Predicting Bioconcentration Factor Using a Metabolism-Based Quantitative Structure-Activity Relationship Model

Julia Magliozzo, Nan An, and Adelina Voultskova-Kostal
Chemistry Department, The George Washington University

Introduction

Bioconcentration refers to the absorption and retention of a substance through an animal’s dermal and/or respiratory surfaces, excluding contributions from dietary uptake. Generally applied to aquatic organisms, it is quantified using a bioconcentration factor, defined as the ratio of substance concentration in the organism versus the water at steady state. In the case of fish, the bioconcentration factor can be considered the result of two competing processes: uptake of the substance through the gills and chemical elimination through respiratory exchange, fecal excretion, metabolic biotransformation of the parent compound, and growth dilution (Arnott and Gobas 2006).

Aside from the direct implications of xenobiotics in fish, the bioconcentration factor is also used as a midpoint in the determination of bioaccumulation factors needed to define water quality criteria for human health by EPA (EPA 2016) and as a criterion to designate substances as persistent, bioaccumulative and toxic (PBT) under the European Union’s REACH regulations (Williams, Panko et al. 2009).

Metabolism-based Approach

To overcome the limitations of existing models Grisoni et al. (2015) suggested accounting for metabolism or processes other than lipid-driven accumulation patterns to develop models with high predictive accuracy and wide applicability domains.

The present work addresses this suggestion by using a subset of data (476 compounds total) from previous work by Dimitrov et al. (2005) to develop a two-step linear QSR model for bioconcentration factor. The two steps are:

1. Tiered classification based on potential metabolism of the chemical. In this step, chemicals are grouped based on structural characteristics as described by 5Miles Arbitrary Target Specification (SMARTS) patterns.
2. Application of the appropriate QSR given the metabolic classification. The disadvantage of a linear model, the non-linearity of the logBCF-octanol-water partition coefficient relationship at very low and very high logP, is overcome by limiting the applicability domain to only those compounds with intermediate logP where linearity is observed (Wen, He et al. 2012). This approach is chosen to explicitly address mechanistic interpretations a priori to model development and shows valuable results.

Classification 1: Gill Permeation

Uptake through the gills is the primary mode by which chemicals are absorbed. At this interface the most important reactions should be those that alter electrophilic compounds as they pass through the gills before any metabolism occurs.

Scheme 1. A generic S_{2} reaction.

A common reaction that may occur is S_{2} substitution. In this reaction a good leaving group such as O, Br, or water is needed and is replaced by a nucophile such as a hydroxide ion or an amine. This can alter the bioconcentration of the compound because the new group can change the interactions of the compound with the surrounding tissues.

Classification 2: Metabolism

The second mode of reaction of the chemicals is in metabolic pathways within the tissues of the fish. These reactions are catalyzed by enzymes, the most relevant of which we believe to be cytochrome P450. Cytochrome P450 is a very active and diverse enzyme that largely falls into the category of EC1 enzymes in the KEGG database. These enzymes are responsible for a breadth of reactions ranging from epoxidation to O,N-dealkylation to oxidation.

Scheme 2. Examples of known P450 catalyzed reactions. (Schlörk et al., 2009)

Compounds from our dataset will be classified as potential substrates for metabolism using SMARTS from previous work by Mu et al. Only those SMARTS pertinent to reactions catalyzed by EC1 enzymes will be used.

QSAR Development and Next Steps

The dataset was limited to compounds whose logD values have a linear relationship with logBCF. LogBCF > 0.5 to 6 for all chemical classes and LogD > 6 to 8 for halogenated biphenyls (Wen et al. 2012) resulting in a final set of 403 compounds. After classifying the entire linear range dataset by gill permeation and metabolism we will establish whether there are meaningful differences between each class relative to the rest of the dataset and to each other. For example, preliminary data suggests that the gill permeation acyl substitution class shows a more constant logBCF given different values of LogD.

Conclusions

We present here a systematic approach for including metabolism or mechanistic interpretation a priori to model development in a QSR with a logBCF endpoint. We have not yet determined whether our two-step process for predicting logBCF will provide more reliable results than existing QSR models. However, regardless of the resulting QSRs, we believe the novelty of our two-tier classification scheme will fill a significant gap in the QSR approach. Given the importance of considering metabolic processes in modeling bioconcentration we hope to provide a classification scheme that will allow for more knowledgeable and realistic bioconcentration QSRs to be developed in the future.

Acknowledgements

We would like to thank Dr. Bryan W. Brooks and Dr. Jona Corrales from CRASH at Baylor University for their insights in fish metabolism. We would also like to acknowledge Dr. Jakub Kostal for his assistance in the computational and SMARTS analysis.

Literate Cited