5-2021

Solvation Properties of Persistent Organic Pollutants and the Relationship to the Bioconcentration Factor

Emily L. Wittkop

Solvation Properties of Persistent Organic Pollutants and the Relationship to the Bioconcentration Factor

Joseph D. Alia, Emily L. Wittkop, Division of Science and Mathematics, University of Minnesota, Morris

Abstract. Persistent organic pollutants (POPs) are a class of substances that can cause major health issues in humans and other animals alike. They are difficult to track through the environment and even more difficult to remove. One of the ways in which we can learn more about these compounds is through the bioconcentration factor (BCF). Little work has been done to differentiate the effects of POP stereoisomers and degradation products on human populations, flora, and fauna. For this work, the selected POPs included Aldrin and related compounds, Chlordane and related compounds, Endosulfan and related compounds, and Hexabromocyclododecane. To learn more about the BCF of the selected compounds, the SMD solvent model was used to find the octanol-water partition coefficient (K_{ow}) and the olive oil-water partition coefficient (K_{vow}). The basis set used was 6-31G (d) and the B3LYP density functional. A linear regression was performed to investigate the potential relationship between the $K_{v/ow}$ and BCF. Other molecules with known BCF values were also used to strengthen the regression. This regression showed a potential correlation between the $K_{v/ow}$ and the BCF that would be strengthened by more information about the BCF of POPs.

Major Findings:
This work had several valuable findings, though the final goal of an equation relating Bioconcentration Factor and Octanol (or Olive Oil) – Water Partition Coefficient was not accomplished. One of the primary findings of significance is that there is a difference between the $K_{v/ow}$s of stereoisomers for the selected POPs. Some of the literature available doesn't make the distinction between stereoisomers, despite the $K_{v/ow}$ sometimes being significantly different. Another finding of this work is that the linear relationship originally proposed by Mackay (1982)[1] is confirmed in all the plots in Figure 1. The statistical accuracy, modelled by the R^2 value, varies across methods, along with the slope, but all methods showed a linear relationship. More about this can be found in the Results and Discussion section.

Introduction:
POPs are compounds that pose significant health risks to humans and other animals. Many of these compounds are used as pesticides. They find their way into the environment and spread through the air, water, and food chain. One of the methods used to identify their movement is the bioconcentration factor (BCF). The BCF is difficult to find experimentally but can also be related to the Octanol-Water Partition Coefficient (K_{ow}), according to several studies, including one from Mackay (1982) that is of particular importance to this work. The K_{ow} can be computed several ways, as is discussed in the Methods section.

The POPs selected for this work were chosen because of their theorized compatibility with the computational methods. A few of the selected POPs can be seen in Figure 2. Those not included in the figure include stereoisomers and degradation products of those shown. Stereochemistry plays an important role when calculating physical properties and is not always accounted for in the literature surrounding these compounds. The degradation products are understudied, likely because they exist only in very small quantities in our environment, but they can be equally as, if not more, harmful as their parent compound.

In this work we hope to identify a linear relationship that can be used to find the BCF from the K_{ow} for the methods used and go on to use this relationship in order to learn more about the understudied compounds, such as the degradation products. Many of the methods used are based on the work of Alia, et al. (2020)[2].

Methods:
Three methods were employed for this work to calculate the K_{ow}:

ALOGPS 2.1: ALOGPS 2.1[3] is an online software that is provided through the Virtual Computational Chemistry Laboratory (VCCLAB). It uses the SMILES code for a compound to calculate solubility and K_{ow}. It is instant, making it great for finding estimates for the K_{ow}, but id doesn't account for stereochemistry.

Equations from Mackay (1982): In Mackay's paper, he uses several equations (equations 1 and 2) that consider both solubility and melting point of a compound to estimate the K_{ow}. We used these equations in this work for another estimate of the K_{ow}. The biggest concern with these equations is finding accurate solubility and melting point data. As we specifically studied compounds that have less data available, it follows that it would be difficult to find accurate values for solubility and melting point.

$$(1)\ \ln K_{ow} = 7.494 - \ln C$$

$$(2)\ \ln K_{ow} = 7.494 - \ln C + \left(1 - \frac{T_m}{T}\right)\ \ T = 298.15\ K$$

DFT: Lastly, we used Gaussian[4] to find both the K_{ow} and the K_{vow}. We used DFT B3LYP/6-31G(d) for the optimizations and the same method alongside SMD for the solvent energy calculations. This gives energy values that can be used in equation 3 to give the K_{ow}. In order to find the K_{vow}, the solvation energy in n-octanol is replaced with the solvation energy in olive oil. The olive oil calculations were done to find a more accurate model for the behavior of these molecules in a biological system.

$$(3)\ logK_{ow} = -\frac{G^{\circ}_{n-octanol} - G^{\circ}_{water}}{RTln(10)}$$

BCF Values: The K_{ow}s were compared to known BCF values that were found from, primarily, two different databases: the Arnot Gobas Database[5], and the NITE database[6]. We used to the BCF from Japanese Carp through out.

Additional Data: More compounds were added to increase the strength of the linear regression. These compounds were selected from Mackay's paper and include some simple compounds, such as benzene, and some other POPs, such as DDT.

Results and Discussion:
Figure 1 shows the plots comparing K_{ow} and BCF for the selected compounds. Outliers were removed for each plot by comparing residuals. The R^2 value for each linear regression is given on the plot. I want to note that, despite the high R^2 value of plot a, which uses the log K_{ow} from ALOGPS 2.1, it is the least accurate because it doesn't account for stereochemistry. The R^2 values for all the other plots are too low to accurately claim that the linear regression is an accurate relationship that can always be used to calculate the BCF of a compound.

Despite the low values of R^2, these regressions and the work still have valuable outcomes. These regressions show computational methods are accurate enough to give similar regressions to other methods of finding the K_{ow}. The data from these methods (except ALOGPS 2.1) also enforced that stereoisomers do not have the same value of K_{ow} and therefore different solvation properties.

It is worth noting that SMD models can have issues with highly halogenated compounds, like the POPs that we selected. This could have added error to our data.

This work could be continued in a few ways. First, confirming and improving the Gaussian input files to have more accurate structures could improve the linear regressions. Second, these regressions and the results could be further analyzed to see if different structural groups have their own regressions separate from one another. Third, other solvent models could be tested with the highly chlorinated compounds to see if those other modelling methods produce more accurate results and therefore a more accurate, and hopefully better, regression.
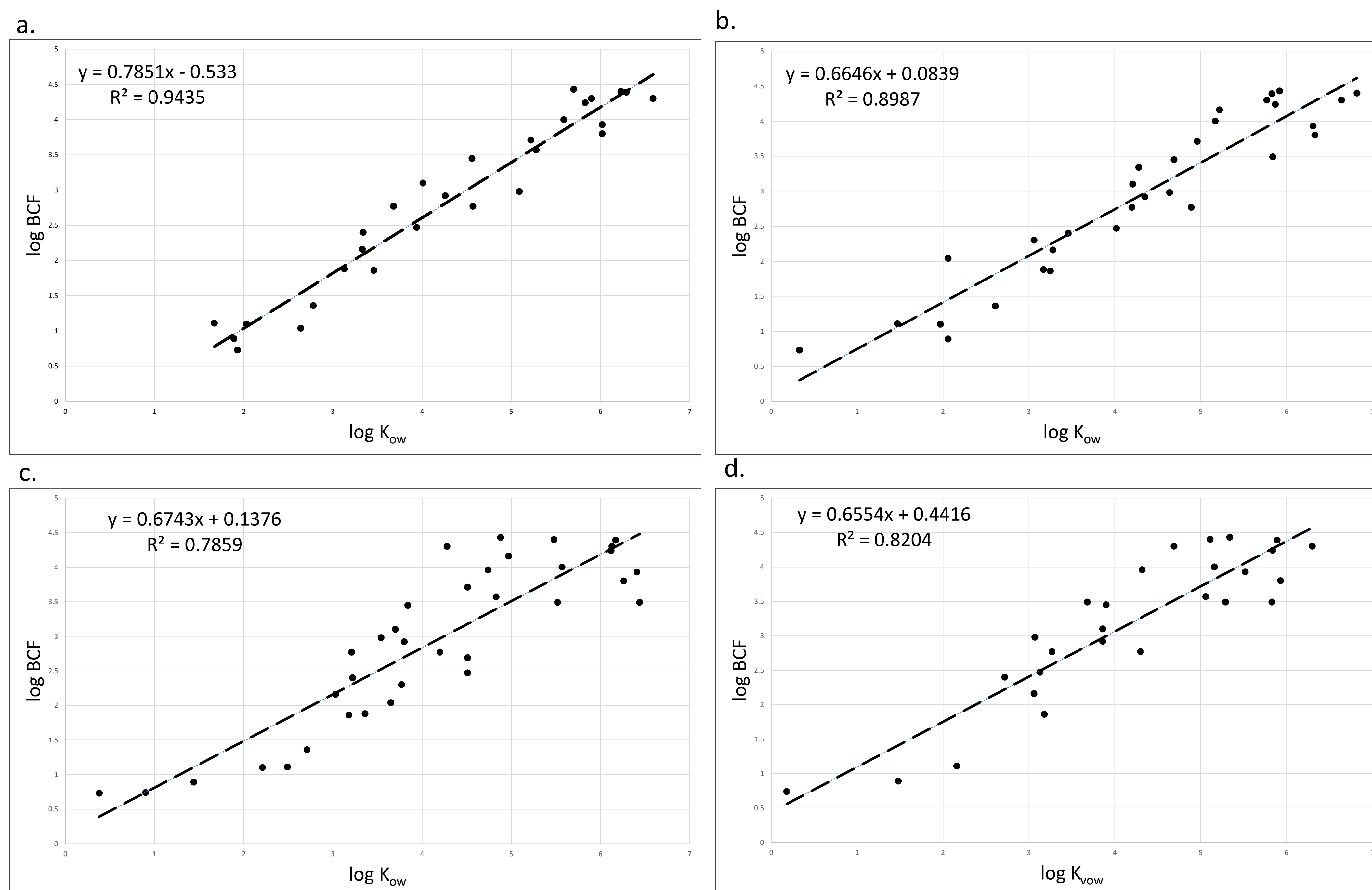

Figure 1. a) log K_{ow} as calculated by ALOGPS 2.1 compared to the log BCF. b) log K_{ow} as calculated by the equations provided by Mackay (1982) compared to the log BCF. c) log K_{ow} as calculated by DFT (B3LYP/6-31G (d), SMD) compared to the log BCF. d) log K_{vow} as calculated by DFT (B3LYP/6-31G(d), SMD) compared to the log BCF.
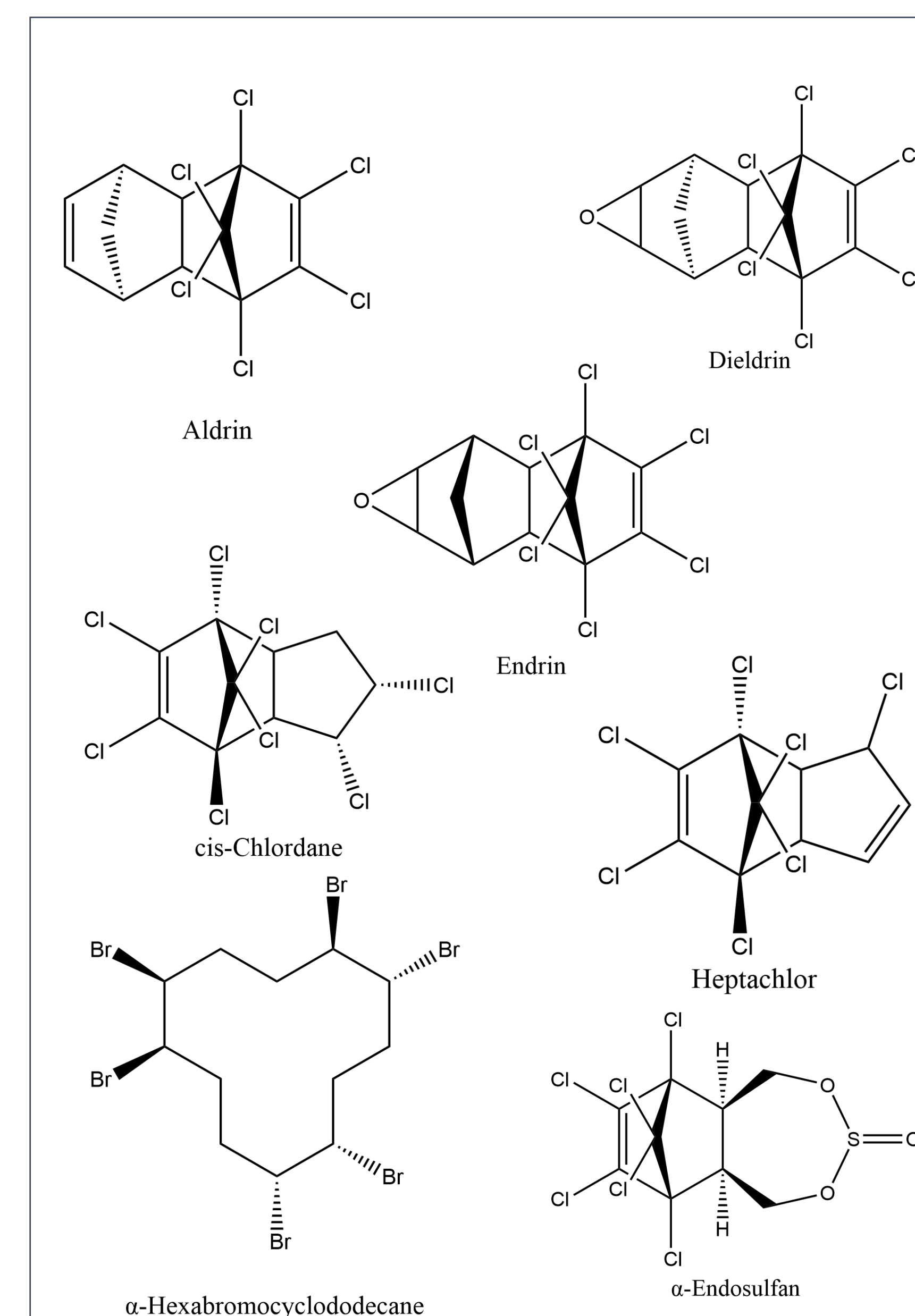

Figure 2. A few of the selected POPs used in this work. The other selected POPs include stereoisomers of the ones shown and degradation products of those shown.

References:
1. Mackay, D. Correlation of Bioconcentration Factors. *Environ. Sci. Technol.* **1982**, 16, 274-278.
2. Alia, J. D.; Karl, S., Kelly, T. D. Quantum Chemical Lipophilicities of Antimalarial Drugs in Relation to Terminal Half-Life. *ACS Omega*, **2020**, 5, 12, 6500-6515. https://doi.org/10.1021/acsomega.9b04140
3. Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; Tanchuk, V. Y.; Prokopenko, V. V. Virtual computational chemistry laboratory - design and description, *J. Comput. Aid. Mol. Des.*, **2005**, 19, 453-63, article.
4. Gaussian 16, Revision C.01, Gaussian, Inc., Wallingford CT, 2016.
5. Arnot, J. A., Gobas, F. A. P. C. A Review of Bioconcentration Factor (BCF) and Bioaccumulation Factor (BAF) Assessments for Organic Chemicals in Aquatic Organisms. *Environ. Rev.* **2006**, 14, 257-297.
6. NITE-CHRIP. National Institute of Technology and Evaluation. https://www.nite.go.jp/en/chem/chrip/chrip_search/html/condUse.html