**Table S2: compilation of QSPRs for primary biodegradation in aquatic matrices**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Endpoint** | **Equation and or model specifications** | **r2** | **Description of training set** | **Reference** |
| Kb  First-order biomass-normalized biodegradation rate coefficients,  (mg of protein/L)-1 (h)-1 | kb = -0.135 + 3.393 x 10^-4 9 SIGMA(ARaman (CH3 str.)), DFT ( B3LYP/6-31G\*) and linear regression | 0.9216 | 6 Dimethylnaphthalenes, Measurements performed in aqueous systems | Alparone (2012) |
| Kb  First-order biomass-normalized biodegradation rate coefficients,  (mg of protein/L)-1 (h)-1 | kb = -0.560 + 2.952 x 10^-4 x SIGMA(ARaman (3N-6)), DFT ( B3LYP/6-31+G\*\*) and linear regression | 0.9409 | 6 Dimethylnaphthalenes, Measurements performed in aqueous systems | Alparone (2012) |
| Kb  First-order biomass-normalized biodegradation rate coefficients,  (mg of protein/L)-1 (h)-1 | kb = -0.154 + 3.103 x 10^-4 x SIGMA(ARaman (CH3 str.)), DFT ( B3LYP/6-31+G\*\*) and linear regression | 0.9216 | 6 Dimethylnaphthalenes, Measurements performed in aqueous systems | Alparone (2012) |
| Kb  Biotransformation second order rate constant | log Kb = -13.7430rW + 0.0351VW + 0.1946pKa - 13.427, PLS | 0.986 | 8 Phenols | Damborsky and Schultz(1997) |
| Kb  Biotransformation second order rate constant | log Kb = -11.237rW + 0.0092Mw + 0.3737pKa - 14.194, PLS | 0.953 | 7 Anilines | Damborsky and Schultz(1997) |
| Kb  First-order biomass-normalized biodegradation rate coefficients,  mg of protein/L)-1(h)-1 | kb = (0.03231 +-0.00181) <a> - (3.92254 +-0.22528), DFT (B3LYP/6-31G\* in gas phase) and linear regression | 0.985 | 7 Dimethylnaphthalenes, Measurements performed in aqueous systems | Librando (2007) |
| Kb  First-order biomass-normalized biodegradation rate coefficients,  mg of protein/L)-1(h)-1 | kb = (0.03345+-0.00127) <a> - (4.67214+-0.18160), DFT (B3LYP/6-31+G\* in gas phase) and linear regression | 0.993 | 7 Dimethylnaphthalenes, Measurements performed in aqueous systems | Librando (2007) |
| Kb  First-order biomass-normalized biodegradation rate coefficient,  mg of protein/L)-1(h)-1 | kb = (0.04673+-0.00211) <a> - (9.17991+-0.41896), DFT (B3LYP/6-31+G\* in aqueous solution) and linear regression | 0.99 | 7 Dimethylnaphthalenes, Measurements performed in aqueous systems | Librando (2007) |
| Kb  rate constant | Kb = -0.977 x 10^-3 Mw + 0.533, Linear regression | 0.954 | 12 Phthalic acids and phthalate esters | Boethling (1986), Lu (2002) |
| Kb  First-Order Biomass-Normalized Rate Coefficient,  L organisms-1 h-1 | No significant correlations were found. No QSPR relationship could be established. Various molecular descriptors, univariate and two- and three-variable regressions for all descriptors generated |  | 22 PAHs, Measurements performed in aqueous nutrient buffer solution | Wammer (2005) |
| Kb  The second-order disappearance rate constant | log kb = 0.53(+-0.03) log kOH + 0.13(+-0.06) log KOW, +-11.8(+-0.2), Linear regression | 0.988 | - | Wolfe (1980) |
| Kb  The second-order disappearance rate constant | log Kb = 0.50(+-0.04) Khydr - 11.4(+-0.1), Linear regression | 0.973 | 6 Phtalate esters at 27 C and 25 C | Wolfe (1980), Steen (1979) |
| Kb  The second-order disappearance rate constant | log Kb = 2.1(+-0.4)Khydr -6(+-1), Linear regression | 0.933 | 4 Phtalate esters at 30 C and 25 C | Wolfe (1980), Steen (1979) |
| Kb  First-Order Biomass-Normalized Rate Coefficient,  L organisms-1 h-1 | log kb = - 14.1 Yvdw - 11.0, Linear regression | 0.924 | 7 Anilines and substituted anilines | Paris and Wolfe (1987), Mamy (2014) |
| K, d-1 | ln k =SIGMA(Njj), Linear regression | - | 11 Miscellaneous organic compounds | Desai et al. (1990), Mamy (2015) |
| K, d-1 | ln k = SIGMA(Njaj), Linear regression | - | 18 Miscellaneous organic compounds, Measurements performed in wastewater | Tabak and Govind (1993), Mamy (2015) |
| Kb  First-Order Biomass-Normalized Rate Coefficients (kb), L organisms-1 h-1 | log kb = - 0.9071 Yvdw - 8.313, Linear regression | 0.908 | 7 Phenols and substituted phenols (7 compounds), Measurements performed in water | Paris et al. (1983), Mamy (2015) |
| DT50  Biodegradation half-life | −logt1/2 = 12.049−0..013 Freq−0.159 N−5.329ELUMO  +16.528NEHOMO  −0.003IRIn, Density functional theory (DFT) and stepwise multiple linear regression analysis (SMLR) method | 0.905 | 17 Polycyclic aromatic hydrocarbons (PAHs), Measurements performed aged contaminated sediments | Xu (2012) |
| DT50  Half-life | DT50 = SIGMA(a0 + a1f1 + a2f2 · · · + amfm + e Open source (BioHCwin, EPI Suite), Multiple linear regression | 0.91 | 121/54 Petroleum hydrocarbons, Measurements performed in different environmental media (mostly aqueous) | Howard (2005) |