solea) in relation to concentrations in water and to lipid metabolism under conditions of starvation*, Aquat. Toxicol., 7 (1985) 119–134.
26. Renberg, L. and Sundström, G., *Prediction of bioconcentration potential of organic compounds using partition coefficients derived from reversed phase thin layer chromatography*, Chemosphere, 7 (1979) 449–459.
27. Harald, J., Geyer, I., Scheunert, I. and Korte, F., *Correlation between the bioconcentration potential of organic environmental chemicals in humans and their n-octanol/water partition coefficients*, Chemosphere, 16 (1987) 239–252.
28. Tao, S., Hu, H., Lu, X., Dawson, R.W. and Xu, F., *Fragment constant method for prediction of fish bioconcentration factors of non-polar chemicals*, Chemosphere, 41 (2000) 1563–1568.
29. Ivanciuc, O., *Artificial neural networks applications. Part 7 - Estimation of bioconcentration factors in fish using solvatochromic parameters*, Rev. Roum. Chim., 43 (1998) 347–354.
30. Dimitrov, S.D., Dimitrova, N.C., Walker, J.D., Veith, G.D. and Mekenyan, O.G., *Predicting bioconcentration factors of highly hydrophobic chemicals. Effects of molecular size*, Pure Appl. Chem., 74 (2002) 1823–1830.
31. Dimitrov, S.D., Dimitrova, N.C., Walker, J.D., Veith, G.D. and Mekenyan, O.G., *Bioconcentration potential predictions based on molecular attributes – an early warning approach for chemicals found in humans, birds, fish and wildlife*, QSAR Comb. Sci., 22 (2003) 58–68.
32. Gobas, F.A.P.C., *A model for predicting the bioaccumulation of hydrophobic organic chemicals in aquatic food-webs: Application to Lake Ontario*, Ecol. Model., 69 (1993) 1–17.
33. Boon, J.P., Oostingh, I., Van der Meer, J., and Hillebrand, T.M.J., *A model for the bioaccumulation of chlorobiphenyl congeners in marine mammals*, Eur. J. Pharmacol., Environ. Toxicol. Pharmacol., 270 (1994) 237–251.
34. Walker, C.H., *Kinetic models for predicting bioaccumulation of pollutants in ecosystems,* Environ. Pollut., 44 (1987) 227–240.
35. Müller, J.F., Hawker, D.W. and Connell, D.W., *Calculation of bioconcentration factors of persistent hydrophobic compounds in the airnegetation system*, Chemosphere, 29 (1994) 623–640.
36. Ivanciuc, O., *Bioconcentration factor QSAR with support vector regression and artificial neural networks*, Internet Electron. J. Mol. Des., 4 (2005) 813–834, http://www.biochempress.com.
37. Bruggeman, W.A., Martron, L.B.J.M., Kooiman, D. and Hutzinger, O., *Accumulation and elimination of di-, tri-, and tetrachlorobiphenyls by goldfish after dietary and aqueous exposure*, Chemosphere, 10 (1981) 811–832.
38. Ivanciuc, O., *Support vector machine identification of the aquatic toxicity mechanism of organic compounds*, Internet Electron. J. Mol. Des., 1 (2002) 157–172, http://www.biochempress.com.
39. Echols, K.R., Gale, R.W., Schwartz, T.R., Huckins, J.N., Williams, L.L.,Meadows, J.C., Morse, D.,Petty, J.D., Orazio, C.E. and Tillitt, D.E., *Comparing polychlorinated biphenyl concentrations and patterns in the Saginaw River using sediment caged fish, and semipermeable membrane devices*, Environ. Sci. Technol., 34 (2000) 4095–4102.
40. Geyer, H.J., Scheunert, I., Bruggemann, R., Steinberg, C., Korte, F. and Kettrup, A., *QSAR for organic chemical bioconcentration in*

Daphnia, algae, and mussels, Sci. Total. Environ., 109/110 (1991) 387–394.
41. Devillers, J., Bintein, S. and Domine, D., *Comparison of BCF models based on log P*, Chemosphere, 33 (1996) 1047–1065.
42. Wei, D., Zhang, A., Wu, C., Han, S. and Wang, L., *Progressive study and robustness test of QSAR model based on quantum chemical parameters for predicting BCF of selected polychlorinated organic compound (PCOCs)*, Chemosphere, 44 (2001) 1421–1428.
43. Saçan, M.T., Erdem, S.S., Özpinar, G.A. and Balcioglu, I.A., *QSPR study on the bioconcentration factors of nonionic organic compounds in fish by characteristic root index and semiempirical molecular descriptors*, J. Chem. Inf. Comput. Sci., 44 (2004) 985–992.
44. Lu, X., Tao, S., Cao, J. and Dawson, R.W., *Prediction of fish bioconcentration factors of nonpolar organic pollutants based on molecular connectivity indices*, Chemosphere, 39 (1999) 987–999.
45. Lu, X., Tao, S., Hu, H. and Dawson, R.W., *Estimation of bioconcentration factors of nonionic organic compounds in fish by molecular connectivity indices and polarity correction factors*, Chemosphere, 41 (2000) 1675–1688.
46. Fox, K., Zauke, G.P. and Butte, W., *Kinetics of bioconcentration and clearance of 28 polychlorinated biphenyl congeners in Zebrafish (Brachydanio rerio)*, Ecotox. Environ. Safety, 28 (1994) 99–109.
47. Burkhard, L.P., Endicott, D.D., Cook, P.M., Sappington, K.G. and Winchester, E.L., *Evaluation of two methods for prediction of bioaccumulation factors*, Environ. Sci. Technol., 37 (2003) 4626–4634.
48. Klein, D.J., *Similarity and dissimilarity in posets*, J. Math. Chem., 18 (1995) 321–348.
49. Klein, D.J. and Babić, D., *Partial orderings in chemistry*, J. Chem. Inf. Comput. Sci., 37 (1997) 656–671.
50. Klein, D.J., *Prolegomenon on partial orderings in Chemistry*, MATCH Commun. Math. Comput. Chem., 42 (2000) 7–21.
51. Klein, D.J. and Bytautas, L., *Directed reaction graphs as posets*, MATCH Commun. Math. Comput. Chem., 42 (2000) 261–289.
52. Brüggemann, R., Schwaiger, J. and Negele, R.D., *Applying Hasse diagram technique for the evaluation of toxicological fish tests*, Chemosphere, 30 (1995) 1767–1780.
53. Brüggemann, R. and Bartel, H.G., *A theoretical concept to rank environmentally significant chemicals*, J. Chem. Inf. Comput. Sci., 39 (1999) 211–217.
54. Pudenz, S., Brüggemann, R., Luther, B., Kaune, A. and Kreimes K., *An algebraic/graphical tool to compare ecosystems with respect to their pollution V: Cluster analysis and Hasse diagrams*, Chemosphere, 40 (2000) 1373–1382.
55. Brüggemann, R., Pundez, S., Carlsen, L., Sørensen, P.B., Thomsen, M. and Mishra, R.K., *The use of Hasse diagrams as a potential approach for inverse QSAR*, SAR QSAR Environ. Res., 11 (2001) 473–487.
56. Carlsen, L., Sørensen, P.B., Thomsen, M. and Brüggemann R., *QSAR's based on partial order ranking*, SAR QSAR Environ. Res., 13 (2002) 153–165.
57. Brüggemann, R., Münzer, B. and Halfon, E., *An algebraic/graphical tool to compare ecosystems with respect to their pollution – the German river "Elbe" as an example - I: Hasse-diagrams*, Chemosphere, 28 (1994) 863–872.
58. Lerche, D., Sørensen, P.B., Larsen H.L., Carlsen, L. and Nielsen, O.J., *Comparison of the combined monitoring-based and modelling-based priority setting scheme with partial order theory and random linear extensions for ranking of chemical substances*, Chemosphere, 49 (2002) 637–649.
59. Lerche, D. and Sørensen, P.B., *Evaluation of the ranking probabilities for partial orders based on random linear extensions,* Chemosphere, 53 (2003) 981–992.
60. Lerche, D., Matsuzaki, S.Y., Sørensen, P.B., Carlsen, L. and Nielsen, O.J., *Ranking of chemical substances based on the Japanese pollutant release and transfer register using partial order theory and random linear extensions*, Chemosphere, 55 (2004) 1005–1025.
61. Sørensen, P.B., Mogensen, B.B., Carlsen, L. and Thomsen, M., *The influence on partial order ranking from input parameter uncertainty: Definition of a robustness parameter*, Chemosphere, 41 (2000) 595–601.
62. Klein, D.J., Schmalz, T.G. and Bytautas, L., *Chemical sub-structural cluster expansions for molecular properties*, SAR QSAR Environ. Res., 10 (1999) 131–156.
63. Ivanciuc, T. and Klein D.J., *Parameter-free structure-property correlation via progressive reaction posets for substituted benzenes,* J. Chem. Inf. Comput. Sci., 44 (2004) 610–617.
64. Ivanciuc, T., Ivanciuc, O. and Klein, D.J., *Posetic quantitative superstructure/activity relationships (QSSARs) for chlorobenzenes*, J. Chem. Inf. Model., 45 (2005) 870–879.
65. Došlić, T. and Klein, D.J., *Splinoid interpolation on finite posets*, J. Comput. Appl. Math., 177 (2005) 175–185.
66. Klein, D.J., *Chemical graph-theoretic cluster expansions*, Int. J. Quantum. Chem., Quantum Chem. Symp., 20 (1986) 153–171.
67. Schmalz, T.G., Živković, T. and Klein, D.J., *Cluster expansion of the Hückel molecular energy for acyclics: Applications to pi resonance theory*, Math. Chem. Comp., 54 (1987) 173–190.
68. Ivanciuc, T., Klein, D.J. and Ivanciuc, O., *Posetic cluster expansion for substitution – reaction diagrams and its application to cyclobutane*, J. Math. Chem. (submitted).
69. Gillet, V.J., Willett, P. and Bradshaw, J., *Identification of biological activity profiles using substructural analysis and genetic algorithms*, J. Chem. Inf. Comput. Sci., 38 (1998) 165–179.
70. Hawkins, D.M., Basak, S C. and Shi, X., *QSAR with few compounds and many features*, J. Chem. Inf. Comput. Sci., 41 (2001) 663–670.