

Regular Article

## Criterion of molecular size to evaluate the bioaccumulation potential of chemicals in fish

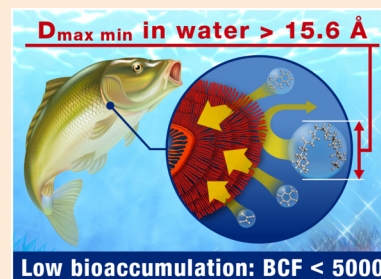
Chiyoiko Miyata,<sup>1,2,\*</sup> Yoshihide Matoba,<sup>1</sup> Makiko Mukumoto,<sup>1</sup> Yoshiaki Nakagawa<sup>2</sup> and Hisashi Miyagawa<sup>2</sup>

<sup>1</sup> Sumitomo Chemical Co., Ltd. Environmental Health Science Laboratory, 1–98 Kasugadenaka 3-chome, Konohana-ku, Osaka 554–8558, Japan

<sup>2</sup> Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Kyoto 606–8502, Japan

(Received June 6, 2021; Accepted October 27, 2021)

To evaluate the bioaccumulation potential of chemicals in fish, a molecular-size descriptor,  $D_{\max \text{ aver}}$  has been used as a weight of evidence under the EU REACH. The  $D_{\max \text{ aver}}$  value, however, is estimated on the basis of 3-D structures of possible stable conformers in a vacuum using OASIS software that requires expertise upon parameter input. We developed a method to calculate the 3-D conformers in water, which is more suitable for bioaccumulation potential evaluation in an aquatic environment, by introducing MD simulation. By examining the relationship of the calculated molecular size of 1665 chemicals with their reported BCF values, we found that 17.1 Å of  $D_{\max \text{ aver}}$  or 15.6 Å of  $D_{\max \text{ min}}$  was a threshold of molecular size in water to predict the low bioaccumulation (*i.e.*,  $\text{BCF} < 5000$ ) of a chemical. Setting this threshold as a new standard would reduce the number of animal tests without compromising the quality of safety evaluation.



**Keywords:** molecular size, quantum chemical calculation, molecular dynamics simulation, bioaccumulation, bioconcentration.

### Introduction

Persistent, bioaccumulative, and toxic chemicals (PBTs) and persistent organic pollutants (POPs) are controlled under various national and international regulatory frameworks, including the Stockholm Convention,<sup>1,2)</sup> due to their negative effects on the environment. Bioaccumulation is the most critical issue for humans and environmental organisms from the viewpoint of the food chain. In Japan, under the Japanese Chemical Substances Control Law (CSCL), since 1974, an aqueous bioconcentration test has been required in principle for every chemical.<sup>3)</sup> In the EU, the bioconcentration test is required for chemicals produced in quantities greater than 100 tons per year under REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) unless the chemical has a low potential for bioaccumula-

tion. Those having a low potential to permeate biological membranes or little chance of direct and indirect exposure by living organisms in the aquatic environment are also exempted from the bioconcentration test.<sup>4,5)</sup> This test is conducted to determine the bioconcentration factor (BCF) of a chemical in fish. The BCF is defined by the ratio of the steady-state concentration of a chemical in fish to that in water after the fish take the chemical in water through gill membranes during a specific time period.<sup>6,7)</sup> If a chemical has a  $\text{BCF} \geq 5000$ , it is regarded as highly bioaccumulative under the Stockholm Convention and the CSCL.<sup>1)</sup>

Currently, typical protocols of the bioconcentration test are relatively laborious and time consuming. They also require large numbers of experimental fish, which is undesirable from an animal welfare standpoint. Thus, alternative methods of assessing the bioaccumulation potential of chemicals have been applied using their physicochemical parameters, such as the octanol–water partition coefficient and the molecular weight (MW).<sup>8)</sup> Some of these methods focus on the permeation process through biological membranes when organisms take in a chemical, and they assume its molecular size as an uptake-constraint factor.<sup>9–12)</sup> The basic idea is that bulky molecules cannot permeate the membranes: for example, a chemical with  $\text{MW} > 800$  is approved as low bioaccumulative, and a bioconcentration study

\* To whom correspondence should be addressed.

E-mail: miyata@sc.sumitomo-chem.co.jp

Published online January 11, 2022

is not required under the CSCL.<sup>2,13)</sup> Mekenyan *et al.* have proposed a  $D_{\max\text{ aver}}$  value as one of the steric factors to characterize the bulkiness of a molecule<sup>14,15)</sup>;  $D_{\max\text{ aver}} > 17.4 \text{ \AA}$  is approved as one indicator in a weight of evidence assessment, a method for making decisions by considering multiple sources of information,<sup>16)</sup> for limited bioaccumulation potential (*i.e.*, BCF < 5000) under REACH.<sup>17)</sup>

The  $D_{\max\text{ aver}}$  value is an arithmetic mean diameter of the smallest spheres accommodating the locally stable conformers. This value can be obtained only by using commercial software OASIS, developed by Mekenyan *et al.*<sup>17,18)</sup> Although the  $D_{\max\text{ aver}}$  value is a leading indicator to describe the steric property of a chemical, this indicator alone is considered insufficient to exempt a chemical from bioconcentration testing.<sup>16,17)</sup> Due to some problems, countries other than the EU have not approved it. The most critical problem appears to be that OASIS estimates the  $D_{\max\text{ aver}}$  value of a chemical based on its conformations in a vacuum,<sup>14,15)</sup> although the bioaccumulation in fish occurs in an aquatic environment. Since the simulation of molecular structures in water requires enormous computational cost in a quantum chemical calculation that considers the interactions with surrounding water molecules, Mekenyan and coauthors selected locally stable conformers within 20 kcal/mol of the most stable one in a vacuum in the OASIS calculation to consider the structural differences between in a vacuum and in water.<sup>15)</sup> However, scientific discussions have not validated this selection, such as a comparison with actual  $D_{\max\text{ aver}}$  of conformers in water. Moreover, in calculating  $D_{\max\text{ aver}}$  values with OASIS software, expert knowledge is required for inputting parameters in the genetic algorithm, such as the size of the permanent population, the number of children, the mutation/crossover ratio, and the generation times.<sup>19)</sup>

In this study, we developed a new method of calculating  $D_{\max\text{ aver}}$  values in a vacuum that requires no expert knowledge. The method combined a commercial program to generate a set of energetically stable conformers in a vacuum with an original program to calculate  $D_{\max\text{ aver}}$  from the generated conformer set. Then, we reproduced the dynamic behavior of molecules in water using the molecular dynamics (MD) simulation method, and the calculated  $D_{\max\text{ aver}}$  values were compared with those in a vacuum. Finally, based on reliable BCF values of 1665 chemicals, we proposed a new threshold of  $D_{\max}$  value in water to judge the low bioaccumulation potential (*i.e.*, BCF < 5000) of chemicals in an aquatic environment.

## Materials and methods

### 1. Calculation method

#### 1.1. Conformers in a vacuum

An initial 3-D structure of a test chemical was drawn using CHEM3D Ver.18 (PerkinElmer Informatics, USA). This initial structure was submitted to the Monte Carlo Multiple Minimum (MCM) search algorithm<sup>20)</sup> of HyperChem 8.0.6. (HyperCube, USA) to generate 3-D conformers by a quantum chemical calculation, in which the dihedral angles of each molecule were ran-

domly selected and rotated by every 120 degrees. The quantum calculation was stopped if the number of the generated structures exceeded 5 million. The number of 5 million was determined by considering the calculation cost. It covers that of all conformers having 14 dihedral angles in a molecule where every angle was rotated by every 120 degrees ( $3^{14}=4,782,969$ ). To verify whether the number of 5 million was sufficient, five different initial 3-D structures with  $D_{\max}$  values different from each other by  $>3 \text{ \AA}$  were submitted to HyperChem as very flexible chemicals with more than 14 dihedral angles.

The free energy gradient method was applied to the geometry optimization to obtain a locally stable conformer. This optimization was performed in two steps. In the first step, the conformers generated by MCM were roughly optimized by the molecular mechanics method using an MM+ force field<sup>21)</sup> with a Polak–Ribiere Conjugate Gradient.<sup>22)</sup> In the second step, each of the locally stable conformers by the molecular mechanics method was re-optimized by the semi-empirical molecular orbital method using AM1/6-31G+<sup>15)</sup> with the Polak–Ribiere Conjugate Gradient. The root-mean-square gradient convergence threshold of the Polak–Ribiere Conjugate Gradient was set to 0.01 kcal/Å mol in each optimization.

After these optimizations, locally stable conformers with potential energies 20 kcal/mol higher than that of the most stable one were excluded in accordance with the method for  $D_{\max\text{ aver}}$  calculation in OASIS.<sup>17)</sup> Conformers with a similar molecular size were also excluded. The molecular similarity was validated by a root-mean-square distance calculation between the two atoms at both ends of the conformers. When differences in the root-mean-square distance were less than 0.1 Å, these conformers were considered similar, and one of them was selected to exclude others. These 3-D conformers obtained by HyperChem are described as “complete locally stable conformers” in the following sections.

The locally stable conformers that emerged around the gentle bottom in the same valley on the potential energy surface were further excluded by an original Excel program using Visual Basic for Applications (VBA) because these similar conformers were likely to provide a biased weight in calculating the  $D_{\max\text{ aver}}$  value. In the program, the root-mean-square distances between all corresponding two atoms of a pair of locally stable conformers were calculated according to Eq. (1),

$$d = \sqrt{\sum_{i,j} (\Delta x_{ij}^A - \Delta x_{ij}^B)^2} \quad (1)$$

where  $\Delta x_{ij}^A$  and  $\Delta x_{ij}^B$  indicate the distance between atoms  $i$  and  $j$  for locally stable conformers A and B, respectively. A smaller  $d$  value indicates a higher similarity between two locally stable conformers. Thus, of similar conformers within a small  $d$  value, one representative conformer was selected to exclude others from the complete locally stable conformers. The maximum number of conformers for  $D_{\max}$  calculation was changed from 10 to 200 in this study by varying the  $d$  value, and  $D_{\max\text{ aver}}$  val-

ues based on the selected conformers were compared with those given by OASIS. The conformers selected by specifying the  $d$  value are described as “filtered stable conformers.”

### 1.2. Conformers in water

Three-dimensional structures of each test chemical in water were generated by MD simulations using Material Studio 2018 (BIOVIA, USA). The simulation was carried out in a cube containing 2000 water molecules with a density of 1.0 g/cm<sup>3</sup> by the Forcite module. The cubic size was set at 39.1 Å in length, which was twofold longer than the limit of the  $D_{\max \text{ aver}}$  of the low bioaccumulation potential (*i.e.*, 17.4 Å). The initial 3-D structure was drawn using CHEM3D. The cubic box containing the test chemical and water molecules was energy-minimized by the Smart algorithm in the Forcite module with a Condensed-Phase Optimized Molecular Potentials for Atomistic Simulation Studies (COMPASS) II force field.<sup>23)</sup> Convergence tolerance for minimization was Medium, with energy of 0.001 kcal/mol, a maximum force of 0.5 kcal/mol·Å, and a maximum displacement of 0.015 Å. After the energy minimization, the MD simulation was performed with a time step of 1.0 fs by using an NTP ensemble in the Forcite module of Material Studio, where a Nosé-Hoover thermostat and a Berendsen barostat<sup>24–26)</sup> were used to keep the temperature and pressure at 298 K and 1×10<sup>5</sup> Pa, respectively. The MD simulations in water for a flexible chemical were performed for 10, 100, 200, 500, and 2000 ps. The snapshots were collected every 1 ps. The 3-D coordinates of all of the atoms in each conformer were used for the following  $D_{\max}$  calculation. The effect of the simulation time was evaluated by the repeatability of the calculated  $D_{\max}$  value and of the range of fluctuation during calculation.

### 1.3. Calculation of $D_{\max}$ values

The coordinates of all atoms of the conformers calculated in sections 1.1 and 1.2 were used to calculate the diameter of a sphere covering the van der Waals radii<sup>27)</sup> of all atoms in a conformer (*i.e.*,  $D_{\max}$ ) using an original Excel program with VBA. In this program, the center of the sphere was first determined,<sup>28)</sup> and then the distance between the sphere's center and the edge of the van der Waals radius of the farthest atom was calculated. The  $D_{\max}$  value was obtained by doubling the distance. The  $D_{\max}$  value was calculated for each of the complete locally stable conformers in a vacuum, of the filtered stable conformers in vacuum, or of the conformers in water.  $D_{\max \text{ aver}}$ ,  $D_{\max \text{ min}}$ , and  $D_{\max \text{ max}}$  values of a test chemical were determined by averaging the  $D_{\max}$  values arithmetically and by selecting the minimum and maximum among the  $D_{\max}$  values, respectively.

The center of the sphere was determined as follows: a moving point was first set at the origin of the 3-D coordinate, and it was moved toward the farthest atom ( $x_1$ ) from the origin by a distance of  $(h_{x_1}+v_{x_1})/2$ , where  $h_{x_1}$  was the distance between the origin and the center of the farthest atom ( $x_1$ ), and  $v_{x_1}$  was the van der Waals radius of the atom. Next, the point was further moved from there toward the farthest atom ( $x_2$ ) by a distance of  $(h_{x_2}+v_{x_2})/2$ , where  $h_{x_2}$  was the distance between the arrival point and the center of  $x_2$ , and  $v_{x_2}$  was the van der Waals radius of

$x_2$ . This operation was repeated  $y$  times. Then the moving distance was changed to  $(h_{x_i}+v_{x_i})/2^2$ , and the point was moved by  $(h_{x_i}+v_{x_i})/2^2$  to the atom farthest from the arrival point by repeating. This operation was repeated  $y$  times. In this way, the point gradually reached the true center of the sphere by decreasing the moving distance to  $(h_{x_i}+v_{x_i})/2^n$ . The number of  $y$  was set at 20, based on the convergence conditions investigated for this method. We judged that the point had arrived at the center when the moving distance got to less than  $1 \times 10^{-6}$  Å.

## 2. Test chemicals

$D_{\max}$  values of 69 chemicals of various sizes and shapes were calculated for the conformers in a vacuum or water (Table 1). These chemicals were described in a report published by the UK Environmental Agency<sup>18)</sup> and consisted of biphenyl ethers, polychlorinated alkanes, polysulfides, styrylphenols, primary alcohols, pigments, and 29 chemicals with various skeletons.

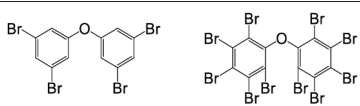
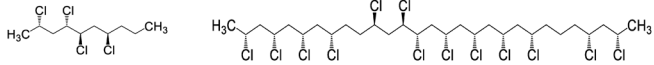
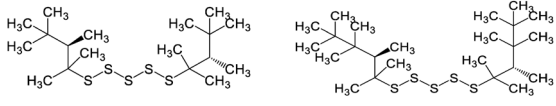
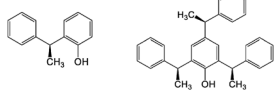
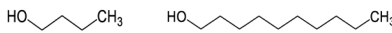
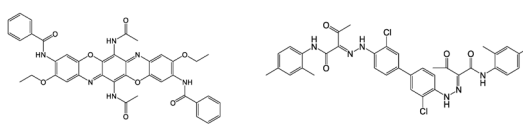
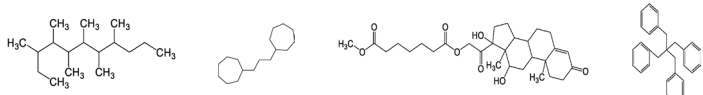
Test chemicals with measured BCF values were collected from the database of the Japanese National Institute of Technology and Evaluation (NITE). Bioconcentration tests with 750 chemicals were conducted by the Japanese Ministry of Economy, Trade and Industry (METI) from 1974 to 1986,<sup>29)</sup> and tests with 1118 chemicals were conducted by applicants under the CSCL from 1987 to 2017.<sup>30)</sup> These tests were performed under the Organisation for Economic Co-operation and Development (OECD) Principles of Good Laboratory Practice<sup>31)</sup> following OECD Test Guideline 305 (I).<sup>6)</sup> The BCF values were used as described in these test results, except for the following cases: When more than two BCF values were available for a single chemical, the highest was selected for conservative analysis. When a BCF value was given for a mixture of different chemicals, the original reports and references were examined, and only the value that could be assigned to a single component was used. Otherwise, the values were excluded from the analysis. Inorganic chemicals and heavy metal-containing organic chemicals were excluded from the investigation because their uptake mechanism to fish differs from passive diffusion through the gill membrane.<sup>32,33)</sup> For ionic chemicals in water, the dissociated structures were used for the calculation to avoid overestimating their molecular size. Perfluoroalkyl chemicals have been reported to bioaccumulate in organisms with the assistance of specific binding proteins in the blood.<sup>34–36)</sup> For these substances,  $D_{\max}$  values are not relevant as the uptake-constraint factor; hence, they were excluded from the investigation. Some chemicals (BCF <100 and MW <100) were excluded from the calculation because they were too small and not relevant to concerns of high bioaccumulation. As a result,  $D_{\max}$  values in water were determined for 1665 chemicals.

## Results and discussion

### 1. $D_{\max}$ values in a vacuum

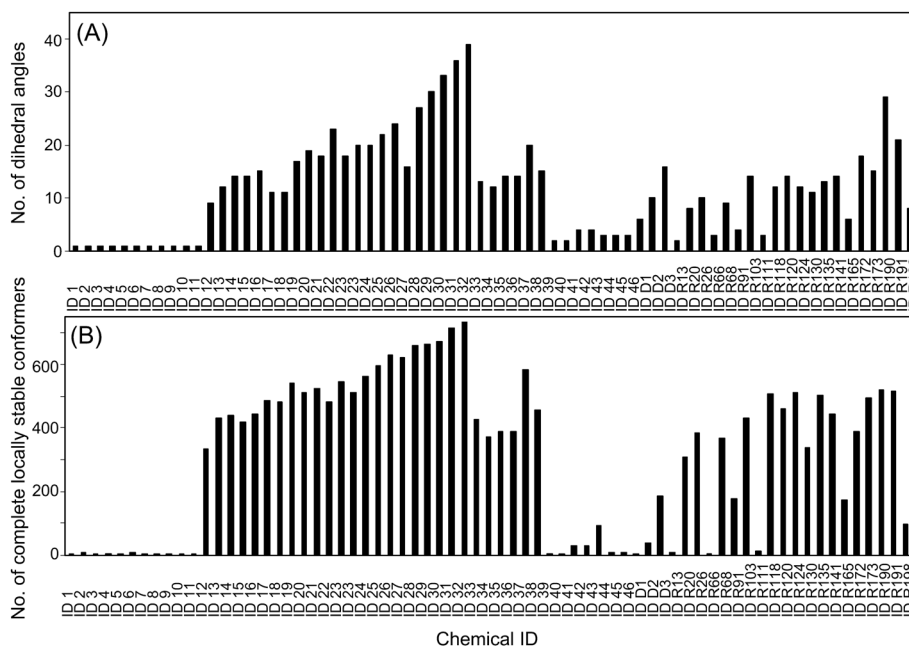
$D_{\max}$  values in a vacuum were estimated by HyperChem and the original program for 69 chemicals listed in the UK report.<sup>18)</sup> The total number of rotational dihedral angles in each chemical ranged from 1 to 39 (Fig. 1A). Five minutes to two weeks

**Table 1.** Classification and typical structures of 69 chemicals<sup>18)</sup> used to calculate  $D_{\max}$  values in a vacuum or in water

Chemical groups	Typical chemical structures	No. of chemicals
Biphenyl ethers		11
Polychlorinated alkanes		22
Polysulphides		6
Styrylphenols		5
Primary alcohols		2
Pigments		3
Other chemicals		29

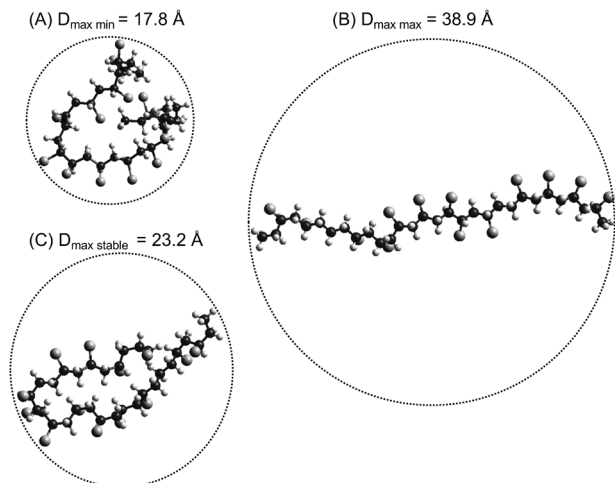
were required to obtain the complete locally stable conformers by changing the respective dihedral angles. The numbers of the complete locally stable conformers ranged from 1 to 717 (Fig. 1B), depending on the number of the dihedral angles in the molecule that gave rise to higher freedom of motion.

A long-chain chlorinated alkane ID33 ( $C_{30}H_{49}Cl_{13}$ ) has the highest rotational freedom and, therefore, has the largest number of 3-D conformers among the chemicals examined.  $D_{\max}$  values ranged from  $D_{\max \min} = 17.8 \text{ \AA}$  (the smallest) to  $D_{\max \max} = 38.9 \text{ \AA}$  (the largest), as shown in Fig. 2. The most stable



**Fig. 1.** Numbers of dihedral angles (A) and the complete locally stable conformers (B) for 69 chemicals with ID numbers described in the UK report.<sup>18)</sup>





**Fig. 2.** Three-dimensional models of locally stable conformers and their  $D_{\max}$  values for chemical ID33 ( $C_{30}H_{49}Cl_{13}$ ).

conformer of ID33 has a  $D_{\max \text{ stable}}$  value of 23.2 Å, whereas the average  $D_{\max}$  value ( $D_{\max \text{ aver}}$ ) was 26.4 Å, based on the complete locally stable conformers.

The initial 3-D structure of each chemical was submitted to HyperChem to generate a set of up to 5 million conformers. In order to verify whether the number of 5 million was sufficient, flexible long-chain chlorinated alkanes (*i.e.*, ID29 to ID33) with more than 14 dihedral angles in a molecule were selected from the 69 test chemicals. Five initial structures that had  $D_{\max}$  values different from each other by  $>3$  Å were submitted to HyperChem to generate the stable conformers and calculate their  $D_{\max \text{ aver}}$  values for ID29 to ID33. Table 2 shows the ranges of the number of the complete locally stable conformers, energy of the most stable conformer, and  $D_{\max \text{ aver}}$  values calculated from the complete locally stable conformers, which were derived from five different input structures of each test chemical. While the number of the complete locally stable conformers varied depending on the shape of the potential energy surface and/or the convergence conditions of each input structure, the differences in the energies and  $D_{\max \text{ aver}}$  values were within 0.3 kcal/mol and  $\pm 0.04$  Å, respectively. As a consequence, we confirmed that 5 million was an appropriate number of random conformer generation even for test chemicals with more than 14 dihedral angles, and the  $D_{\max \text{ aver}}$  value was independent of a structure input into HyperChem. This high repeatability suggests that everyone can obtain the same results

using this method without any expert knowledge.

For the 69 chemicals written about in the UK report,  $D_{\max}$  values calculated from the complete locally stable conformers were correlated with those by OASIS<sup>18)</sup> with a slope of linear regression equation of 0.992 and a coefficient of determination ( $R^2$ ) of 0.985 (Fig. 3A). However, when the potential energy surface gradient was gradual, some similar conformers could converge into different multiple locally stable conformers satisfying the convergence threshold (*i.e.*, 0.01 kcal/Å mol). If such conformers were deemed to be independent, their  $D_{\max}$  values could substantially affect the calculation of  $D_{\max \text{ aver}}$  making it unreliable by overweighting their contribution. To eliminate the effect of highly similar conformers on the calculation of  $D_{\max \text{ aver}}$  the filtering conditions for excluding them were examined. By adjusting the  $d$  value as defined by Eq. (1), the maximum number of filtered stable conformers was changed from 10 to 200. The maximum number of complete locally stable conformers in this chemical set was 717. The calculation of  $D_{\max \text{ aver}}$  for each set of filtered stable conformers revealed that 100 conformers had the highest agreement with  $D_{\max \text{ aver}}$  given by OASIS, where the slope of the linear regression equation was 0.996, and the coefficient of determination ( $R^2$ ) was 0.996 (Fig. 3C). When the conformers were filtered to 200, the degree of agreement was lower, probably due to the incomplete removal of similar conformers (Fig. 3D). On the other hand, when the number of filtered conformers was less than 50, some of the structures that importantly contributed to  $D_{\max \text{ aver}}$  were likely excluded, also lowering the degree of agreement (Fig. 3B shows the result using 30 conformers). The differences between the  $D_{\max \text{ aver}}$  values calculated from the 100 filtered stable conformers and those calculated by OASIS were within  $\pm 0.60$  Å for the 69 chemicals, while, taking the complete locally stable conformers into consideration, the differences were within  $\pm 1.17$  Å. Thus, the maximum number of filtered stable conformers for each chemical was set to 100 in the following investigations.

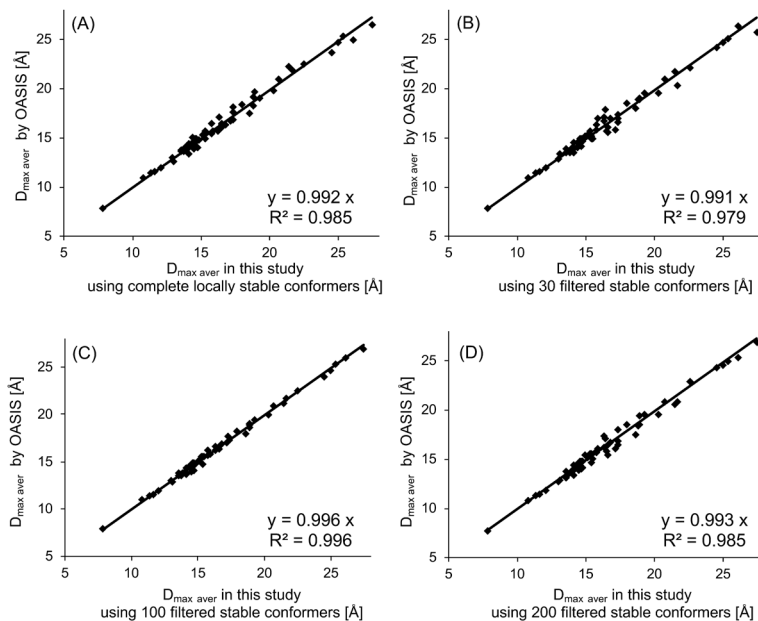
To summarize the results in a vacuum, the computational parameters are clearly defined and unequivocal in this study, and the calculation results are not affected by the initial conformation of a chemical. Everyone can perform the same quality of calculations as OASIS without expertise, which is definitely advantageous.

## 2. $D_{\max}$ values in water

The effect of simulation time on the  $D_{\max}$  calculation was examined using the most flexible chemical ID33 ( $C_{30}H_{49}Cl_{13}$ ) among

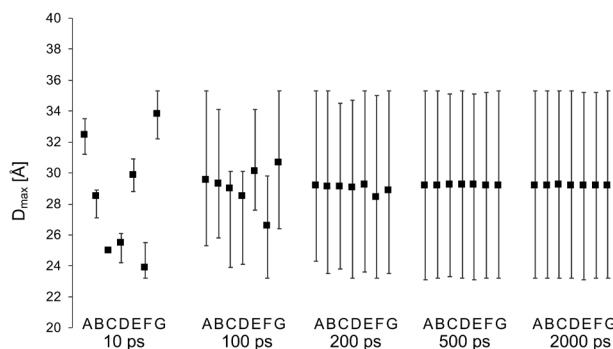
**Table 2.** Repeatability of  $D_{\max \text{ aver}}$  calculation in a vacuum by HyperCHEM starting from five different 3-D structures of each test chemical

Chemical ID	Formula	No. of dihedral angles	No. of complete locally stable conformers	Energy of most stable conformer (kcal/mol)	$D_{\max \text{ aver}}$ of complete locally stable conformers (Å)
ID29	$C_{22}H_{37}Cl_9$	27	648–659	32.2–32.3	$21.77 \pm 0.02$
ID30	$C_{24}H_{40}Cl_{10}$	30	657–666	37.5–37.8	$23.40 \pm 0.03$
ID31	$C_{26}H_{43}Cl_{11}$	33	660–674	39.9–40.1	$23.63 \pm 0.02$
ID32	$C_{28}H_{46}Cl_{12}$	36	661–672	36.2–36.3	$24.87 \pm 0.04$
ID33	$C_{30}H_{49}Cl_{13}$	39	698–717	50.3–50.6	$26.35 \pm 0.02$



**Fig. 3.** Comparison of  $D_{\max \text{ aver}}$  calculated by OASIS with those calculated in this study from the complete locally stable conformers (A) and from the filtered stable conformers with the numbers of 30 (B), 100 (C), and 200 (D).

the 69 chemicals. Seven different 3-D structures (A to G) of the chemical ID33 were created by CHEM3D, whose  $D_{\max}$  values were 32.1, 28.2, 25.0, 25.8, 30.3, 23.2, and 35.3 Å, respectively. Of these, structure F represented the smallest among the available initial structures, while G represented the largest. The MD simulations for each structure in water were performed for a specified time up to 2000 ps. Figure 4 shows the averages and the ranges of  $D_{\max}$  values after simulation times of 10, 100, 200, 500, and 2000 ps for each initial structure. The conformers for the  $D_{\max}$  calculation were sampled every 1 ps in principle, although the conformers for the 10 ps simulation were sampled every 0.1 ps to increase the number of the conformers. The smallest  $D_{\max}$  values ( $D_{\max \text{ min}}$ ) in the 10 ps simulation ranged from 23.2 (structure F) to 32.2 Å (structure G), and the difference was 9.0 Å. On the other hand, the largest  $D_{\max}$  values ( $D_{\max \text{ max}}$ )

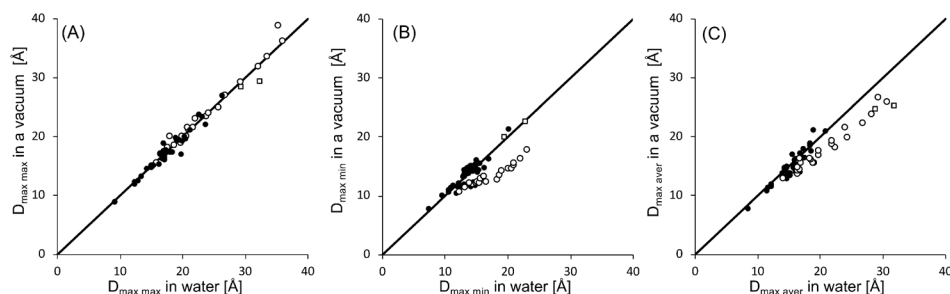


**Fig. 4.** Average (black square) and range (vertical bar) of  $D_{\max}$  values in water during calculation times of 10, 100, 200, 500, and 2000 ps when starting from seven different 3-D structures (A to G) of chemical ID33 ( $C_{30}H_{49}Cl_{13}$ ).

ranged from 25.2 (structure C) to 35.3 Å (structure G), and the difference was 10.1 Å. These results indicated that both the  $D_{\max \text{ min}}$  and  $D_{\max \text{ max}}$  values were strongly affected by the initial structures of a test chemical in the case of the 10 ps simulation, and such a short simulation time was insufficient to cover a wide range of possible structural changes of a chemical having a flexible structure like ID33. By extending the simulation time from 10 to 2000 ps, the differences of  $D_{\max \text{ min}}$ ,  $D_{\max \text{ max}}$ , or  $D_{\max \text{ aver}}$  values became smaller, ultimately being less than 0.1 Å between any pair of two initial structures after 500 ps or more. In fact, the standard deviations of  $D_{\max \text{ min}}$ ,  $D_{\max \text{ max}}$ , and  $D_{\max \text{ aver}}$  among these seven structures in the 10 ps simulation were 3.50, 4.03, and 3.84 Å, respectively; however, those in the 500 ps simulation decreased to 0.07, 0.09, and 0.03 Å, respectively. Eventually,  $D_{\max}$  values ranged from 23 to 35 Å depending on the simulation time; however, the maximum or minimum values of  $D_{\max}$  obtained after 500 ps simulations were unchanged by extending the simulation time as in the case of chemical ID33. When several other test chemicals were examined in the same manner, all simulations converged in a shorter simulation time than that of ID33. Therefore, the simulation time was set to 500 ps in the following investigations, taking into account the calculation costs.

### 3. Comparison of $D_{\max}$ values in a vacuum and in water

$D_{\max}$  values of the 69 chemicals in water were compared with those in a vacuum as calculated in this study (Fig. 5). Although the maximum  $D_{\max}$  values ( $D_{\max \text{ max}}$ ) in water were almost the same as those in a vacuum (Fig. 5A), the minimum  $D_{\max}$  values ( $D_{\max \text{ min}}$ ) in water were different from those in a vacuum in the case of polychlorinated alkanes with a high degree of freedom (open circles in Fig. 5B). For example, the differences of  $D_{\max \text{ min}}$



**Fig. 5.** Comparison of  $D_{\max \max}$  (A),  $D_{\max \min}$  (B), and  $D_{\max \text{aver}}$  values (C) in water based on MD simulations with those in a vacuum based on quantum chemical calculations for 69 chemicals (closed circles) including ID12 to ID33 (open circles) and ID D2 and D3 (open squares). A line with slope 1 is shown in each figure.

values in water and in a vacuum for ID12 to ID33 were 1.8 to 6.5 Å, where the values in water were larger than those in a vacuum. These chemicals can adopt highly compact conformations in a vacuum, but such compact conformations would hardly be formed in water due to their interaction with water molecules. As a result, the average values of  $D_{\max}$  ( $D_{\max \text{aver}}$ ) in a vacuum were not in good agreement with those in water because some of the stable conformers in a vacuum cannot exist in the aqueous environment (Fig. 5C). Similarly, two pigments, ID D2 and D3 (open squares in Fig. 5C; see their structures in Table 1), with higher molecular weights (*i.e.*, around 700) also showed different  $D_{\max \text{aver}}$  in a vacuum and in water. In general,  $D_{\max \text{aver}}$  values in water were larger than those in a vacuum for bulky chemicals. These results indicate that the MD calculations help correctly estimate the 3-D stereo structures of molecules in water, especially for some flexible or bulky molecules.

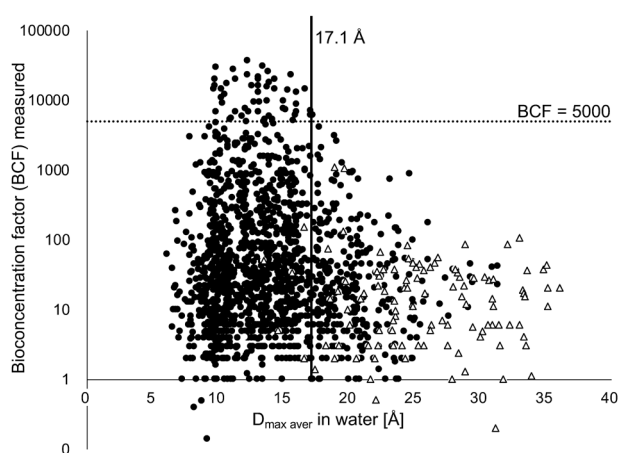
#### 4. Evaluation of bioaccumulation potential with the $D_{\max}$ value in water

Mekenyan *et al.* selected the conformers in a vacuum that can exist within 20 kcal/mol of the most stable conformer without any validations, and the 17.4 Å of  $D_{\max \text{aver}}$  based on these conformers has been proposed as an indicator of low bioaccumulation.<sup>16,17</sup> In this study, we calculated the molecular sizes in water for chemicals that had been subject to bioconcentration studies and tried to find a size threshold to predict the low bioaccumulation of chemicals.

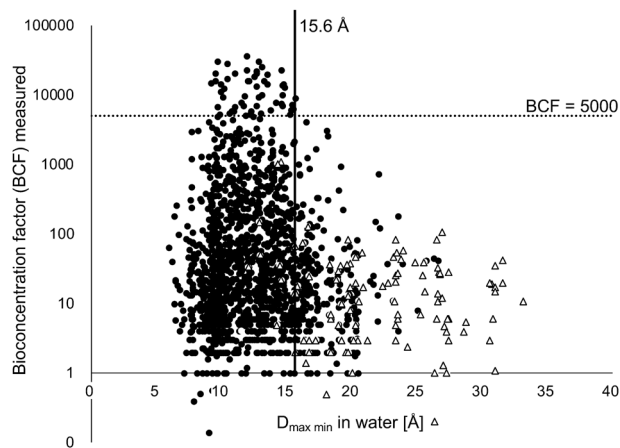
Figure 6 shows the relationship between the measured BCF values and the  $D_{\max \text{aver}}$  values in water for the 1665 chemicals, including 116 substances with MW >800 (open triangle), together with a dotted line at BCF=5000. The maximum  $D_{\max \text{aver}}$  value for the chemicals with BCF  $\geq$  5000 was 17.1 Å among the 1665 chemicals. The number of chemicals with  $D_{\max \text{aver}} > 17.1$  Å was 436, and all of them had low bioaccumulation (*i.e.*, BCF < 5000). Therefore, a criterion of the low bioaccumulation in terms of  $D_{\max \text{aver}}$  in water could be set as  $> 17.1$  Å, which was in good agreement with the  $D_{\max \text{aver}}$  in a vacuum of 17.4 Å approved under REACH.<sup>17</sup> This result indirectly justifies the approximation by Mekenyan *et al.* to predict the  $D_{\max}$  values in water, where only the conformers in a vacuum within 20 kcal/mol of the most stable one were selected for the calculation.

Under the CSCL, chemicals with MW >800 are approved as low bioaccumulative and are not required to be tested in fish.<sup>2,13</sup> When applying this standard, only 116 of the examined 1665 chemicals (open triangles in Fig. 6) are exempted from bioconcentration testing using fish. In this study, the  $D_{\max}$  values derived from realistic 3-D conformers in water were considered more appropriate than MW as the indicator of molecular bulkiness that may govern the penetration of biological membranes. Based on the  $D_{\max}$  value in water calculated in this study, it was expected that 436 chemicals with  $D_{\max \text{aver}} > 17.1$  Å could be excluded from testing.

While the  $D_{\max \text{aver}}$  values have conventionally been deemed the criterion of low bioaccumulation, the  $D_{\max \min}$  values more likely reflect on the stereo structures of compact conformers that can be more closely associated with gill membrane permeation. Thus, from the viewpoint of screening assessment, the use of  $D_{\max \min}$  values as thresholds seems more conservative and easily accepted by regulatory authorities. Figure 7 shows the relationship between the measured BCF values and the  $D_{\max \min}$  values in water for the same set of chemicals as in Fig. 6. Here, the maximum  $D_{\max \min}$  value for the chemicals with BCF  $\geq$  5000 was 15.6 Å. The number of chemicals regarded as having low bioac-



**Fig. 6.** Relationship between  $D_{\max \text{aver}}$  values in water and measured BCF values for 1665 chemicals (closed circles) including substances >800 MW (open triangles). The dotted line denotes the threshold of BCF=5000.



**Fig. 7.** Relationship between  $D_{\max \min}$  values in water and measured BCF values for 1665 chemicals (closed circles) including substances  $>800$  MW (open triangles). The dotted line denotes the threshold of  $BCF=5000$ .

accumulation by the  $D_{\max \min}$  threshold was 357, which was somewhat less than that by the  $D_{\max \text{ aver}}$  threshold used in this study. This result also confirmed that the  $D_{\max \min}$  threshold was more conservative than that of the  $D_{\max \text{ aver}}$ .

This study deduced two thresholds,  $D_{\max \text{ aver}} >17.1 \text{ \AA}$  and  $D_{\max \min} >15.6 \text{ \AA}$ , of low bioaccumulation evaluation. These values are determined based on the realistic molecular size in water, so they are considered more reliable thresholds than the conventional values that assume the molecular size in a vacuum. In addition to reflecting more realistic structures, the 3-D stereo structures of some chemicals were larger in water than in a vacuum (Fig. 5). This means that even the chemicals with  $D_{\max}$  values below the threshold in a vacuum may have values above the threshold in water in some cases; thus, the use of the  $D_{\max}$  value in water can be more advantageous for reducing the number of fish experiments. Between the two threshold values based on  $D_{\max \text{ aver}}$  and  $D_{\max \min}$ , it seems appropriate to use the former, since it reflects all conformers that can exist in water. However, from a regulatory point of view, setting a margin on the threshold is preferable due to safety concerns. Therefore, we conclude that the more conservative threshold based on  $D_{\max \min}$  is more suitable. In any case, setting these new thresholds in water would contribute to finding a more accurate screening method for determining low bioaccumulative chemicals with  $BCF <5000$  and reducing animal testing.

### Conclusion

To evaluate the bioaccumulation potential in an aquatic environment,  $D_{\max}$  values of chemicals in water were calculated by an MD simulation. Obtained  $D_{\max \text{ aver}}$  values in water were generally comparable to those in a vacuum, but some of the most compact conformers in a vacuum (*i.e.*,  $D_{\max \min}$ ) were not reproduced in water for 69 chemicals in the UK report. Thus, calculating 3-D conformations in water seemed to be helpful for a more accurately screening bioaccumulation potential in fish.

Analysis using the measured BCF values of 1665 chemicals revealed that chemicals in water with  $D_{\max \text{ aver}} >17.1 \text{ \AA}$  or  $D_{\max \min} >15.6 \text{ \AA}$  had BCF values of less than 5000. This threshold would contribute to reducing animal testing without diminishing the conventional screening accuracy based on molecular weight.

### References

- 1) D. Wijk, R. Chénier, T. Henry, M. D. Hernando and C. Schulte: Integrated Approach to PBT and POP Prioritization and Risk Assessment. *Integr. Environ. Assess. Manag.* **5**, 697–711 (2009).
- 2) S. Jacobi, B. Brown, M. Claessens, I. Colombo, M. Embry, J. Franklin, B. Hidding, M. L. Paumen, N. McGrath, C. Miyata, K. Woodburn and H. Vrijhof: Information to be considered in a weight-of-evidence-based PBT/vPvB assessment of chemicals (Annex XIII of REACH). *ECETOC Special Report No. 18* (2014).
- 3) Japan METI: “Act on the Evaluation of Chemical Substances and Regulation of Their Manufacture, etc. (Chemical Substances Control Law).” Ministry of Economy, Trade and Industry, Tokyo, Japan. <http://www.japaneselawtranslation.go.jp/law/detail/?id=3350&vm=02&re=01> (Accessed 10 Sep., 2020).
- 4) Regulation (EC) No 1907/2006-Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) <https://osha.europa.eu/en/legislation/directives/regulation-ec-no-1907-2006-of-the-european-parliament-and-of-the-council> (Accessed 10 Sep., 2020).
- 5) ECHA: REGULATION (EC) No 1907/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 (2006).
- 6) OECD guidance for testing chemicals: Bioaccumulation in Fish: Aqueous and Dietary Exposure (2012).
- 7) Guideline of Chemical Substances Control Law: Testing methods for new chemical substances, Bioconcentration test using fish (Water exposure method) (2018) (in Japanese).
- 8) J. W. Nichols, M. Bonnell, S. D. Dimitrov, B. I. Escher, X. Han and N. I. Kramer: Bioaccumulation Assessment Using Predictive Approaches. *Integr. Environ. Assess. Manag.* **5**, 577–597 (2009).
- 9) S. Nakai, K. Takano and S. Saito: Bioconcentration Prediction under the Amended Chemical Substances Control Law of Japan. *R&D Report SUMITOMO KAGAKU 2006-I* (2006).
- 10) Y. Sakuratani, Y. Noguchi, K. Kobayashi, J. Yamada and T. Nishihama: Molecular size as a limiting characteristic for bioconcentration in fish. *J. Environ. Biol.* **29**, 89–92 (2008).
- 11) A. Opperhuizen, E. W. Volde, F. A. P. C. Gobas, D. A. K. Liem and J. M. D. Steen: Relationship between bioconcentration in fish and steric factor of hydrophobic chemicals. *Chemosphere* **14**, 1871–1896 (1985).
- 12) F. Grisoni, V. Consonni, S. Villa, M. Vighi and R. Todeschini: Molecular size cutoff criteria for screening bioaccumulation potential: Fact or fiction? *Chemosphere* **127**, 171–179 (2015).
- 13) A. V. Weisbrod, L. P. Burkhard, J. Arnot, O. Mekenyan, P. H. Howard, C. Russom, R. Boethling, Y. Sakuratani, T. Traas, T. Bridges, C. Lutz, M. Bonnell, K. Woodburn and T. Parkerton: Workgroup report: Review of fish bioaccumulation databases used to identify persistent, bioaccumulative, toxic substances. *Environ. Health Perspect.* **115**, 255–261 (2007).
- 14) S. D. Dimitrov, N. C. Dimitrova, J. D. Walker, G. D. Veith and O. G. Mekenyan: Bioconcentration potential predictions based on molecular attributes—An early warning approach for chemicals found in humans, birds, fish and wildlife. *QSAR Comb. Sci.* **22**, 58–68 (2003).
- 15) S. D. Dimitrov, N. C. Dimitrova, J. D. Walker, G. D. Veith and O. G. Mekenyan: Predicting bioconcentration factors of highly hydropho-



- bic chemicals. Effects of molecular size. *Pure Appl. Chem.* **74**, 1823–1830 (2002).
- 16) ECHA: Guidance on information requirements and chemical safety assessment. Part C: PBT/vPvB assessment. Version 2.0. European Chemicals Agency, Helsinki, Finland. (2014).
  - 17) ECHA: Guidance on information requirements and chemical safety assessment. Chapter R.11: PBT assessment. Version 3.0. European Chemicals Agency, Helsinki, Finland. (2017).
  - 18) D. N. Brooke and M. J. Crookes: Calculation of molecular dimensions related to indicators for low bioaccumulation potential. Science report SCHO0109BPGT-E-P. Environment Agency, Bristol, UK. (2009).
  - 19) O. Mekenyan, D. Dimitrov, N. Nikolova and S. Karabunarliev: Conformational Coverage by a Genetic Algorithm. *J. Chem. Inf. Comput. Sci.* **39**, 997–1016 (1999).
  - 20) G. Chang, W. C. Guida and W. C. Still: An internal-coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* **111**, 4379–4386 (1989).
  - 21) A. Hocquet and M. Langgard: An Evaluation of the MM+ Force Field. *J. Mol. Model.* **4**, 94–112 (1998).
  - 22) L. Grippo and S. Lucidi: A globally convergent version of the Polak-Ribiere conjugate gradient method. *Math. Program.* **78**, 375–391 (1997).
  - 23) H. Sun: COMPASS: An *ab Initio* force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. *J. Phys. Chem. B* **102**, 7338–7364 (1998).
  - 24) S. Nosé: A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
  - 25) W. G. Hoover: Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
  - 26) H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak: Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
  - 27) A. Bondi: van der Waals Volumes and Radii. *J. Phys. Chem.* **68**, 441–451 (1964).
  - 28) K. Ishihata: *JPSJ Magazine*. **43**, 1009–1015 (2002).
  - 29) National Institute of Technology and Evaluation: The result of bio-concentration. (judged from 1974 to 1986). <https://www.nite.go.jp/data/000022750.xlsx> (Accessed 31 May, 2021).
  - 30) National Institute of Technology and Evaluation: The result of bio-concentration (judged from 1987 to 2017). <https://www.nite.go.jp/data/000108118.xlsx> (Accessed 31 May, 2021).
  - 31) OECD Principles of Good Laboratory Practice (1997).
  - 32) National Institute of Technology and Evaluation: Prediction of bio-concentration by category approach. [https://www.nite.go.jp/chem/qsar/category\\_approach.html](https://www.nite.go.jp/chem/qsar/category_approach.html) (in Japanese). (Accessed 31 May, 2021).
  - 33) Y. Ikenaga, Y. Sakuratani and J. Yamada: Evaluation of Bioconcentration Factors by Category Approach for Chemicals. National Institute of Technology and Evaluation. <https://www.nite.go.jp/data/000064116.pdf> (Accessed 31 May, 2021).
  - 34) J. W. Martin, S. A. Mabury, K. R. Solomon and D. C. Muir: Progress toward understanding the bioaccumulation of perfluorinated alkyl acids. *Environ. Toxicol. Chem.* **32**, 2421–2423 (2013).
  - 35) C. A. Ng and K. Hungerbühler: Bioconcentration of perfluorinated alkyl acids: How important is specific binding? *Environ. Sci. Technol.* **47**, 7214–7223 (2013).
  - 36) C. A. Ng and K. Hungerbühler: Bioaccumulation of perfluorinated alkyl acids: Observations and models. *Environ. Sci. Technol.* **48**, 4637–4648 (2014).