



# Modeling adsorption of brominated, chlorinated and mixed bromo/chloro-dibenzo-*p*-dioxins on C<sub>60</sub> fullerene using Nano-QSPR

Piotr Urbaszek<sup>1</sup>, Agnieszka Gajewicz<sup>1</sup>, Celina Sikorska<sup>2</sup>, Maciej Haranczyk<sup>3</sup> and Tomasz Puzyn<sup>\*1</sup>

## Full Research Paper

[Open Access](#)**Address:**

<sup>1</sup>Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland, <sup>2</sup>Laboratory of Molecular Modeling, Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland and <sup>3</sup>IMDEA Materials Institute, C/Eric Kandel 2, 28906 Getafe, Madrid, Spain

**Email:**

Tomasz Puzyn<sup>\*</sup> - t.puzyn@qsar.eu.org

<sup>\*</sup> Corresponding author

**Keywords:**

brominated; chlorinated; dioxins; fullerenes; QSPR; sorption

*Beilstein J. Nanotechnol.* **2017**, *8*, 752–761.

doi:10.3762/bjnano.8.78

Received: 07 November 2016

Accepted: 14 March 2017

Published: 31 March 2017

Associate Editor: J. Frommer

© 2017 Urbaszek et al.; licensee Beilstein-Institut.

License and terms: see end of document.

## Abstract

Many technological implementations in the field of nanotechnology have involved carbon nanomaterials, including fullerenes such as the buckminsterfullerene, C<sub>60</sub>. The unprecedented properties of such organic nanomaterials (in particular their large surface area) gained extensive attention for their potential use as organic pollutant sorbents. Sorption interactions can be very hazardous and useful at the same time. This work investigates the influence of halogenation by bromine and/or chlorine in dibenzo-*p*-dioxins on their sorption ability on the C<sub>60</sub> fullerene surface. Halogenated dibenzo-*p*-dioxins (PXDDs, where X = Br or Cl) are ever-present in the environment and accidentally produced in many technological processes in only approximately known quantities. If all combinatorial Br and/or Cl dioxin substitution possibilities are present in the environment, the experimental characterization and investigation of sorbent effectiveness is more than difficult. In this work, we have developed a quantitative structure–property relationship (QSPR) model ( $R^2 = 0.998$ ), predicting the adsorption energy [kcal/mol] for 1,701 PXDDs adsorbed on C<sub>60</sub> (PXDD@C<sub>60</sub>). Based on the QSPR model reported herein, we concluded that the lowest energy PXDD@C<sub>60</sub> complexes are those that the World Health Organization (WHO) considers to be less dangerous with respect to the aryl hydrocarbon receptor (AhR) toxicity mechanism. Therefore, the effectiveness of fullerenes as sorbent agents may be underestimated as sorption could be less effective for toxic congeners than previously believed.

## Introduction

### Dioxin congeners are present and dangerous

Studies on chlorinated dibenzo-*p*-dioxins (PCDDs) as representatives of persistent organic pollutants (POPs) [1] are an important area of the environmental sciences and scientific research [2–5]. PCDDs are usually represented by 2,3,7,8-tetrachloro dibenzo-*p*-dioxin (TCDD), considered as one of the most dangerous and toxic for living organisms upon long-term exposure [6]. Previous studies have shown that other PXDDs (halogenated-brominated and/or chlorinated dioxins, where X = Br or Cl) can also cause toxic effects [7] and induce diverse enzymes and receptors such as aryl hydrocarbon hydroxylase (AHH) and 7-ethoxyresorufin-O-deethylase (EROD) [8]. Most of the toxic effects caused by dioxins are thought to be mediated through a specific protein complex known as the aryl hydrocarbon receptor (AhR) [9]. To interact with AhR, the dioxin structure must penetrate into the cell. This task is easiest for dioxins with symmetrically substituted halogen atoms, such as TCDD. Inside the cell it interacts with the AhR receptor, proteins, and finally, by entering the nucleus, it reacts with the so-called dioxin responsive element (DRE) region on the mRNA surface and causes errors in the translation process and synthesis of new proteins. Most studies are focused on chlorinated dioxins, but brominated dioxins can also be found in environmental samples [7]. Furthermore, in some cases, brominated dioxins show an even higher AhR receptor binding affinity than 2,3,7,8-TCDD [6,7,10]. Because of natural processes occurring in the environment, the total amount of chlorinated PXDD derivatives is constantly decreasing, while the amount of brominated and mixed congeners (molecules based on the same carbon skeleton differing by the number and type of substituents) is increasing [11].

### Fullerene C<sub>60</sub> – opportunities and risks of possible surface interactions

Fullerene C<sub>60</sub> [12–14], discovered in 1985, has a soccer ball-like structure [15] with a chemical structure representative of carbon nanostructures. Its unique properties and shape make C<sub>60</sub> and its derivatives promising candidates for various applications, including sorbents, cancer therapeutics, drug delivery systems, computer sensors, etc. [16–19]. With the further development of nanotechnology, C<sub>60</sub> will be produced and used in large amounts. Over time, fullerene structures will be found in the environment more often and in higher concentrations.

Aromatic structures render fullerenes as good acceptors of  $\pi$ -electrons. On the other hand, aromatic systems like halogenated dioxins are classified as  $\pi$ -donors [20]. Recent studies have proved that halogens, such as bromine or chlorine, have a more positive region on the surface opposite to the X–C bond direction as well as an equatorial belt of negative potential, so

that they can display different properties depending on the angle of approach [21]. Regarding one of the toxicity mechanisms for nanoparticles proposed at the NATO Advance Research Workshop, dioxin–fullerene interactions and complex formation can be dangerous because of the ability of nanoparticles like fullerenes to penetrate biological barriers and act as a vector, transferring dioxins or other pollutants inside the cells [22].

The potential applications of C<sub>60</sub> as smoke filters or air-cleaning agents are examples where it may be employed to improve environmental conditions [13,18]. On the other hand, if the sorption interactions can occur spontaneously in the environment, they may bring as much hope as they do risk [23–26].

## Objective

The possibility of surface interactions between fullerenes and organic compounds has raised the question: How many halogenated PXDDs congeners will create a PXDD@C<sub>60</sub> complex based on weak  $\pi$ – $\pi$  interactions, and what is the influence of halogen substitution of dioxin congeners in these interactions?

The main goal of this study was to calculate the adsorption energy for the representative subset of dioxin congeners and to develop a model to predict the energy for a large subset of structurally similar compounds. At the same time, the goal was to demonstrate that it is possible to predict the influence of the substitution pattern (i.e., type, number, and location of halogen substituents) in dioxin molecules on the final adsorption energy of the complex by using *in silico* methods.

Our model presented here may provide important information in designing new fullerene applications and assessing the risk of those interactions according to the differences in toxicity caused by the number and type of substitution. The investigations have been performed with quantitative structure–property relationship modelling for nanomaterials (Nano-QSPR) – a method of defining a mathematical function that connects the structure of the investigated nanomaterial (fullerene) and the POPs (dioxin) with a modeled property (energy of the PXDD@C<sub>60</sub> complex). It is a computational technique that, to the best of our knowledge, is the first published example of the use of Nano-QSPR to predict interactions between fullerenes and numerous organic pollutants.

## Results

### Nano-QSPR model

Based on the adsorption energy values ( $\Delta E_{\text{ads}}$ ) for 32 Br/Cl dibenzo-*p*-dioxin congeners adsorbed on a C<sub>60</sub> fullerene surface and carefully selected structural descriptors, we developed

a Nano-QSPR model, employing a hybrid genetic algorithm, partial least squares linear regression (GA-PLS), as the modeling method. The developed Nano-QSPR model utilizes only four descriptors for predicting the adsorption energy values for 1,669 PXDD@C<sub>60</sub> materials as follows:

$$\begin{aligned} \Delta E_{\text{ads}}[\text{kcal/mol}] &= 7.37 \cdot \#H - 6.76 \cdot TE - 0.11 \cdot D_x - 0.07 \cdot D_y \\ R^2 &= 0.998; Q^2_{\text{CV}} = 0.987; Q^2_{\text{Ext}} = 0.956; \\ \text{RMSE}_C &= 1.179; \text{RMSE}_{\text{CV}} = 1.468; \text{RMSE}_P = 2.285 \end{aligned} \quad (1)$$

where #H is the number of hydrogen atoms in the dioxin molecule, TE is the total energy of the molecule, and D<sub>x</sub> and D<sub>y</sub> are the dipole moments of dioxin molecule along the *x* and *y* axis, respectively. More detailed information about the statistical description of the obtained model is available in Supporting Information File 1.

### Applicability domain and OECD guidelines

The developed Nano-QSPR model has been comprehensively validated according to the Organization for Economic Co-operation and Development (OECD) QSAR validation recommendations [27] and fulfills all the validation criteria. The presented model has a well-defined endpoint ( $\Delta E_{\text{ads}}$  - adsorption energy of a C<sub>60</sub>@PXDD complex) and well-known algorithms (GA-PLS). According to OECD guidelines, it is recommended to define its applicability domain (AD). This is a theoretical space for which the predictions are most reliable and applicable. The applicability domain for models based on theoretical data can be verified by use of the leverage values [28,29] and the values predicted by the model presented on the same plot (so-called Insubria graph or Insubria plot) [30]. This approach allows for verification of the AD for training, validation and prediction sets at once. The leverage value represents the distance of a compound from the training set (TS) centroid. Given this, it is possible to determine whether a predicted model response is an effect of interpolation (compound within the AD) or extrapolation (compound outside the AD) of the model [31,32]. The Insubria graph is available in Supporting Information File 1 (Figure S1) and proves that all of the congeners from training and validation sets are located in the space of applicability domain (AD). Moreover, 1,563 congeners out of 1,669 from the prediction set lie in the space of AD. Only 84 congeners (4.94% of all 1,701 dioxin congeners) have leverage values higher than critical  $h^* = 0.625$ , and for those C<sub>60</sub>@PXDDs, the predicted  $\Delta E_{\text{ads}}$  may be less reliable. We consider this to be a very good result for the Nano-QSPR model based on 24 congeners in the training set, 8 congeners in the validation set and 1,669 in the prediction set.

The developed Nano-QSPR model was internally and externally validated [28]. The leave-one-out cross-validation (LOO

method) was the algorithm chosen for internal validation and allowed the model's robustness to be calculated –  $Q^2_{\text{CV}}$  (cross-validation coefficient) and  $\text{RMSE}_{\text{CV}}$  (root mean square error of cross validation). The external validation was performed with an independent set of congeners (8 dioxins not used during model development). With the external validation, we have calculated the predictive power of the model defined by  $Q^2_{\text{Ext}}$  (0.956), the external validation coefficient, and  $\text{RMSE}_P$  (2.285), the root mean square error of prediction [33]. The selection of a final number of latent vectors was based on the cross-validation results (the lowest value of the root mean square error of cross validation,  $\text{RMSE}_{\text{CV}}$ ).

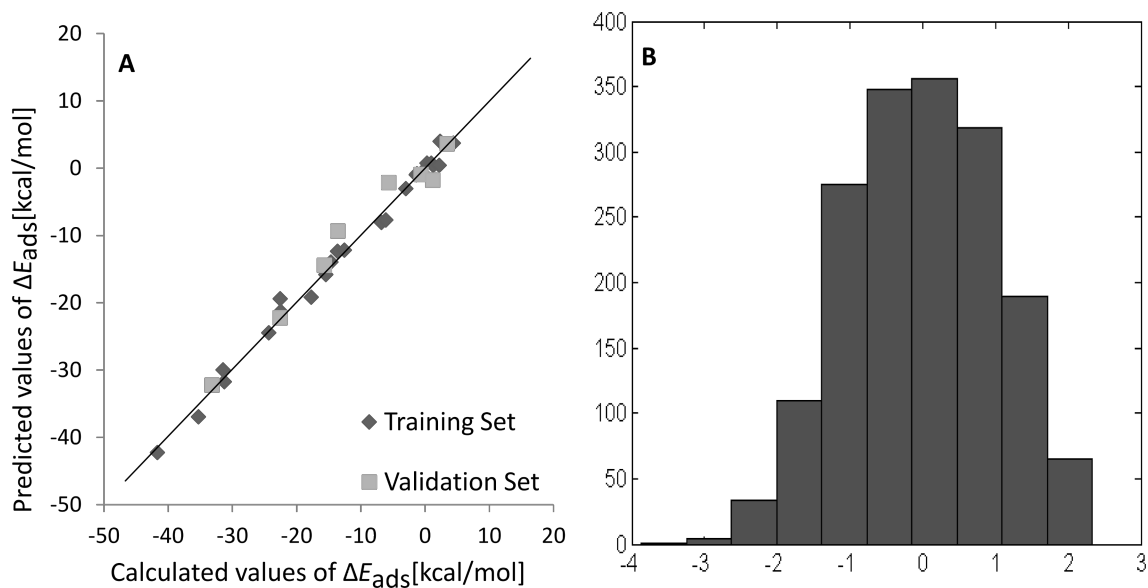
Four latent vectors (LVs), as a set of 4, were selected by the GA descriptors: #H, TE, D<sub>x</sub> and D<sub>y</sub> together explained 100% of the variance (58.08% + 40.42% + 1.30% + 0.20%) of the X-block and 99.22% variance of the Y-block (40.63% + 1.70% + 22.73% + 34.16%). According to OECD guidelines, when  $R^2$ ,  $Q^2_{\text{CV}}$ , and  $Q^2_{\text{Ext}}$  values are close to 1, and  $\text{RMSE}_C$ ,  $\text{RMSE}_{\text{CV}}$  and  $\text{RMSE}_P$  are as low as possible, the developed model can be considered as robust, well-fitted and having good predictive abilities. The Nano-QSPR model presented in this study fulfills the guidelines given by OECD, which is also shown in Figure 1a (a satisfactory correlation between the calculated and predicted  $\Delta E_{\text{ads}}$  values for the training and validation sets). Moreover, the histogram performed for autoscaled  $E_{\text{ads}}$  values for all 1,701 PXDD congeners (Figure 1b) shows that the values have a normal distribution.

### Mechanistic interpretation of the Nano-QSPR model

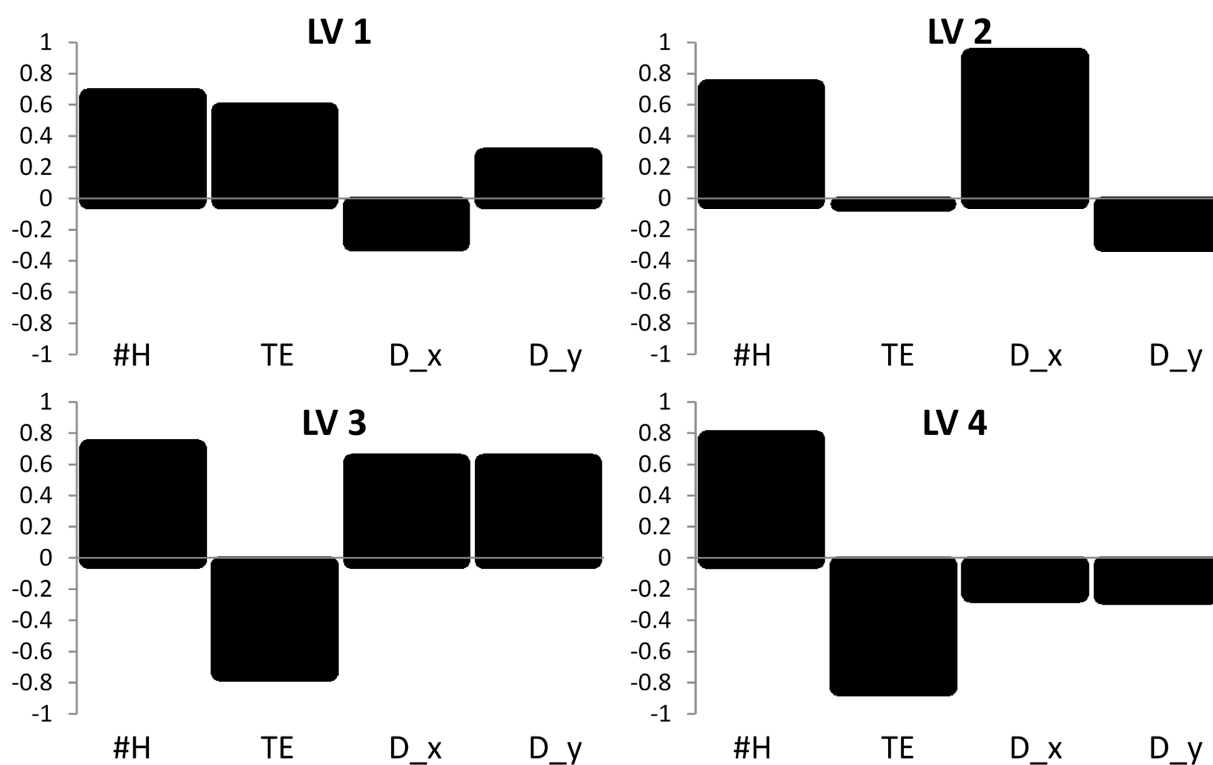
By analyzing scatter plots and loading values of the LVs, it is possible to interpret the obtained model. LV loadings show the contribution of a particular descriptor to a given latent vector (Figure 2), while score plots present training and validation set congeners in the space of LVs (Figure 3).

As presented in Figure 2 and Figure 3a–c, the first latent vector (LV1) is mainly related to the #H, TE, and D<sub>y</sub> descriptors. Such a combination suggests that congeners with hydrogen atoms, located along the *y* axis (in positions 1,4,6,9), and halogen atoms, mainly in lateral positions (2,3,7,8), have higher values of adsorption energy. D<sub>y</sub> and D<sub>x</sub> loadings with opposite signs suggest that higher values of predicted  $\Delta E_{\text{ads}}$  will have dioxins substituted by more electro-negative atoms (preferably chlorine than bromine) in nonlateral positions along the *y* axis.

The second latent vector (LV2) (Figure 2, Figure 3a), is a combination of D<sub>x</sub> and #H, and is related to those congeners with a high dipole moment along the *x* axis. Congeners mostly substi-



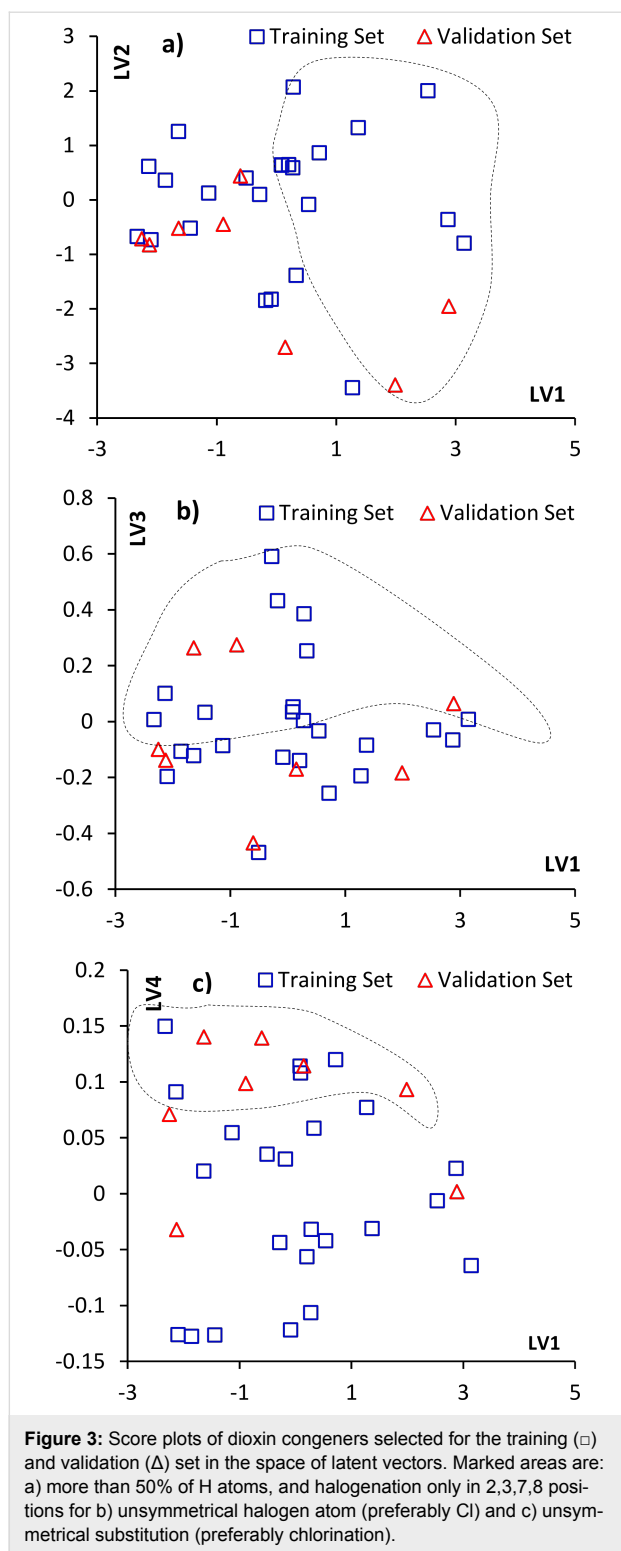
**Figure 1:** a) Plot of calculated and predicted values of  $\Delta E_{\text{ads}}$  energy. b) Histogram of calculated and predicted values of  $\Delta E_{\text{ads}}$ .



**Figure 2:** Significance and loading values of individual latent vectors (LVs).

tuted in positions 2,3,7,8 and filled with hydrogen atoms in positions 1,4,6,9 will have the highest values of LV2. Because of the differences in the electro-negativity of bromine and chlo-

rine (Br = 2.96, Cl = 3.16, Pauling scale), congeners more chlorinated than brominated would have a higher dipole moment along the x axis of dioxin, and in effect, higher predicted values



of sorption energies. It should be noted that structures considered as toxic have lateral halogen atoms and will have higher  $E_{\text{ads}}$  values, unless the number of hydrogen atoms are decreased by chlorine or bromine substitution in 1,4,6 or 9 positions.

All four descriptors have a significant contribution in the third latent vector (LV3), but only TE has a negative loading value. As shown in Figure 2 and Figure 3b, LV3 separates congeners with few hydrogen atoms substituted in the dioxin structure, and unsymmetrical halogen substituents located along both the  $x$  and  $y$  axis. The halogen atom will preferably be chlorine because of the negative TE contribution in this LV and higher influence of chlorine substitution to the dipole moment of congeners.

As shown in Figure 2 and Figure 3c, the last latent vector (LV4) is a combination of #H and negative TE, with small negative D<sub>x</sub> and D<sub>y</sub> contributions, which can be interpreted as a complementation of LV3 and separates unsymmetrical substitutions in congeners with a predominance of chlorine atoms.

Since the developed Nano-QSPR model fulfilled all OECD recommendations, including the mechanistic interpretation, we applied the model to predict the adsorption energy  $\Delta E_{\text{ads}}$  for the rest of the brominated or/and chlorinated dibenzo-*p*-dioxin congeners. All predicted data with values for particular descriptors from the prediction set are available in Supporting Information File 1, Table S3.

## Discussion

In the literature, there are only very few examples of experimental studies aimed at interactions between fullerenes or other carbon nanomaterials with particles such as proteins [34], porphyrines [35], toxic water pollutants [14], solid phases [36], or other materials [37]. Dibenz-*p*-dioxins and dibenzofurans produced during incineration of nanomaterials have also been studied [38]. There are also a few studies aimed at using *in silico* methods, such as semi-empirical or density functional theory (DFT) calculations, for exploring interactions on nanoparticle surfaces [39–41].

Our results provide new knowledge about: i) general sorption interaction mechanisms of dioxin congeners on the C<sub>60</sub> surface, and ii) applicability of the Nano-QSPR approach for predictions of organic pollutant congeneric groups.

As far as we know, the proposed study is the first attempt to use the Nano-QSPR approach for predictions of the interaction between the congeneric family of pollutants and the C<sub>60</sub> nanoparticle.

## Toxicity results

In this study, we consider toxicity to be well-described for dioxins and dioxin-like compounds as an aryl hydrocarbon receptor (AhR) interaction mechanism. AhR is a cytosolic transcription factor. Normally, the inactive protein is bound

to several co-chaperone proteins. Upon binding to a dioxin structure or dioxin-like compound, the chaperones dissociate, resulting in an AhR translocation into the nucleus and dimerization with aryl hydrocarbon receptor nuclear translocator (ARNT) protein, which leads to interaction with the dioxin responsive element (DRE) of the nucleic acids and synthesis of new proteins or changes in gene transcription [9]. Dioxins have the largest binding affinity to AhR proteins, which are symmetrically substituted in 2,3,7,8 positions. Chlorine or other halogen atoms in positions 2,3,7, and 8 in dioxin structures are essential for toxicity and also prevent the early enzymatic destruction of dioxin. Each additional chlorine in the 2,3,7,8-structure decreases the toxicity according to the AhR mechanism. However, the congeners can still cause a toxic response during enzymatic reactions or oxidation processes inside the organisms.

Less toxic (according to AhR interactions) dioxin congeners will have a halogen substitution in 1,4,6,9 positions and lower dipole moment along the *x* axis. In our study, congeners considered to be less toxic have the lowest adsorption energy values, which can suggest that they will form complexes with C<sub>60</sub> more effectively (see Supporting Information File 1 for numerical data). Therefore, the predicted adsorption energy values may suggest that the highest sorption potential will appear for those dioxin congeners which are considered as less toxic. More toxic congeners may have a lower potential to be adsorbed on the fullerene surface while competing with other congeners because of their highest adsorption energies (predicted and calculated). Those predictions are strongly correlated with lateral halogen substitution and a high dipole moment distribution along the *x* axis. It may be suggested that unmodified fullerenes as sorbents can be selectively effective in adsorbing dioxins or other congeners.

The obtained results do not consider the strength of interactions between the AhR receptor and dioxin structure. It is clear that dioxin adsorption on the fullerene surface is possible. We can also assume that, according to one of the proposed mechanisms

for nanoparticle toxicity, carbon structures such as fullerenes – because of their sorption abilities – will have the potential to act as vectors. This would allow more pollutants to enter to the cells of the organism. At this point, we have no further knowledge if dioxins adsorbed on the surface would be able to interact with receptors like AhR. It is also possible that they prefer to stay adsorbed, and in fact, might be deactivated as toxic agents.

### Confrontation with WHO TEF values

The obtained results which suggest that most of the PXDD congeners have higher or lower potential to adsorb on the C<sub>60</sub> surface have created the need to verify the sorption potential according to structural dissimilarities. The comparison of predicted  $\Delta E_{\text{ads}}$  values with toxic equivalency factors (TEFs) recommended by the World Health Organization (WHO) [42] seems to indicate that more dangerous congeners will have lowest sorption potential (highest predicted energy of the complex). As presented in Table 1, congeners with official WHO TEF values have relatively high predicted adsorption energies compared with the predicted values for congeners like 1,4,6,9-tetrachlorinated congeners ( $\Delta E_{\text{ads}}$  from –4 to –22 kcal/mol), or 2,3,7,8-tetrabromo-substituted congeners ( $\Delta E_{\text{ads}}$  from –19 to –22 kcal/mol) (see Supporting Information File 1).

As shown in Figure 1b and Table 1, and also in Table S2 in Supporting Information File 1, for some of the congeners, we observed slightly positive  $E_{\text{ads}}$  values in calculations and in predictions as well. Please note that all of the calculations were performed at the standard 298.15 K temperature in gas phase. More importantly, the M06-2X functional is one of the best and it is recommended for modeling of weak interactions, but still it is only an approximation of a real molecule geometry and energy. Errors obtained for this functional can fluctuate around 0.3–0.7 kcal/mol, depending on the type of calculated structure and other calculation parameters [43–45]. Furthermore, adding an error range to the obtained results will cause slightly positive  $E_{\text{ads}}$  values, which should be interpreted as congeners with practically no interaction with the fullerene surface.

**Table 1:** Predicted adsorption energies for dioxin congeners compared with official WHO TEF values.

IUPAC name	$\Delta E_{\text{ads}}$ [kcal/mol], calculated	$\Delta E_{\text{ads}}$ [kcal/mol], predicted	WHO TEF [42]
2,3,7,8-tetrachlorodibenzo- <i>p</i> -dioxin	0.958	0.733	1
1,2,3,7,8-pentachlorodibenzo- <i>p</i> -dioxin	prediction set	0.599	1
1,2,3,4,7,8-hexachlorodibenzo- <i>p</i> -dioxin	prediction set	–0.738	0.1
1,2,3,6,7,8-hexachlorodibenzo- <i>p</i> -dioxin	prediction set	–0.749	0.1
1,2,3,7,8,9-hexachlorodibenzo- <i>p</i> -dioxin	prediction set	0.095	0.1
1,2,3,4,6,7,8-heptachlorodibenzo- <i>p</i> -dioxin	prediction set	–2.132	0.01
1,2,3,4,6,7,8,9-octochlorodibenzo- <i>p</i> -dioxin	prediction set	–2.315	0.0003

It should also be noted that adsorption, as in many other reactions and processes, has an energetic barrier to overcome during the process. In this study, only the energies of single molecules of a fullerene and congeners were calculated, and the energies of the PXDD@C<sub>60</sub> complexes were calculated without obtaining the transition state energy. A small energetic barrier to overcome during the adsorption process and a slightly higher energy of the complex (comparable to the sum of single molecule energies for calculations in a gas phase (at 298.15 K)) may seem doubtful according to Hess's law (Equation 2). Additionally, it should be highlighted that all of the PXDD@C<sub>60</sub> materials are considered to be thermodynamically stable systems, and all of the Hessian matrix eigenvalues were found to be positive, so we are confident that these structures correspond to the minima on the DFT ground state potential energy surface.

## Conclusion

In conclusion, it can be stated that sorption interactions between fullerenes and halogenated dibenzo-*p*-dioxin congeners – based on weak dispersion interactions and  $\pi$ – $\pi$  stacking between – is possible. It is also strongly dependent on the type and amount of halogen substituents. Because the information about experimentally measured dioxin–fullerene sorption tendencies is still very limited [46], further investigation on the sorption mechanisms on carbon nanoparticles surfaces is essential in order to evaluate the risk, further applications and toxicity assessment. Moreover, the presented Nano-QSPR approach seems to be very helpful in this type of study and in predictions of such weak forces such as dispersion interactions.

Taking into account the inaccuracy of the computational calculations, the predicted values of  $E_{\text{ads}}$  (presented in Table S3 of Supporting Information File 1) suggest that all halogenated dibenzo-*p*-dioxin congeners will interact with the C<sub>60</sub> surface. The analysis of the obtained predictions shows that congeners that have hydrogen atoms located along the *y* axis (in 1,4,6,9 positions) and halogen atoms mainly in lateral (2,3,7,8) positions will have higher values of adsorption energy. The obtained Nano-QSPR model shows the dependency between the value of predicted  $E_{\text{ads}}$  and the dipole moment. As a result, congeners that are more chlorinated than brominated would have a higher dipole moment along the *x* axis of the dioxin structure and higher predicted values of  $E_{\text{ads}}$ . The valuable observation is that structures considered as toxic have lateral halogen atoms and will have higher  $E_{\text{ads}}$  values, unless the number of hydrogen atoms are decreased by chlorine or bromine substitution in positions 1,4,6 or 9. Keeping in mind that brominated and mixed dioxin congeners are present in the environment in mostly unknown concentrations, we hope that the presented results will be considered as a strong signal that further experimental and theoretical studies on sorption mecha-

nisms of organic pollutants on carbon nanomaterial surfaces are critical.

## Experimental

### Congener characterization

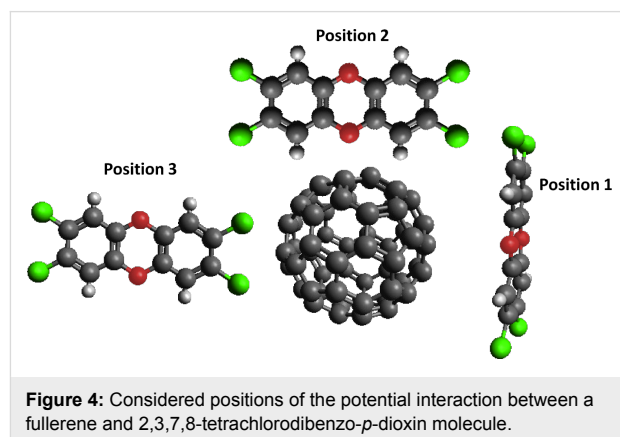
A set of 1,701 congeners containing all combinatorial possibilities of bromine, chlorine and mixed (Br/Cl) substitutions of dibenzo-*p*-dioxins was generated as a part of the Persistent Organic Pollutants Big Data project by using ConGENER software [47], and described in more detail in our previous study [48]. A set of 26 so-called molecular descriptors was calculated for each congener at the semi-empirical PM6 level. A descriptor, by definition, is an experimentally measured or calculated numerical parameter describing a particular molecule (e.g., dipole moment, number of halogen atoms, melting temperature, molecular mass, total energy). The list of 26 descriptors used in the project and more detailed information about their calculation is available in Supporting Information File 1, Table S1.

### Subset selection

A significant lack of experimental values and time-consuming calculations, which would be required for analysis of all of the 1,701 PXDDs@C<sub>60</sub> complexes, led us to select a representative subset of congeners and apply Nano-QSPR modeling of adsorption energies based on highly reliable calculations for selected compounds. A representative subset of 32 PXDDs (2% of the whole 1,701 PXDDs set) was selected by using the Kennard–Stone algorithm (KS) [49,50]. The list of selected congeners and details about the algorithm can be found in Supporting Information File 1, Table S2. This part was performed by using a script m.file [50] in MATLAB 2012 software [51].

### $\Delta E_{\text{ads}}$ calculations for the selected subset

In the first stage, three different starting positions (Figure 4) of the 2,3,7,8-tetrachlorodibenzo-*p*-dioxin molecule in the space close to the C<sub>60</sub> structure were examined to verify if the initial



dioxin position has an impact on its final placement as well as on the result of the optimization. The influence of the starting position was checked by molecular mechanics calculations. The influence of the initial distance between the dioxin and fullerene structure was also examined. Calculations on the M06-2X DFT level were performed at this stage for a 2.5–5 Å distance between the PXDD molecule and C<sub>60</sub> [52]. Supporting Information File 1 provides further details.

A M06-2X DFT functional developed by Truhlar's group is a hybrid DFT method with partially implemented experimental parameters from different databases. Since it is recognized as one of the best existing methods for showing weak interactions like Van der Waals forces and  $\pi$ - $\pi$  interactions [53–56], it was applied in the presented study. The 6-31++G(d,p) basis set was used for calculations in this part. Since for each PXDD@C<sub>60</sub> complex studied in this work all of the Hessian matrix eigenvalues were found to be positive, we are confident that these structures correspond to the minima in the DFT ground state potential energy surface. The adsorption energy ( $\Delta E_{\text{ads}}$ ) was calculated by subtracting the energy of the dioxin–fullerene system from the sum of the separated dioxin and fullerene molecules, according to Hess's law:

$$\Delta E_{\text{ads}} = \Delta E_{\text{PXDD-C60}} - (\Delta E_{\text{PXDD}} + \Delta E_{\text{C60}}). \quad (2)$$

Following the recommendations of the authors of the M06-2X approximations [53], the basis set superposition error (BSSE) was not included. The choice of theoretical methods to obtain the sorption energies for PXDD@C<sub>60</sub> complexes is reasoned because the values calculated by hybrid DFT methods, especially those calculated with M05-2X and M06-2X, are in good agreement with experimental measurements. What is more, they are appropriate to obtain weak interactions for organic compounds like pesticides or halogenated persistent organic pollutants [55,57–59]. All calculations were performed with the Gaussian 09 program [60].

## Nano-QSPR modeling

The Nano-QSPR method is based on the assumption that the variance in the physicochemical properties of compounds is determined by the variance in their molecular structures. Therefore, it is possible to predict the missing data from the calculated molecular parameters and a suitable mathematical model established for a group of similar chemicals [29]. For details of the QSPR procedure please see Supporting Information File 1.

Holland's genetic algorithm (GA) [61] was used for the selection of the optimal combination of molecular descriptors and redundancy elimination in the structural data. Partial least

squares (PLS) regression was applied as the method of modeling to solve the common problem of co-linearity within a set of descriptors. The PLS method is based on a linear transition of the original variables (descriptors) to a defined number of novel, "latent" variables (latent vectors, LVs) [62]. The use of this method usually results in well-fitted, stable models with high predictive ability [63,64]. The PLS method uses orthogonal latent vectors for regression instead of original descriptors, which is why the coefficients presented in the model equation cannot be individually interpreted. GA-PLS calculations were performed with MATLAB 2012 [51] and PLS Toolbox 7.3 [65] software packages.

To avoid overestimation and to confirm the stability and predictive ability of the developed Nano-QSPR model, a detailed validation procedure was performed, following the recommendations by the Organization for Economic Co-operation and Development (OECD) [29,66]. The details of the procedure are described in Supporting Information File 1.

## Supporting Information

### Supporting Information File 1

Adsorption of dibenzo-*p*-dioxins on the surface of C<sub>60</sub> fullerenes and calculations and QSPR predictions of the influence of halogenation.

Details about the molecular descriptor calculation method, the usage of the Kennard–Stone algorithm, and quantum mechanical calculations can be found in this file. Also, details about the development of the Nano-QSPR model and its statistical characterization are described. Predicted adsorption energies for all chlorinated and/or brominated dibenzo-*p*-dioxin congeners are provided for training and validation sets and for the prediction set.

[<http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-8-78-S1.pdf>]

## Acknowledgements

This material is based on research sponsored by the Polish National Science Center (grant no. UMO-2011/01/M/NZ7/01445). M.H. acknowledges support from the Spanish Ministry of Economy and Competitiveness (RYC-2013-13949). The authors want to acknowledge CI TASK and MCSR OLEMISS computational centers. The authors also want to acknowledge Dr. Lidia Chomicz (University of Gdansk), Dr. Bakhtior Rasulev (Jackson State University) and Prof. Janusz Rak (University of Gdansk) for scientific consultations and advices.

The authors contributed to the manuscript as follows: P.U. and T.P. conceived and designed the experiments; M.H. performed



the molecular descriptor calculations for PXDD congeners; P.U. performed the quantum mechanical calculations; C.S. consulted the quantum mechanical calculations and helped in the results discussion. P.U. developed the Nano-QSPR model and performed the chemometric analysis; A.G. supervised the Nano-QSPR modeling and was a helpful consultant during the interpretation of the obtained results.

## References

- Lerche, D.; van de Plassche, E.; Schwegler, A.; Balk, F. *Chemosphere* **2002**, *47*, 617–630. doi:10.1016/S0045-6535(02)00028-0
- Bruzy, L. P.; Hites, R. A. *Environ. Sci. Technol.* **1995**, *29*, 2090–2098. doi:10.1021/es00008a031
- Falandysz, J.; Rose, M.; Fernandes, A. R. *Environ. Int.* **2012**, *44*, 118–127. doi:10.1016/j.envint.2012.03.006
- Fernandes, A.; Mortimer, D.; Gem, M.; Dicks, P.; Smith, F.; White, S.; Rose, M. *Food Addit. Contam., Part A* **2009**, *26*, 918–927. doi:10.1080/02652030902803026
- Fernandes, A. R.; Rose, M.; Mortimer, D.; Carr, M.; Panton, S.; Smith, F. *J. Chromatogr. A* **2011**, *1218*, 9279–9287. doi:10.1016/j.chroma.2011.10.058
- Tuppurainen, K.; Ruuskanen, J. *Chemosphere* **2000**, *41*, 843–848. doi:10.1016/S0045-6535(99)00525-1
- Samara, F.; Wyrzykowska, B.; Tabor, D.; Touati, D.; Gullett, B. K. *Environ. Int.* **2010**, *36*, 247–253. doi:10.1016/j.envint.2009.12.005
- Safe, S. H. *Annu. Rev. Pharmacol.* **1986**, *26*, 371–399. doi:10.1146/annurev.pa.26.040186.002103
- Landers, J. P.; Bunce, N. J. *Biochem. J.* **1991**, *276*, 273–287. doi:10.1042/bj2760273
- Landers, J. P.; Winhall, M. J.; McCready, T. L.; Sanders, D. A. R.; Rasper, D.; Nakai, J. S.; Bunce, N. J. *J. Biol. Chem.* **1991**, *266*, 9471–9480.
- Unger, M.; Asplund, L.; Haglund, P.; Malmvärn, A.; Arnoldsson, K.; Gustafsson, Ö. *Environ. Sci. Technol.* **2009**, *43*, 8245–8250. doi:10.1021/es901705r
- Dresselhaus, M. S.; Dresselhaus, G.; Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes*; Academic Press: San Diego, 1996. doi:10.1016/B978-012221820-0/50000-9
- Jensen, A. W.; Wilson, S. R.; Schuster, D. I. *Bioorg. Med. Chem.* **1996**, *4*, 767–779. doi:10.1016/0968-0896(96)00081-8
- Chowdhury, S.; Balasubramanian, R. *Adv. Colloid Interface Sci.* **2014**, *204*, 35–56. doi:10.1016/j.cis.2013.12.005
- Kroto, H. W.; Heath, J. R.; O'Brien, S. C.; Curl, R. F.; Smalley, R. E. *Nature* **1985**, *318*, 162–163. doi:10.1038/318162a0
- Sayes, C. M.; Fortner, J. D.; Guo, W.; Lyon, D.; Boyd, A. M.; Ausman, K. D.; Tao, Y. J.; Sitharaman, B.; Wilson, L. J.; Hughes, J. B.; West, J. L.; Colvin, V. L. *Nano Lett.* **2004**, *4*, 1881–1887. doi:10.1021/nl0489586
- Park, S.; Srivastava, D.; Cho, K. J. *Nanotechnology* **2001**, *12*, 245–249. doi:10.1088/0957-4484/12/3/308
- Da Ros, T.; Spalluto, G.; Prato, M. *Croat. Chem. Acta* **2001**, *74*, 743–755. <http://hrcak.srce.hr/131951>
- Sherigara, B. S.; Kutner, W.; D'Souza, F. *Electroanalysis* **2003**, *15*, 753–772. doi:10.1002/elan.200390094
- Keiluweit, M.; Kleber, M. *Environ. Sci. Technol.* **2009**, *43*, 3421–3429. doi:10.1021/es8033044
- Adams, H.; Cockcroft, S. L.; Guardigli, C.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Spey, S. E.; Urch, C. J.; Ford, R. *ChemBioChem* **2004**, *5*, 657–665. doi:10.1002/cbic.200400018
- Linkov, I.; Steevens, J.; Adlakha-Hutcheon, G.; Bennett, E.; Chappell, M.; Colvin, V.; Davis, J. M.; Davis, T.; Elder, A.; Hansen, S.; Hakkinen, P. B.; Hussain, S. M.; Karkan, D.; Korenstein, R.; Lynch, I.; Metcalfe, C.; Ramadan, A. B.; Satterstrom, F. K. *J. Nanopart. Res.* **2009**, *11*, 513–527. doi:10.1007/s11051-008-9514-9
- Yang, K.; Zhu, L. Z.; Xing, B. S. *Environ. Sci. Technol.* **2006**, *40*, 1855–1861. doi:10.1021/es052208w
- Zhu, D. Q.; Pignatello, J. J. *Environ. Sci. Technol.* **2005**, *39*, 2033–2041. doi:10.1021/es0491376
- Zhang, S. J.; Shao, T.; Kose, H. S.; Karanfil, T. *Environ. Sci. Technol.* **2010**, *44*, 6377–6383. doi:10.1021/es100874y
- Gajewicz, A.; Rasulev, B.; Dinadayalane, T. C.; Urbaszek, P.; Puzyn, T.; Leszczynska, D.; Leszczynski, J. *Adv. Drug Delivery Rev.* **2012**, *64*, 1663–1693. doi:10.1016/j.addr.2012.05.014
- OECD Principles for the validation, for regulatory purposes, of (quantitative) structure-activity relationship models. <http://www.oecd.org/dataoecd/33/37/37849783.pdf> (accessed Jan 22, 2007).
- Tropsha, A.; Gramatica, P.; Gombar, V. K. *QSAR Comb. Sci.* **2003**, *22*, 69–77. doi:10.1002/qsar.200390007
- Gramatica, P. *QSAR Comb. Sci.* **2007**, *26*, 694–701. doi:10.1002/qsar.200610151
- Gramatica, P.; Cassani, S.; Roy, P. P.; Kovarich, S.; Yap, C. W.; Papa, E. *Mol. Inf.* **2012**, *31*, 817–835. doi:10.1002/minf.201200075
- Eriksson, L.; Jaworska, J.; Worth, A. P.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P. *Environ. Health Perspect.* **2003**, *111*, 1361–1375.
- Netzeva, T. I.; Worth, A. P.; Aldenberg, T.; Benigni, R.; Cronin, M. T. D.; Gramatica, P.; Jaworska, J. S.; Kahn, S.; Klopman, G.; Marchant, C. A.; Myatt, G.; Nikolova-Jeliazkova, N.; Patlewicz, G. Y.; Perkins, R.; Roberts, D. W.; Schultz, T. W.; Stanton, D. T.; van de Sandt, J. J. M.; Tong, W. D.; Veith, G.; Yang, C. H. Current Status of Methods for Defining the Applicability Domain of (Quantitative) Structure–Activity Relationships. *The Report and Recommendations of ECVAM Workshop 52, ATLA 33*; 2005; pp 155–173.
- Gramatica, P.; Papa, E. *QSAR Comb. Sci.* **2005**, *24*, 953–960. doi:10.1002/qsar.200530123
- Benyamini, H.; Shulman-Peleg, A.; Wolfson, H. J.; Belgorodsky, B.; Fadeev, L.; Gozin, M. *Bioconjugate Chem.* **2006**, *17*, 378–386. doi:10.1021/bc050299g
- Mitra, R.; Bauri, A. K.; Bhattacharya, S. *Spectrochim. Acta, Part A* **2012**, *96*, 485–492. doi:10.1016/j.saa.2012.05.040
- Vallant, R. M.; Szabo, Z.; Bachmann, S.; Bakry, R.; Najam-ul-Haq, M.; Rainer, M.; Heigl, N.; Petter, C.; Huck, C. W.; Bonn, G. K. *Anal. Chem.* **2007**, *79*, 8144–8153. doi:10.1021/ac0712392
- Nie, B.; Rotello, V. J. *Phys. Chem. Solids* **1997**, *58*, 1897–1899. doi:10.1016/S0022-3697(97)00083-8
- Vejerano, E. P.; Holder, A. L.; Marr, L. C. *Environ. Sci. Technol.* **2013**, *47*, 4866–4874. doi:10.1021/es304895z
- Shukla, M. K.; Dubey, M.; Zakar, E.; Namburu, R.; Czyznikowska, Z.; Leszczynski, J. *Chem. Phys. Lett.* **2009**, *480*, 269–272. doi:10.1016/j.cplett.2009.09.031
- Alparone, A.; Librando, V.; Minniti, Z. *Chem. Phys. Lett.* **2008**, *460*, 151–154. doi:10.1016/j.cplett.2008.05.028
- Gritti, F.; Guiochon, G. *J. Chromatogr. A* **2004**, *1053*, 59–69. doi:10.1016/S0021-9673(04)00978-1

42. Van den Berg, M.; Birnbaum, L. S.; Denison, M.; De Vito, M.; Farland, W.; Feeley, M.; Fiedler, H.; Hakansson, H.; Hanberg, A.; Haws, L.; Rose, M.; Safe, S.; Schrenk, D.; Tohyama, C.; Tritscher, A.; Tuomisto, J.; Tysklind, M.; Walker, N.; Peterson, R. E. *Toxicol. Sci.* **2006**, *93*, 223–241. doi:10.1093/toxsci/kfl055
43. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101. doi:10.1063/1.2370993
44. Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 5121–5129. doi:10.1021/jp060231d
45. Pieniazek, S. N.; Clemente, F. R.; Houk, K. N. *Angew. Chem., Int. Ed.* **2008**, *47*, 7746–7749. doi:10.1002/anie.200801843
46. Richarz, A.-N.; Cronin, M.; Madden, J.; Lubinski, L.; Mokshina, E.; Urbaszek, P.; Puzyn, T.; Kuz'min, V. *Toxicol. Lett.* **2013**, *221*, S246. doi:10.1016/j.toxlet.2013.05.609
47. Haranczyk, M.; Puzyn, T.; Sadowski, P. *QSAR Comb. Sci.* **2008**, *27*, 826–833. doi:10.1002/qsar.200710149
48. Haranczyk, M.; Urbaszek, P.; Ng, E. G.; Puzyn, T. *J. Chem. Inf. Model.* **2012**, *52*, 2902–2909. doi:10.1021/ci300289b
49. Kennard, R. W.; Stone, L. A. *Technometrics* **1969**, *11*, 137–148. doi:10.1080/00401706.1969.10490666
50. Daszykowski, M.; Walczak, B.; Massart, D. L. *Anal. Chim. Acta* **2002**, *468*, 91–103. doi:10.1016/S0003-2670(02)00651-7
51. *Matlab 2013b*; The MathWorks Inc., <http://www.mathworks.com>.
52. Puzyn, T.; Suzuki, N.; Haranczyk, M.; Rak, J. *J. Chem. Inf. Model.* **2008**, *48*, 1174–1180. doi:10.1021/ci800021p
53. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *120*, 215–241. doi:10.1007/s00214-007-0310-x
54. Scott, A. M.; Gorb, L.; Mobley, E. A.; Hill, F. C.; Leszczynski, J. *Langmuir* **2012**, *28*, 13307–13317. doi:10.1021/la3027286
55. Michalkova, A.; Gorb, L.; Hill, F.; Leszczynski, J. *J. Phys. Chem. A* **2011**, *115*, 2423–2430. doi:10.1021/jp112240d
56. Gu, J. D.; Wang, J.; Leszczynski, J.; Xie, Y. M.; Schaefer, H. F., III. *Chem. Phys. Lett.* **2008**, *459*, 164–166. doi:10.1016/j.cplett.2008.05.049
57. Arnold, W. A. *Environ. Sci.: Processes Impacts* **2013**, *16*, 832–838. doi:10.1039/C3EM00479A
58. da Silva, G. J. *J. Phys. Chem. A* **2012**, *116*, 5317–5324. doi:10.1021/jp303806w
59. de la Luz, A. P.; Iuga, C.; Alvarez-Idaboy, J. R.; Ortiz, E.; Vivier-Bunge, A. *Int. J. Quantum Chem.* **2012**, *112*, 3525–3534. doi:10.1002/qua.24159
60. *Gaussian 09*, Revision b.01; Gaussian, Inc.: Wallingford CT, 2009.
61. Holland, J. H. *Sci. Am.* **1992**, *267*, 66–72. doi:10.1038/scientificamerican0792-66
62. Wold, S.; Sjöström, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. doi:10.1016/S0169-7439(01)00155-1
63. Gajewicz, A.; Haranczyk, M.; Puzyn, T. *Atmos. Environ.* **2010**, *44*, 1428–1436. doi:10.1016/j.atmosenv.2010.01.041
64. Chen, J. W.; Quan, X.; Zhao, Y. Z.; Yang, F. L.; Schramm, K.-W.; Kettrup, A. *Bull. Environ. Contam. Toxicol.* **2001**, *66*, 755–761. doi:10.1007/s001280073
65. *PLS toolbox*, 7.3; Eigenvector Research Inc., 2013.
66. Organisation for Economic Co-operation and Development, Environment Directorate. *Guidance document on the validation of (quantitative) structure-activity relationships [(Q)SAR] models*; Mar 2007.

## License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The license is subject to the *Beilstein Journal of Nanotechnology* terms and conditions: (<http://www.beilstein-journals.org/bjnano>)

The definitive version of this article is the electronic one which can be found at:  
doi:10.3762/bjnano.8.78