Table S3: compilation of QSPRs for chlorination HOCl

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Endpoint** | **Equation and/or model specifications** | **R2** | **S** | **Description of training set** | **Reference** |
| log(kHOCl) Second-order rate constants | log (kPhO-) = 4.46 (+-0.15) - 4.90 (+-0.44) SIGMA(s- o,m,p), Linear regression | 0.94 | 0.45 | 35 Dissociated phenols | Lee and von Gunten (2012) |
| log(kHOCl) | log (k3amine) = 4.79 (+-0.29) - 1.96 (+-0.80) SIGMA(s\*), Linear regression | 0.86 | 0.27 | 8 Tertiary amines | Lee and von Gunten (2012) |
| log(kHOCl) | log (k1amine) = 9.70 (+-0.32) - 2.10 (+-0.18) SIGMA(s\*) | 0.97 | 0.43 | 22 Primary amines & amine derivatives | Lee and von Gunten (2012) |
| log(kO3) | logkaromatic (O3) = 4.50 (+-0.38) + 0.96 (+-0.11) logkaromatic (HOCl), Linear regression  This QSPR can convert log(KO3) and log(KHOCl) into each other | 0.94 | 0.66 | 24 Aromatics | Lee and von Gunten (2012) |
| log(KHOCl) Second-order rate constants | log(k4) = 44.34.3 - 2.1SIGMA(s\*), Linear regression | 0.975 |  | 9 dissociated chloro- and bromophenols | Acero (2005) |
| log(KHOCl) | log kArO- = - (10.7 +- 2.2)Σσ+ o,m,p + 4.43 (+-0.35), Linear regression | 0.96 | - | 8 dissociated chlorophenols | Rule et al. (2005), Chen (2015) |
| log(KHOCl) (log(k3)) Rate constant for phenol + HOCl + H+ > products | log(k3) = 2.37-4.26 (+-0.29) SIGMA(s o,m,p), Linear regression | 0.99 | - | 4 Para-substituted phenols | Gallard (2002), Chen et al (2015) |
| log(KHOCl) (log(k3)) Rate constant for phenol + HOCl + H+ > products | log(k3) = 7.32-3.95 (+-0.29) SIGMA(s o,m,p), Linear regression | 0.989 | - | 4 Resorcinols | Gallard (2002), Chen et al (2015) |
| log(KHOCl) (log(k5)) Rate constant for phenolate + HOCl > products | log(k5) = 4.15-3.00 (+-0.22) SIGMA(s o,m,p), Linear regression | 0.951 | - | 11 Monovalent anionic monohydroxybenzenes | Gallard (2002) |
| log(KHOCl) (log(k5)) Rate constant for phenolate + HOCl > products | log(k5) = 6.71-3.36 (+-0.25) SIGMA(s o,m,p), Linear regression | 0.971 | - | 7 Monovalent anionic meta-dihydroxybenzenes | Gallard (2002) |
| log(KHOCl) (log(k3)) Rate constant for ED- + HOCl > products | log(k3) = 0.78pKa -3.63, Linear regression | 0.93 | - | 10 Monovalent anionic endocrine disruptors | Deborde (2004) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = 4.56 (± 0.45) N – 14.41 (± 1.95), Linear regression | 0.993 | - | 7 Inorganic anions  Swain–Scott plot. Nucleophilicity data (N), obtained from Hine (1962) | Gerritsen and Margerum (1990), Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = 1.14 (± 0.26) σ∗ + 7.24 (± 0.17), Linear regression | 0.884 | - | 14 Secondary and primary amines, T=25C | Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = -2.24 (± 0.82) σ∗ + 4.92 (± 0.32), Linear regression | 0.907 | - | 7 Tertairy amines, T=25C | Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = -11.82 (± 1.26) σi - 5.12 (± 0.72), Linear regression | 0.997 | - | 5: phenoxide ion, phenol, anisole and butylphenylether, T=22-25C | Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = -3.58 (± 1.14) Σ(σo,p,m) + 6.83 (± 0.42), Linear regression | 0.908 | - | 8 Substituted phenoxide ions (PhO- ) and 1,3-dihydroxybenzene anions (BOHO- and BO2 2- ),  T=22-25C | Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = -3.87 (± 0.66) Σσo,p,m + 1.93 (± 0.55), Linear regression | 0.882 | - | 22 Phenoxide ions (PhO-), 1,3-dihydroxybenzene anions (BOHO- and BO2 2- ),  T=22-25C | Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants at given pH or pH 7 (25 C) | log k = -3.77 (± 0.44) Σσo,p,m + 1.88 (± 0.37), Linear regression | 0.943 | - | 21 Phenoxide ions (PhO-), 1,3-dihydroxybenzene anions (BOHO- and BO2 2- ), but without Cl2BO2 2-,  T=22-25C | Deborde and von Gunten (2008) |
| log(KHOCl) (log kapp) Apparent rate constants log kHOCl | log kHOCl = 0.99 (± 0.08) log kO3 – 4.47 (± 0.63), Linear regression | 0.964 | - | 24 Aromatic compounds (mostly phenols) | Deborde and von Gunten (2008) |
| Log(KHOCl) Rate constant, mol-1 h-1 | log kHOCl = 29.44 EHOMO + 16.74 | 0.95 |  | 5 Phosphorothioates | Duirk et al. (2009), Mamy (2014) |
| Log(KHOCl) Rate constant, mol-1 h-1 | log kHOCl = 41.36 EHOMO + 20.45 | 0.99 |  | 3 Phosphorodithioates | Duirk et al. (2009), Mamy (2014) |
| log(KHOCl), log (k2))  HOCl + C6H4X(O-) > products (k2) | log k2 = -3.86 + 0.95 pKa, Linear regression | - | - | 34 Monovalent phenolates | Rebenne (1996), Soper and Smith |
| log(KHOCl) (log k)  Rate constant for k2 and k3 (monovalent/divalent anions), ArX(OH)O- and ArX(O-)2 | log k = -3.07 + 1.01pKa, Linear regression | 0.974 | - | 8 Monovalent/divalent anionic substituted resorcinols | Rebenne (1996) |
| Removal, %  Percentage removed by chlorination (HOCl/OCl) | % chlorine removal = 106.8 + 0.791%(ozone removal) + 7.89(#rtvFG) + 4.80(QPlog Pow) + 0.175(FISA) - 15.0(IP), Several 3D molecular descriptors and physicochemical properties for the development of multiple linear regression analysis | 0.705 | - | 73 Endocrine-disrupting compounds (EDCs) and pharmaceuticals and personal care products (PPCPs) | Lei and Snyder (2007) |
| log(K/n)  Measured rate constant per site of attack | log(k/n) = 3.2 - 2.4(SIGMA(s)avg), Linear least-squares fit,  Hammett-type correlation, where n is the number of sites for attack | 0.89 | - | 17 Substituted resorcinols and phenols | Rebenne (1996) |
| ΔrG and ΔG#  Energy barriers for N- and/or C-hydroxylation | Double-hybrid DFT method, double-hybrid DFT methods (B2PLYP-D and BK-PLYP) and a G3B3 composite scheme | - |  | 5? Amines: ammonia; methylamine;  diethylamine; piperidine; morpholine | Sakic (2015) |