Supplementary Information

In Silico Modeling Method for Computational Aquatic Toxicology of Endocrine Disruptors: A Software-Based Approach Using QSAR Toolbox

Marie-Léonie Bohlen, Hyun Pyo Jeon, Young Jun Kim, and Baeckkyoung Sung*

KIST Europe Forschungsgesellschaft mbH, 66123 Saarbrücken, Germany

*Correspondence: sung@kist-europe.de

Fig. S1 and the following Eqs. S1 and S2 show the statistical relationship between the descriptor (K_{OW}) and endpoint (LC_{50}) of the EDs listed in Table 1. In the log-log plot, the linear regression for the predicted and experimental values showed a comparable slope ($\log_{10}LC_{50}$ / $\log_{10}K_{OW}$ = -0.509 and -0.414, respectively) with a relatively high coefficient of determination for the prediction (adjusted r^2 = 0.805 and 0.437 for prediction and observation, respectively).

Computational prediction: $$log_{10}LC_{50} = -0.509 \cdot log_{10}K_{ow} + 2.43 \quad \text{(Eq. S1)}$$

Experimental observation: $$log_{10}LC_{50} = -0.414 \cdot log_{10}K_{ow} + 2.12 \quad \text{(Eq. S2)}$$

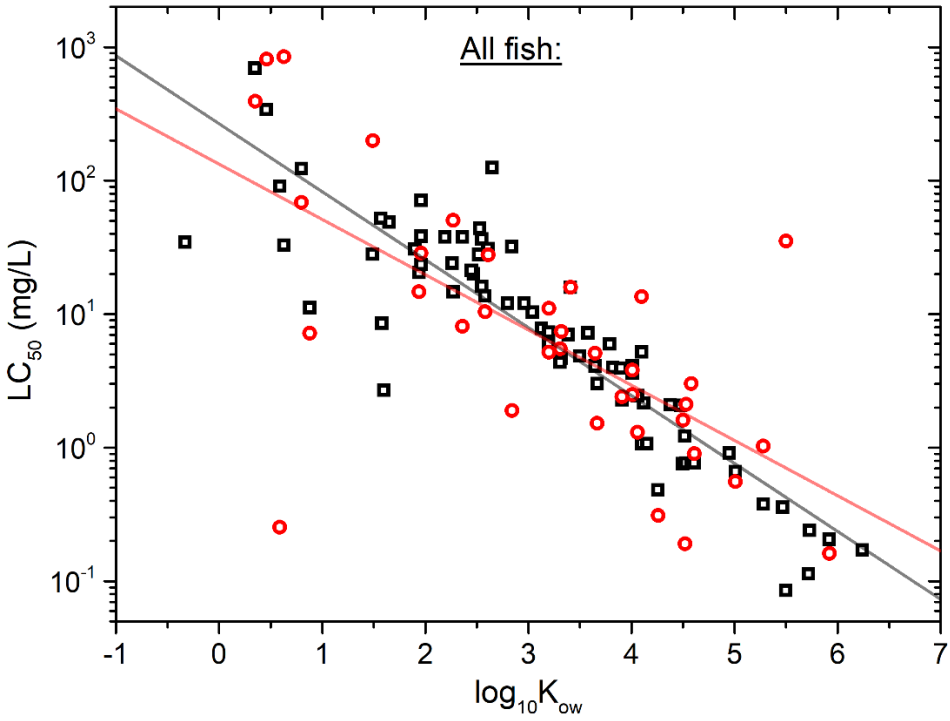

Figure S1: Relationships between 96-h LC_{50} and $\log_{10}K_{OW}$ values of EDs listed in Table 1. For all fish species, the values for predicted (black empty squares, n = 74) and experimental (red empty circles, n = 36) are plotted together in a log scale as a function of $\log_{10}K_{OW}$. For the predicted LC_{50}, the average values are displayed. The gray (Eq. 1) and light-red (Eq. 2) lines represent the linear regression for each case (predicted and experimental, respectively).

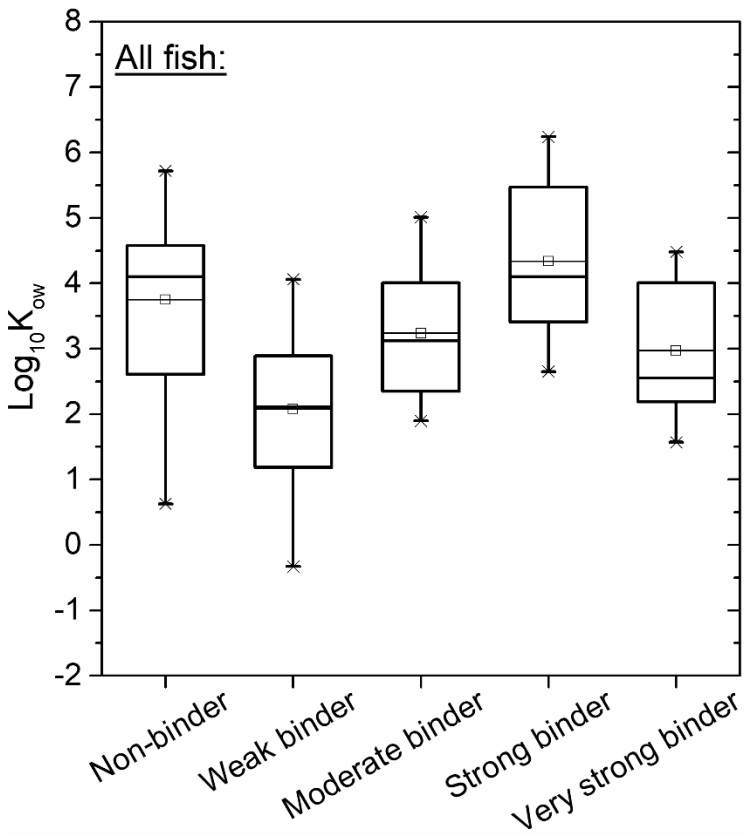

Figure S2: Distributions of the $\log_{10}K_{OW}$ depending on the estrogen receptor (ER) binding affinity of EDs in Table 1 for all fish (n = 8-20 for each category). A box plot represents: (a) mean (small square with a horizontal bar), (b) 1st and 3rd quartiles (lower and upper ends of the box, respectively), (c) median (horizontal segment inside the box), (d) 5th and 95th percentile (lower and upper error bars, respectively), (e) 1th and 99th percentile (lower and upper ×, respectively), and (f) minimum and maximum (lower and upper −, respectively).

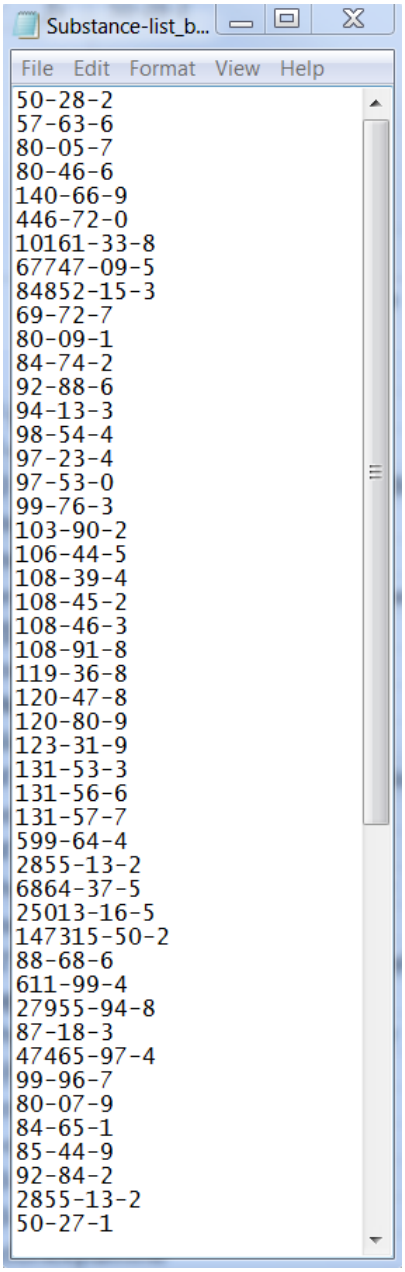

Figure S3: Example of a substance list of CAS numbers for batch mode processing. Each CAS number is listed in a single row. The file was prepared in a text editor and must be saved as text file (.txt).