High structural and molecular parameter diversity among chemicals with similar log P and log Koc values

Carr J Smith¹, Thomas A Perfetti², Gene M Ko³, and Suzanne B Hartigan⁴

Abstract
Chemicals possessing persistence (P) and high mobility (M) can present a hazard to drinking water resources by traversing natural barriers like riverbanks and artificial barriers found in water treatment plants. If the chemical is also toxic (T), i.e. classifiable as a PMT, the agent might be of particular concern as a potential drinking water contaminant. During routine water sampling, detection and quantitation of polar substances with high mobility can be problematic. The German Environment Agency (UBA) is considering the use of the Log Koc value as a proxy for mobility (M). Log Koc is related to Log P by the equation Log Koc = 0.69 Log P + 0.22. In this study, we demonstrate that chemicals with log P values at or very close to 2.0, 3.0 or 4.0 (and their concomitant log Koc values) can vary significantly in their chemical structures, molecular weights, molar volumes, and calculated molar refractivity (CMR), which is related to the mean polarizability of a molecule. The large degree of potential diversity in chemical structure and molecular parameters related to chemical behavior at a particular log P or log Koc value suggests that log Koc might not contain enough information to function as a standalone surrogate for the mobility (M) of a chemical, i.e. as related to its ability to move from a drinking water resource through the water plant purification process.

Keywords
Chemical mobility, log Koc, log P, QSAR, water quality

Introduction
Certain persistent (P) and mobile (M) chemicals inadvertently released into the environment can pass through natural barriers like riverbanks and soil layers above aquifers, and through artificial barriers found in water plants, eventually working their way into drinking water.¹ The time required for an environmental chemical to traverse a particular natural barrier and make its way to a water treatment plant varies greatly, from a few days for surface water run-off, 1 to 2 weeks to percolate through a riverbank, or up to the scale of years to enter groundwater wells.¹ Several factors can influence the time lag between a chemical’s release into the environment and its entry into drinking water. First, higher quantity emissions increase the probability of downstream contamination. Second, persistent chemicals (P) have a longer half-life thereby increasing the chance of eventually entering drinking water. Third, chemicals with higher mobility (M) are more likely to complete the journey from point of environmental entry to drinking water.¹

The German Environment Agency (UBA) is considering the use of the Log Koc value as a proxy for mobility (M).²–⁶ The Soil Adsorption Coefficient is denoted by either Kd or Kf. The Soil Adsorption Coefficient (Kd or Kf) measures the amount of a chemical adsorbed onto soil per amount of water. [The Kf designation comes from the Freundlich solid-water distribution coefficient.] Soil

¹ Albemarle Corporation, Charlotte, NC, USA
² Perfetti & Perfetti, LLC, Winston-Salem, NC, USA
³ Electromagnetic Systems Group, General Atomics, San Diego, CA, USA
⁴ American Chemistry Council, Washington, DC, USA

Corresponding author:
Carr J Smith, PhD, DABT, Albemarle Corporation, 6400 Brindlewood Court Mobile, Alabama, Charlotte, NC 36608, USA.
Email: carr.smith@albemarle.com

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Adsorption Coefficient values vary greatly because the organic chemical compositional percentage of soil varies greatly. Adsorption of chemicals onto soil occurs predominantly by partition into the soil organic matter. Therefore, the Kd or Kf is normalized to the amount of organic carbon in a soil and is expressed as either Koc or Kfoc, which are interchangeable for practical purposes. The Koc is known as the organic carbon-water partition coefficient. Log Koc is related to Log P (Log of the octanol/water partition coefficient) by the equation Log Koc = 0.69 Log P + 0.22.

In this study, we demonstrate that 3 sets of 25 chemicals with log P values at or very close to 2.0, 3.0 or 4.0 (and their concomitant log Koc values), respectively, can vary significantly in their Hansch Quantitative Structure-Activity Relationship (QSAR) parameters. The parameters examined in this study are the calculated base 10 logarithm of the octanol–water partition coefficient (ClogP), the McGo- wan molecular volume (MgVol), and the calculated molar refractivity (CMR). These QSAR parameters represent hydrophobic (ClogP) effects, steric/size (MgVol) considerations, and the total polarizability (CMR) of a molecule of the substance of a chemical on its biological activity. These parameters have proven extremely useful in developing QSAR models describing the quantitative relationships between the biological activity of chemicals and their physicochemical characteristics.

Previously, Smith et al.11 showed that smaller molecular volumes were found to be associated with higher levels of tumorigenicity. Lower rather than higher levels of lipophilicity were found to be associated with higher levels of tumorigenicity. Positive Ames test results were positively correlated with overall tumorigenicity and with possession of structural alerts of carcinogenicity. Since larger organic molecules have more chemical reaction centers, it was not surprising that higher ClogP values were positively correlated with the number of structural alerts of carcinogenicity. The results from this earlier study11 demonstrated the ability to devise rational rules for relative tumorigenicity in rodents that correlated with known parameters of toxicity.

Log P values of 2.0 represent 100-fold more solubility in lipid than in water, log P values of 3.0 represent 1000-fold more lipid solubility, and log P values of 4.0 represent 10,000-fold more lipid solubility. Thus, the 75 molecules (3 sets of 25 molecules) examined in this study span a wide range of lipid solubility. As log Koc is linearly related to log P, the range of log Koc values spanned in this study is also very large. The large degree of potential diversity in chemical structure and molecular parameters related to chemical behavior at a particular log P or log Koc value suggests that log Koc might not contain enough information to function as a standalone surrogate for the mobility (M) of a chemical as related to its ability to move from a drinking water resource through the water plant purification process.

### Methods

The octanol-water partition coefficients (P) for the 75 chemicals were reported in the literature and located by developing a web scraper searching for chemicals from the Super Natural II Database (http://bioinf-applied.charite.de/super-natural_new/index.php) with log P values of 2.0, 3.0 and 4.0.12 Super Natural II is a highly curated online database for natural products. Initially designed in 2006, the database contains 325,508 natural products extracted from various resources, including vendor information. Natural II offers both search and analysis options. It provides the toxicity prediction for the database compounds.12 The log Koc values were calculated from the literature values for log P using the equation10:

\[ \text{Log Koc} = 0.69 \text{ Log P} + 0.22 \]

### Calculation of Clog P, CMR and MgVol

Bio-Loom (version 1.6; Biabyte Corp., Claremont, CA, USA)13 was used to compute the three parameters used in our QSAR analysis from the simplified molecular input line entry system representation of each chemical compound: ClogP, CMR, and MgVol (Online Appendix 4). The utility of Bio-Loom for comparative QSAR (C-QSAR) analysis in comparative correlation analysis has been discussed in Hansch and Leo.14 The parameters used in this study are also discussed in detail in Hansch and Leo.14 In brief, ClogP is the calculated logarithm of the partition coefficient in octanol/water and is a measure of hydrophobicity (or lipophilicity) of a chemical.14,15 MgVol is the molar volume calculated by the method of Abraham and McGo- wan16,17 and CMR is the calculated molar refractivity (MR) for the whole molecule. MR is calculated as follows:

\[ \text{MR} = \left[ \left( n^2 - 1 \right) / \left( n^2 + 2 \right) \right] \times \left[ \text{MW} / d \right] \]

where n is the refractive index, MW is the molecular weight, and d is the density of a substance. Since there is very little variation in n,16 MR is largely a measure of volume with a small correction for polarizability. The MR values are scaled by 0.1. MR can be used for a substituent or for the whole molecule. Clog P and CMR are for the neutral form of partially ionized chemicals. CMR values obtained are calculated using the same program as that used to calculate ClogP.13 Note that the Clog P values are for the neutral form of acids and bases that may be partially ionized. If the degree of ionization is about the same for a set of congeners, the ionization factor can be neglected; otherwise, good correlation can be obtained using electronic terms.14,18 The correlation between experimental Log P and Clog P values for 13,815 chemicals in the CLOG program, which is a part of Bio-Loom,15 is 0.98 (experimental Log P = 1.00 Clog P − 0.03 (n = 13,815, r = 0.98, s = 0.35)). Clog P parameter that was used in this study has been widely used and cited by the QSAR community, both for environmental studies and for drug design.19–30 A very
high correlation \((r = 0.98)\) between experimental Log P and Clog P gives confidence in using Clog P values whenever experimental Log P values are not available.

**Statistical methods**

**Analysis of variance (ANOVA) for one factor.** The following scenario is considered. A set of data consisting of \(N\) scores \(x’s\) composed of \(n\) groups of equal size \(k\) each so that \(N = nk\). Index these scores as \(x_{i,j}\) for \(1 \leq i \leq n, 1 \leq j \leq k\). The score \(x_{i,j}\) falls in the \(i^{th}\) group and index \(j\) is for the \(j^{th}\) score \(x\) in the \(i^{th}\) group.

Let \(\bar{x}_i = \left(\frac{1}{k}\sum_{j=1}^{k} x_{i,j}\right)\) be the mean for the \(x’s\) in group \(i\).

Let \(\bar{x} = \left(\frac{1}{N}\sum_{i=1}^{n} \sum_{j=1}^{k} x_{i,j}\right)\) be the mean for all the \(x’s\).

A test for the statistical significance of differences in the means \(\bar{x}_i\) is requested. In section 40, Intraclass Correlation as an Example of the Analysis of Variance in Statistical Methods for Research Workers, by RA Fisher (Fifth Edition 1934), the following reasoning is employed.

First \(\frac{1}{N}\sum_{i=1}^{n} \sum_{j=1}^{k} (x_{i,j} - \bar{x})^2 = \frac{1}{k}\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2 + \frac{1}{N}\sum_{i=1}^{n} \sum_{j=1}^{k} (x_{i,j} - \bar{x}_i)^2\)

Thus, the total sum of squares \(\sum_{i=1}^{n} \sum_{j=1}^{k} (x_{i,j} - \bar{x})^2\) can be written as the between families \(k\) \(\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2\) plus the within families \(\sum_{i=1}^{n} \sum_{j=1}^{k} (x_{i,j} - \bar{x}_i)^2\) sums of squares.

Under the assumption (Null Hypothesis) that all the \(x’s\) are independent observations taken from the same population with a normal distribution (fixed mean and standard deviation) the total, between families, and within families sum of squares after each being divided by the variance of this normal distribution will have \(\chi^2\) distributions with \(nk - 1, n - 1\), and \(n(k - 1)\) degrees of freedom, respectively.

Fisher goes on to define the correlation \(\rho = A / (A + B)\), where \(A = A = \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2\right) / (n - 1) - B/k\) and \(B = \left(\frac{1}{k} \sum_{j=1}^{k} (x_j - \bar{x})^2\right) / (n(k - 1))\).

The modification \(F(n - 1, N - n) = \left(\frac{\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2}{n - 1}\right) / \left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{k} (x_{i,j} - \bar{x}_i)^2 / (N - n)}{N - n}\right)\) is used by Excel ANOVA.

**Results**

For the chemicals with log P values at or near 2.0 (for log P of 2, concomitant log Koc value is 1.6), the molecular weights vary widely from a low value of 150.104 to a high of 888.451 (Table 1). In ascending order, the molecular weights of the chemicals with log P values at or near 2.0 are as follows: 150.104; 150.105; 164.131; 220.085; 256.167; 260.105; 270.183; 288.16; 288.17; 308.116; 328.058; 348.157; 350.19; 354.067; 358.069; 376.225; 404.255; 420.19; 432.323; 480.183; 482.217; 483.226; 486.262; 594.377; and 888.451. For the chemicals with log P values at or near 3.0 (for log P of 3, concomitant log Koc value is 2.29), the molecular weights range from 125.12 to 694.263 (Table 1). In ascending order, the molecular weights of the chemicals with log P values at or near 3.0 are as follows: 125.12; 156.151; 198.068; 203.095; 222.126; 250.157; 252.173; 317.128; 328.067; 330.183; 345.214; 342.194; 352.154; 356.126; 385.213; 389.257; 395.185; 398.137; 424.271; 436.163; 476.241; 489.154; 493.273; 530.324; and 694.263. For the chemicals with log P values at or near 4.0 (for log P of 4, concomitant log Koc value is 2.98), the molecular weights range from 216.151 to 839.555 (Table 1). In ascending order, the molecular weights of the chemicals with log P values at or near 4.0 are as follows: 216.151; 218.167; 228.115; 250.063; 254.094; 259.194; 260.095; 266.182; 304.167; 320.046; 324.184; 339.183; 351.147; 357.984; 376.142; 377.126; 402.11; 439.307; 450.14; 451.076; 472.319; 486.173; 538.387; 583.295; and 839.555. In summary, chemicals sharing very similar log P values of 2.0, 3.0 and 4.0 display quite wide ranges of molecular weights indicative of the great diversity of chemical structures capable of having similar or the same log P or log Koc values.

For the chemicals with log P values at or near 2.0 (for log P of 2, concomitant log Koc value is 1.6), the volumes of 1 mole of each compound at Standard Temperature and Standard Pressure (MgVols) display a wide range from a low value of 1.34 to a high of 6.54 (Table 1). In ascending order, the MgVols of the chemicals with log P values at or near 2.0 are as follows: 1.34; 1.39; 1.41; 1.6; 1.92; 2.07; 2.12; 2.19; 2.21; 2.31; 2.32; 2.4; 2.42; 2.66; 2.68; 2.94; 3.03; 3.16; 3.5; 3.56; 3.58; 3.63; 3.78; 4.66; and 6.54. For the chemicals with log P values at or near 3.0 (for log P of 3, concomitant log Koc value is 2.29), the MgVols display a wide range from a low value of 1.25 to a high of 4.89 (Table 1). In ascending order, the MgVols of the chemicals with log P values at or near 3.0 are as follows: 1.25; 1.47; 1.54; 1.58; 1.77; 2.05; 2.1; 2.16; 2.34; 2.49; 2.54; 2.65; 2.71; 2.84; 2.96; 2.97; 3.18; 3.19; 3.36; 3.39; 3.53; 3.74; 4.24; and 4.89. For the chemicals with log P values at or near 4.0 (for log P of 4, concomitant log Koc value is 2.98), the MgVols fall over a wide range from a low value of 1.61 to a high value of 6.72 (Table 1). In ascending order, the MgVols of the chemicals with log P values at or near 4.0 are as follows: 1.61; 1.76; 1.86; 1.91; 1.93; 1.94; 1.96; 2.2; 2.22; 2.34; 2.38; 2.55; 2.69; 2.72; 2.74; 2.78; 2.83; 2.98; 3.09; 3.59; 3.62; 3.88; 4.4; 4.41; and 6.72. In summary, chemicals sharing very similar log P values of 2.0, 3.0 and 4.0 display quite wide ranges of molar volumes.
Table 1. A diverse set of chemicals with log p values at or very near 2.0, 3.0 or 4.0.

| logP | P     | logKoc | Koc  | ClogP | MVol | CMR     | Name                                                                 | Molecular Weight | Formula           | Chemical Class            | Structure          |
|------|-------|--------|------|-------|------|---------|----------------------------------------------------------------------|------------------|-------------------|--------------------------|--------------------|
| 1.99 | 97.72372 | 1.5931 | 39.18321 | 2.09  | 1.6  | 5.79    | N-(3-methylphenyl)-2-oxo-1,3-oxazolidine-3-carboxamide               | 220.085          | C11H12N2O3        | Aromatic amine           | ![Structure](image1) |
| 1.99 | 97.72372 | 1.5931 | 39.18321 | 1.82  | 2.94 | 10.25   | (1S,3S,4S,6S,9S,10S,11S,13S)-6,11-dihydroxy-5,5,9-trimethyl-14-methylidene-15-oxotetraacyclo[11.2.1.01,10.04,9]hexadecan-3-yl acetate | 376.225          | C22H32O5         | Oxotetraacyclohexadecanyl acetate | ![Structure](image2) |
| 1.99 | 97.72372 | 1.5931 | 39.18321 | 1.85  | 1.41 | 4.98    | 2-ethyl-3-methyl-5H,6H,7H,8H-imidazo[1,2-a]pyridine                  | 164.131          | C10H16N2        | Imidazopyridine            | ![Structure](image3) |
| 1.99 | 97.72372 | 1.5931 | 39.18321 | 1.9   | 2.4  | 8.87    | 2-(3,4-dihydroxyphenyl)-5-hydroxy-4-oxochromen-7-yl (2R)-2-hydroxypropanoate | 358.069          | C18H14O8         | Oxochromenyl-hydroxypropanoate | ![Structure](image4) |
| 1.99 | 97.72372 | 1.5931 | 39.18321 | 2.07  | 1.34 | 4.79    | UNPD31574                                                               | 150.104          | C10H14O          | Cyclohexadiene            | ![Structure](image5) |

(continued)
| logP  | P    | logKoc | Koc  | ClogP | M&vol | CMR | Name                                                                 | Molecular Weight | Formula | Chemical Class                  | Structure |
|-------|------|--------|------|-------|--------|-----|----------------------------------------------------------------------|------------------|---------|---------------------------------|-----------|
| 1.9902| 97.76874 | 1.593238 | 39.19566 | 2.03  | 3.03   | 10.69 | (2 S)-N-phenyl-2-((1 S,2 R,6 R,8 R,9 R)-4,4,11,11-tetramethyl-3,7,10,12-pentaoxatricyclo[7.3.0.02,6]dodecan-8-yl)formamido propanamide | 420.19          | C21H28N2O7 | Amide                        |           |
| 1.991 | 97.949 | 1.59379 | 39.24551 | 1.86  | 2.31   | 8.65  | 4-methyl-1-(2-oxo-2-phenylethyl)-3H-1,4-benzodiazepine-2,5-dione       | 308.116         | C18H16N2O3 | Benzodiazepinedione             |           |
| 1.991 | 97.949 | 1.59379 | 39.24551 | 1.86  | 1.92   | 6.92  | (8 S)-8-hydroxy-6,6,8-trimethyl-5 H,7 H-azuleno[5,6-c]furan-4,9-dione | 260.105         | C15H16O4 | Azulenosurandione               |           |
| 1.9926| 98.31052 | 1.594894 | 39.3454 | 1.92  | 2.42   | 9.49  | (3aR,6aR)-5-(2,3-dimethylphenyl)-3-(thiophene-2-carbonyl)-3aH,6aH-pyrrolo[3,4-d][1,2]oxazole-4,6-dione | 354.067         | C18H14N2O4S | Pyrroloxazolinedione           |           |
| logP | P   | logKoc | Koc  | ClogP | MgVol | CMR   | Name                                                                 | Molecular Weight | Formula       | Chemical Class                  | Structure                                                                 |
|------|-----|--------|------|-------|-------|-------|----------------------------------------------------------------------|------------------|--------------|---------------------------------|--------------------------------------------------------------------------|
| 2    | 100 | 1.6    | 39.81 | 1.39  | 2.12  | 3.63  | 12.98 Methyl [(2 R)-2-[(2 R)-1-acetyl-4-(phenylcarbamoyl)piperazin-2-yl]formamido]-3-(4-methoxyphenyl)propanoate | 482.217          | C25H30N4O6   | Piperazinylformamidomethoxyphenylpropanoate                          | ![Chemical Structure](https://example.com/structure1.png)               |
| 2    | 100 | 1.6    | 39.81 | 1.92  | 2.66  | 9.37  | (3aS,4 R,6 R,7 R,7aS)-6-ethenyl-3a-hydroxy-6-methyl-3-methyliden-2-oxo-7-(prop-1-en-2-yl)-tetrahydro-1-benzofuran-4-yl 2-(hydroxymethyl)prop-2-enoate | 348.157          | C19H24O6     | Benzofuran propanoate                                                     | ![Chemical Structure](https://example.com/structure2.png)               |
| 2    | 100 | 0.92   | 39.83 | 1.97  | 2.33  | 7.49  | Methyl [(3 S,4 E,6 S)-3-hydroxy-6-isopropyl-3-methyl-9-oxodec-4-enoate | 270.183          | C15H26O4     | Methyloxodecenoate                                                      | ![Chemical Structure](https://example.com/structure3.png)               |

(continued)
| logP | P   | logKoc | Koc | ClogP | MgVol | CMR | Name                                                                 | Molecular Weight | Formula                  | Chemical Class                        | Structure |
|------|------|--------|-----|-------|-------|-----|---------------------------------------------------------------------|-------------------|--------------------------|---------------------------------------|-----------|
| 2.001| 100.2305| 1.60069| 39.87402| 3.58 | 12.81 | (1 S,2 R,5 S,7 R,11 S,12 S,15 S,16 S)-15-[(1 R)-1-[(1 S,4 R,6 S)-1,6-dimethyl-2-oxa-3,7-dioxabicyclo[4.1.0]heptan-4-yl]-1-hydroxyethyl]-6-hydroxy-2,16-dimethyl-8-oxapentacyclo[9.7.0.02,7.07,9.012,16]octadec-4-en-3-one | 486.262          | C28H38O7                 | Oxapentacyclooctacendecane                                      |           |
| 2.001| 100.2305| 1.60069| 39.87402| 1.95 | 6.96  | (3 R,4 R,5 R,7 S)-4,5-dihydroxy-3,5,7,9-tetramethyl-1-oxacycloundec-9-en-2-one | 256.167          | C14H24O4                 | Oxacycloundecane                                                |           |
| 2.0014| 100.3229| 1.600966| 39.89937| 3.56 | 12.49 | (1 S,3aS,5 S,7aR)-5-hydroxy-3,3,5-trimethyl-N-[(1-methylpyrrol-2-yl)methyl]-7a-[3-(morpholin-4-yl)-3-oxopropyl]-hexahydroinden-1-aminium | 432.323          | C25H42N3O3               | Morpholinylxopropylhexahydroindenaminium ion                  |           |
| 2.0014| 100.3229| 1.600966| 39.89937| 2.08 | 9.84  | (2 R,4 S,5 S)-5-[(thiophen-2-ylformamido)methyl]-1-azabicyclo[2.2.2]octan-1-ium | 350.19           | C18H128N3O2S             | Azabicyclooctanium ion                                              |           |
| 2.0016| 100.3691| 1.601104| 39.91205| 3.78 | 13.57 | (2E)-N-[(2 R,3 S)-2,3-dihydroxy-8-[(2 S,5 R)-3-[(6 S)-6-hydroxyoct-7-en-2,4-dyn-1-yl]oxiran-2-yl]oxyl]-3-(3,4-dihydroxyphenyl)prop-2-erasamide | 483.226          | C27H33NO7               | Dihydroxyphenylpropenamide                                         |           |
| logP | P   | logKoc | Koc  | ClogP | MgVol | CMR | Name                                                                 |
|------|------|--------|------|-------|-------|-----|----------------------------------------------------------------------|
| 2.0017 | 100.3922 | 1.601173 | 39.91839 | 1.82  | 3.5   | 13.05 | N-[(3 S,7 S,10 R,12 S)-4-(2,5-dimethylbenzoyl)-2,9-dioxo-1,4,8-triazatricyclo[8.3.0.0\(3,7\)]tridecan-12-yl]-2-(thiophen-2-yl)acetamide |
| 2.002 | 100.4616 | 1.60138 | 39.93742 | 1.8   | 2.07  | 7.94  | (2 R,3 R)-2,5,7-trihydroxy-2H-9',11'-dioxaspiro[1-benzopyran-3,4'-cyclo[6.3.0.0\(2,7\)]undecane]-1',3'(6'),7'-trien-4-one |
| 2.002 | 100.4616 | 1.60138 | 39.93742 | 1.94  | 6.54  | 22.65 | (1 R,2 R,5 S,10 S,11 R,14 S,15 S,16 R)-14-acetyl-10,11,14-trihydroxy-5-[(2 R,4 S,5 S,6 S)-4-hydroxy-5-[(2 S,4 R,5 S,6 R)-6-hydroxy-4-methoxy-6-methyloxan-2-yl]oxy]-6-methyloxan-2-yl]oxy]·2.15-dimethyltetraacyclo[8.7.0.0\(2,7\)]heptadecen-14-yl]-6-methylheptan-4-one |
| 2.002 | 100.4616 | 1.60138 | 39.93742 | 2.1   | 4.66  | 15.83 | (2 R)-2-[(1 S,2 S,3 R,5 S,9 R,10 S,11 S,13 S,14 S,15 S,5 S)-2,15-dimethyl-tetraacyclo[8.7.0.0\(2,7\)]heptadecen-14-yl]-6-methylheptan-4-one |

(continued)
| logP    | P      | logKoc | Koc  | ClogP | MgVol | CMR | Name                                                                                   | Molecular Weight | Formula          | Chemical Class                           | Structure                                                                 |
|---------|--------|--------|------|-------|-------|-----|----------------------------------------------------------------------------------------|-------------------|------------------|------------------------------------------|--------------------------------------------------------------------------|
| 2.0023  | 100.531| 1.601587| 39.95646| 1.84  | 2.19  | 8.06 | (2R,4S,5R)-5-[5-(furan-2-yl)-2-methyl[1,3]oxazol-3-yl]-2-(hydroxymethyl)-1-azabicyclo[2.2.2]octan-1-ium | 288.171          | C16H22N3O2      | Azabicyclooctanium ion                  |                                                                           |
| 2.0024  | 100.542 | 1.601656| 39.96281| 2.11  | 2.19  | 8.15 | (1S,12S,14R)-14-hydroxy-9-methoxy-4-methyl-11-oxa-4-azatetracyclo[8.6.1.01,12.06,17]heptadeca-6,8,10(17),15-tetraen-4-ium | 288.16           | C17H22NO3       | Azatetracycloheptadeca-tetraenium ion    |                                                                           |
| 2.0024  | 100.542 | 1.601656| 39.96281| 1.89  | 3.16  | 11.31| 1-[(2S,5S,6S)-6-(cyclopropylcarbamoyl)-5-hydroxyoxan-2-yl]ethyl-4-(2-methoxyphenyl)piperazinium | 404.255          | C22H34N3O4      | Methoxyphenypiperazinium ion             |                                                                           |
| 2.99    | 977.2372| 2.2831 | 19.19111| 3.13  | 2.71  | 9.31 | (2R,3R,4R,4aS,8aS)-4-[2-(furan-3-yl)ethyl]-2,4-dihydroxy-3,4a,8,8a-tetramethyl-hexahydronaphthalen-1-one | 334.214          | C20H30O4       | Tetramethyl-hexahydronaphthalene         |                                                                           |
Table 1. (continued)

| logP  | P    | logKoc | Koc  | MolVol | CMR | Name                                                                                       |
|-------|------|--------|------|--------|-----|--------------------------------------------------------------------------------------------|
| 2.9901| 977.4623 | 2.283169 | 191.9416 | 3.2   | 4.79 | (1 R,3 S)-1-isopropoxy-3-methylcyclohexane                                                   |
| 2.9902| 977.6874 | 2.283238 | 191.972 | 3.08  | 2.65 | N’-[(1E)-(4-hydroxy-3-methoxyphenyl)methylidene]-3-(2-methyl-1,3-benzodiazol-1-yl)propanehydrazide |
| 2.9903| 977.9125 | 2.28307 | 192.0026 | 2.81 | 2.97 | (3 R,6'R,7'S,8'aS)-7'-(1,3-dimethoxy-3-oxoprop-1-en-2-yl)-6'-ethyl-2-oxo-2',3',4,5,6,7,8'B'octahydro-1'H-spiro[indole-3,1'-indolizin]-4'-ium |
| 2.9908| 979.039  | 2.283652 | 192.1551 | 2.82 | 3.19 | (2 S)-1-(4-hydroxypiperidin-1-yl)-2-((2 R)-7-(2-methoxyethoxy)-5,8-dimethyl-1,2,3,4-tetrahydrophenanthren-2-yl)propan-1-one |
| 2.991 | 979.49 | 2.28379 | 192.2162 | 3.16 | 2.49 | (1 S,8 R,12 R,13 R,14 R)-8,15,15-trimethyl-6,19-dioxapentacyclo[10.6.1.01,14.03,11.05,9]nonadeca-3(11),4,9-triene-4,13-diol |
| 2.991 | 979.49 | 2.28379 | 192.2162 | 3.16 | 2.1  | (3 R,3aS,8 S,8aS)-8-hydroxy-3-isopropyl-8a-methyl-2,3a,6,7,8-hexahydro-1'H-azulene-5-carboxylic Acid |

(continued)
| logP  | P    | logKoc | Koc  | ClogP | MVol | CMR | Name                                                                 |
|-------|------|--------|------|-------|------|-----|----------------------------------------------------------------------|
| 2.991 | 979.49 | 2.28379 | 192.2162 | 3.09  | 2.84 | 10.32 | 3-[(2S,4R,5S,6R)-4,5-dihydroxy-6-methyloxan-2-yl]oxy]-1-ethyl-8-hydroxyanthracene-9,10-dione |
| 2.991 | 981.2959 | 2.284342 | 192.4607 | 3.74  | 13.9 | 493.273 | (2S,5aS,8aR)-2-[(4-fluorophenyl)methyl]carbamoyl-1-methyl-6-[[1-methyl-1,3-benzodiazol-2-yl]methyl]-5-oxo-octahydropyrrolo[3,2-a][1,4]diazepin-1-ium |
| 2.992 | 981.7479 | 2.28448 | 192.5218 | 2.86  | 3.53 | 476.241 | Methyl (1S,2S,4R,7R,8S,11R,12R,16R,16-pentamethyl-5,15-dioxapentacyclononadecane-7-carboxylate |
| 2.992 | 981.7479 | 2.28448 | 192.5218 | 2.94  | 7.77 | 328.067 | (1S,3R,10S,12S)-12-bromo-1,11,11-trimethyl-2,5-dioxatricyclo[8.4.0.03,7]tetradec-7-en-6-one |
| 2.992 | 981.7479 | 2.28448 | 192.5218 | 2.83  | 1.25 | 125.12 | Octanonitrile |

(continued)
| logP | P  | logKoc | Koc  | ClogP | MVol | CMR | Name                                                                 |
|------|-----|--------|------|-------|------|-----|----------------------------------------------------------------------|
| 2.992 | 981.7479 | 2.28448 | 192.2518 | 2.54 | 9.36 | (3S)-4',5,5',7-tetramethoxy-2H-spiro[1-benzopyran-3,7'-bicyclo[4.2.0]octane]-1(6')2',4'-trien-4-one | 356.126 | C20H12O6 | Spiro benzopyran/bicyclo octane tranone |
| 2.992 | 981.7479 | 2.28448 | 192.2518 | 1.54 | 5.85 | 4-hydroxybenzophenone                                                | 198.068 | C13H10O2 | Hydroxybenzophenone |
| 2.9923 | 982.4263 | 2.284687 | 192.26136 | 2.65 | 9.61 | (1S,9R,10R,17R)-N,N-diethyl-16-oxo-2-oxa-15-azatetracyclo[7.5.3.01,10.03,8]heptadeca-3,5,7-triene-17-carboxamide | 342.194 | C20H26N2O3 | Azatetracycloheptadecatriene carboxamide |
| 3 | 1000 | 2.29 | 194.9845 | 4.89 | 17.67 | 2-(2R,4R,5R,6R)-4-[(2R,5S,6S)-4,5-dihydroxy-6-methyloxan-2-yl]oxy]-6-methyloxan-2-yl]oxy]-5-hydroxy-6-methyloxan-2-yl]-1,6,10-trihydroxy-8-methyltetraene-5,12-dione | 694.263 | C37H42O13 | Methyltetraenionedione |
| 3 | 1000 | 2.29 | 194.9845 | 2.98 | 2.05 | 2-(2R,4aR)-8-(hydroxymethyl)-4a-methyl-2,3,4,5,6,7-hexahydro-1H-naphthalen-2-yl]propenoic Acid | 250.157 | C15H12O3 | Hexahydonaphthalenyl propenoic acid |
| 3 | 1000 | 2.29 | 194.9845 | 3.01 | 4.24 | Methyl (2R,6R)-6-[(2S,5S,7R,11R,12S,14R,15R)-5,12-dihydroxy-2,6,6,11,15-pentamethyl-9,17-dioxotetraacyclo[8.7.0.02,7.011,15]heptadec-1(10)-en-14-yl]-2-methyl-4-oxoheptanoate | 530.324 | C31H46O7 | Dioxotetraacycloheptadecenyl]-2-methyl-4-oxoheptanoate |
| logP  | P     | logKoc | Koc  | ClogP | MgVol | CMR | Name                                                                 | Molecular Weight | Formula       | Chemical Class                  | Structure |
|------|-------|--------|------|-------|-------|-----|---------------------------------------------------------------------|------------------|---------------|---------------------------------|-----------|
| 3.0005 | 1001.152 | 2.290345 | 195.1394 | 2.9 | 1.58 | 5.83 | 5,5-dimethyl-4-methylidene-3-phenyl-1,3-oxazolidin-2-one | 203.095 | C12H13NO2 | Oxazolidinone                 | ![Structure](image1) |
| 3.0014 | 1003.229 | 2.290966 | 195.4186 | 3.04 | 3.39 | 12.3 | (2R,4S,5R)-5-[6-[4-(dimethylamino)phenyl]-2-methylpyrimidin-4-yl]-2-[[2-methoxyacetamido]methyl]-1-azabicyclo[2.2.2]octan-1-ium | 424.271 | C24H34N5O2 | Azabicyclooctanium ion        | ![Structure](image2) |
| 3.0017 | 1003.922 | 2.291173 | 195.5118 | 3.06 | 2.96 | 10.84 | (11R)-13-[(3,3-dimethylbutanoyl)-5-(furan-2-yl)-1,9,13-triazatricyclo[9.4.0.0^3,8]pentadeca-3,5,7-triene-2,10-dione | 395.185 | C22H25N3O4 | Triazatricyclopentadecatrienedione | ![Structure](image3) |
| 3.0017 | 1003.922 | 2.291173 | 195.5118 | 3.14 | 2.34 | 9.24 | 1-methyl-2-{4-methylquinazolin-2-ylamino]quinazolin-4-one | 317.128 | C18H15N5O | Methylquinazolinylaminoquinazoline | ![Structure](image4) |
| logP  | P      | logKoc | Koc | ClogP | MyVol | CMR | Name                                                                 |
|-------|--------|--------|-----|-------|-------|-----|----------------------------------------------------------------------|
| 3.0019| 1004.384| 2.29131| 195.5739| 3.18  | 11.79 |     | 2-[[5-hydroxy-4-oxo-2-phenylchromen-7-yl)oxy]-N-[3-(2-oxopyrrolidin-1-yl)propyl]acetamide |
| 3.002 | 1004.616| 2.29138| 195.605  | 3.13  | 6.22  |     | (2 S)-5-(2H-1,3-benzodioxol-5-yl)-2-methylpentan-1-ol                  |
| 3.0021| 1004.847| 2.291449| 195.6361| 3.19  | 12.7  |     | 2-[[furan-2-ylmethyl]-N-{4-methoxy-6-methyl-2 H5 H7 H8M-{1,3}dioxolo[4,5-g]isoquinolin-9-yl]-1,3-dioxoisoindole-5-carboxamide |
| 3.99  | 9772.372| 2.9731 | 939.9397| 3.85  | 2.34  | 8.3 | (2 S,4E,6E)-12-(pyridin-3-yl)dodeca-4,6-dien-2-ol                      |
| 3.99  | 9772.372| 2.9731 | 939.9397| 4.17  | 4.41  | 14.89 | (1 S)-4-[[2 R,4aR,6 R,8aS)-6-[[2 S,5 R)-5-(2-hydroxyprop-2-yl)-2-methyloxolan-2-yl]-Ba-methyl-6-hydroxy-2H-pyran[3,2-b]pyran-2-yl]-1-[(2 R5 R)-5-(2-hydroxypropan-2-yl)-2-methyloxolan-2-yl]pent-4-en-1-ol |

(continued)
| logP | P    | logKoc | Koc  | ClogP | MgVol | CMR | Name                                                                 |
|------|------|--------|------|-------|-------|-----|----------------------------------------------------------------------|
| 3.9902 | 9776.874 | 2.973238 | 940.2384 | 4.04  | 3.62  | 12.96 | (3 R,4 S)-3-[(5-(4-tert-butylphenyl)-1,2-oxazol-3-yl)methyl]-4-[2-(4-methylpiperazin-1-yl)-2-oxoethyl]piperidinium ion |
| 3.9906 | 9785.883 | 2.973514 | 940.8362 | 3.88  | 2.2   | 8.77  | 2-[(2-bromophenyl)methylidene]-3-oxo-1-benzofuran-6-yl Acetate        |

(continued)
### Table 1. (continued)

| logP | P   | logKoc | Koc  | ClogP | MVol | CMR | Name                                                                 |
|------|-----|--------|------|-------|------|-----|---------------------------------------------------------------------|
| 3.99 | 9794.9 | 2.97379 | 941.4343 | 4.03  | 2.98  | 10.8 | ((3 R,3aS,6S,6aR)-6-(4-methylbenzamido)-hexahydrofurano[3,2-b]furan-3-yl N-[3-(trifluoromethyl)phenyl]carbamate |
| 3.99 | 9794.9 | 2.97379 | 941.4343 | 4.05  | 1.96  | 6.94 | (7 R)-4,10-dimethylidene-7-(prop-1-en-2-yl)cyclodec-5-en-1-one        |
| 3.99 | 9797.155 | 2.973859 | 941.5838 | 3.85  | 2.55  | 9.43 | (4aR,10aS)-6,6-dimethyl-2-phenyl-1,4,5,6,7,8-hexahydrophthalazine-1,4-dione |
| 3.99 | 9794.9 | 2.97379 | 941.4343 | 3.9   | 1.93  | 6.84 | 2-[2(R)-5,6-dimethyl-1,2,3,4-tetrahydronaphthalen-2-yl]propan-2-ol     |
| 3.99 | 9794.9 | 2.97379 | 941.4343 | 3.85  | 2.55  | 9.43 | (4aR,10aS)-6,6-dimethyl-2-phenyl-3 H,4aH5 H7 H8 H9 H,10 H,10aH-cyclohexa[4]phthalazine-1,4-dione |

Molecular Weight: 450.14  
Formula: C22H21F3N2O5  
Chemical Class: Hexahydrofuranyl N-(trifluoromethyl)phenyl carbamate  
Structure: ![Structure Image]
### Table 1. (continued)

| logP | P    | logKoc | Koc  | ClogP | MVol | CMR  | Name                                                                 | Molecular Weight | Formula     | Chemical Class                      | Structure                        |
|------|------|--------|------|-------|------|------|----------------------------------------------------------------------|------------------|-------------|-------------------------------------|----------------------------------|
| 3.992 | 9817.479 | 2.97448 | 942.9312 | 3.97  | 2.78 | 10.67 | 2-[(3,4-dimethyl-2-oxochromen-7-yl)oxy]-N-(1-methylindol-4-yl)acetamide | 376.142          | C22H12O2N4O4 | Methylindolyl acetamide             | ![Methylindolyl Acetamide](image1) |
| 3.992 | 9817.479 | 2.97448 | 942.9312 | 3.86  | 2.72 | 10.14 | Dimethyl[2-(2,3,4-trimethoxyphenanthren-1-yl)ethyl]amine              | 339.183          | C21H25NO3  | Trimethoxyphenanthryl ethylamine    | ![Trimethoxyphenanthryl Ethylamine](image2) |
| 3.992 | 9817.479 | 2.97448 | 942.9312 | 4.17  | 2.83 | 10.82 | (7 S,13 R)-9-hydroxy-13-(4-hydroxyphenyl)-7-phenyl-4,14-dioxatricyclo[8.4.0.0^2,7.0^11,15]tetradec-1,3(8),9-triene-5,11-dione | 402.11           | C24H18O6   | Dioxatricycloheptadeca-dienedione   | ![Dioxatricycloheptadeca-dienedione](image3) |
| 3.992 | 9817.479 | 2.97448 | 942.9312 | 3.81  | 3.88 | 13.47 | (2 S,5 S,6 S,7 R,11 S,15 R)-5-hydroxy-6-(hydroxymethyl)-14-[(2 R,5 R)-5-methoxy-4-oxoheptan-2-yl]-2,11,15-trimethyloctadecylo[8.7.0.0^2,7.0^11,15]heptadec-1(10),13-dien-12-one | 472.319          | C29H44O5   | Trimethyltricycloheptadeca-dienone   | ![Trimethyltricycloheptadeca-dienone](image4) |
| logP | P   | logKoc | Koc | ClogP | MgVol | CMR | Name                                                                 | Molecular Weight | Formula     | Chemical Class      | Structure                                                                 |
|------|------|--------|-----|-------|--------|-----|----------------------------------------------------------------------|------------------|------------|---------------------|--------------------------------------------------------------------------|
| 4    | 1000 | 2.98   | 954.9926 | 4.04  | 3.59   | 13.8 | (4 R,7 S,8aS)-4-[2-(benzylcarbamoyl)ethyl]-7-[(4-chlorophenyl)carbamothioyl]amino]-1-oxo-octahydropyrrolo[1,2-a]pyrazin-5-um | 486.173         | C24H29ClN5O2S | Oxo-octahydropyrrolo[1,2-a]pyrazinium ion | ![Structure](image1) |
| 4.008 | 10018.44 | 2.980552 | 956.2072 | 3.86  | 1.76   | 7.22 | 15-methyl-9,13-dioxatetracyclo[8.7.0.0^2,7.0^12,16]heptadeca-1(10),2,4,6,11,14,16-heptadecan-8-one | 250.063         | C16H10O3      | Methyl-dioxatetracycloheptadeca–heptadecan-8-one | ![Structure](image2) |
| 4.001 | 10023.05 | 2.98069  | 956.5111 | 3.83  | 2.38   | 8.27 | (4 S,4aR,5 S,8aR)-4a,5-dimethyl-9-oxo-4 H,5 H,6 H,7 H,8 H,8aH-naphtho[2,3-b]furan-4-yl 2-methylpropanoate | 304.167         | C18H24O4      | Naphthofuranyl 2-methylpropanoate | ![Structure](image3) |
| 4.0012 | 10027.67 | 2.980828 | 956.8151 | 3.91  | 3.09   | 12.12 | (11 R)-5-(4-chlorophenyl)-13-(thiophene-2-carbonyl)-1,9,13-triazatricyclo[9.4.0.0^3,8]pentadeca-3,5,7-triene-2,10-dione | 451.076         | C23H18ClN3O3S | Triazatricyclopentadecatriene-dione | ![Structure](image4) |
| 4.0012 | 10027.67 | 2.980828 | 956.8151 | 3.98  | 2.22   | 8.23 | 1-(adamantan-1-yl)-3-buthylthiourea | 266.182         | C15H26N2S    | Admantanyl butylthiourea | ![Structure](image5) |
| logP  | P    | logKoc | Koc  | ClogP | MgVol | CMR | Name                                                                 |
|------|------|--------|------|-------|-------|-----|----------------------------------------------------------------------|
| 4.0014 | 10032.29 | 2.980966 | 957.1191 | 3.97  | 1.61  | 4.68 | Tert-butyl 3,3,3-trifluoro-2,2-bis(trifluoromethyl)propanoate         |
|       |      |        |      |       |       |     | Bis(trifluoromethyl)propanoate                                       |
| 4.0016 | 10036.91 | 2.981104 | 957.4233 | 3.97  | 4.4   | 15.54 | 2-[(1 S,2 R,10 S,11 S,14 R,15 S,17 S)-14,17-dihydroxy-2,15-dimethyl-5-oxatetracyclo[8.7.0.0^2,7.0^11,15]heptadec-6-en-14-yl]-2-oxoethyl 3-[(2-(4-fluorophenyl)ethyl)carbamoyl]propanoate |
|       |      |        |      |       |       |     | Oxoethyl fluorophenyl ethyl carbamoyl propanoate                      |
| 4.0019 | 10043.84 | 2.981311 | 957.8798 | 3.85  | 2.69  | 10.01 | N-(2-methylpropyl)-2-[(2-oxo-4-phenylchromen-7-yl)oxy]acetamide      |
|       |      |        |      |       |       |     | Oxophenylchromenyl oxacetamide                                       |
| 4.0019 | 10043.84 | 2.981311 | 957.8798 | 4.12  | 6.72  | 23.32 | (2 S,4aS,6 S,6aS,6bR,8aS,9 R,10 S,11 R,12aR,12bR,14bS)-2,9-bis(hydroxymethyl)-2,6a,6b,9,12a-pentamethyl-11-[(4-[[methylammonio]methyl]phenyl)methoxy]-6-penyl-10-[(2 S,3 R,4 S,5 R)-3,4,5-trihydroxyoxan-2-yl]oxy)-1,3,4,5,6,7,8,8a,10,11,12a,12b,13,14b-tetradecahydro pyranose |
|       |      |        |      |       |       |     | Pyranose                                                             |
Table 1. (continued)

| logP | P     | logKoc | Koc  | ClogP | MgVol | CMR | Name                                      | Molecular Weight | Formula | Chemical Class                  | Structure |
|------|-------|--------|------|-------|-------|-----|-------------------------------------------|------------------|---------|----------------------------------|-----------|
| 4.002 | 10046.16 | 2.98138 | 958.032 | 3.84 | 1.86 | 6.9 | 4-[(3-methoxyphenyl)ethyl]phenol          | 228.115          | C15H16O2 | Methoxyphenethyl phenol          | ![](image) |
| 4.0028 | 10064.68 | 2.981932 | 959.2504 | 3.87 | 1.91 | 7.36 | Chalepensin                               | 254.094          | C16H14O3 | Furanocoumarin                  | ![](image) |
| 4.0028 | 10064.68 | 2.981932 | 959.2504 | 3.87 | 1.94 | 7.93 | 1-[indolo[3,2-b]quinolin-10-yl]ethanone   | 260.095          | C17H12N2O | Indoloquininyl estanone          | ![](image) |

\[ \log K_{oc} = 0.69 \log P + 0.22 \]
indicative of the great diversity of chemical structures capable of having similar or the same log P or log Koc values.

For the chemicals with log P values at or near 2.0 (for log P of 2, concomitant log Koc value is 1.6), the CMRs of each compound display a wide range from a low value of 4.79 to a high of 22.65 (Table 1). In ascending order, the CMRs of the chemicals with log P values at or near 2.0 are as follows: 4.79; 4.98; 5.08; 5.79; 6.92; 6.96; 7.49; 7.94; 8.06; 8.15; 8.65; 8.87; 9.37; 9.49; 9.84; 10.25; 10.69; 11.31; 12.49; 12.81; 12.98; 13.05; 13.57; 15.83; and 22.65. For the chemicals with log P values at or near 3.0 (for log P of 3, concomitant log Koc value is 2.29), the CMRs display a wide range from a low value of 3.9 to a high of 17.67 (Table 1). In ascending order, the CMRs of the chemicals with log P values at or near 3.0 are as follows: 3.9; 4.79; 5.83; 5.85; 6.22; 7.15; 7.17; 7.77; 9.08; 9.24; 9.31; 9.36; 9.61; 10.09; 10.32; 10.83; 10.84; 11.08; 11.79; 12.1; 12.3; 12.7; 13.9; 14.27; and 17.67. For the chemicals with log P values at or near 4.0 (for log P of 4, concomitant log Koc value is 2.98), the CMRs fall over a wide range from a low value of 4.68 to a high value of 23.32 (Table 1). In ascending order, the CMRs of the chemicals with log P values at or near 4.0 are as follows: 4.68; 6.84; 6.9; 6.94; 7.22; 7.36; 7.93; 8.23; 8.27; 8.3; 8.77; 9.43; 10.01; 10.14; 10.43; 10.67; 10.8; 10.82; 12.12; 12.96; 13.47; 13.8; 14.89; 15.54; and 23.32. In summary, chemicals sharing very similar log P values of 2.0, 3.0 and 4.0 display quite wide ranges of CMRs indicative of the great diversity of chemical structures capable of having similar or the same log P or log Koc values.

Table 2 shows the results from an analysis of variance (ANOVA) for one factor in which the means of the molecular parameters MgVol, CMR, and molecular weight were shown to not be related to the log Koc values derived from log Ps at or very near to 2, 3 or 4. The p value for the comparison of the means of MgVol values at the three different log Koc values was 0.896056. The p value for the comparison of the means of CMR values at the three different log Koc values was 0.759057. The p value for the comparison of the means of molecular weights at the three different log Koc values was 0.889793. Therefore, MgVol, CMR and molecular weight are not related to log Koc.

Discussion

There are additional examples that illustrate the utility of employing Hansch molecular parameters toward better understanding of complex toxicological issues impacting the environment. Previously, Garg and Smith conducted a QSAR study to address an important problem encountered in the prediction of the bioconcentration factor (BCF) of highly hydrophobic chemicals. They noted that the linear relationship between the BCF and hydrophobic parameter, i.e. calculated octanol-water partition coefficient (ClogP), breaks down for highly hydrophobic chemicals. Their results suggested that a non-linear relationship between BCF and the hydrophobic parameter, along with inclusion of additional molecular size, weight and/or volume parameters, should be considered while developing a QSAR model for more reliable prediction of the BCF of highly hydrophobic chemicals.

In the current study, 75 randomly selected compounds with log P values at or very near 2.0, 3.0 or 4.0 were selected searching the Super Natural II database with a web scraper. Many more chemicals could have been selected via the same method, but 25 chemicals in each log P/log Koc category well illustrated the high degree of structural diversity that can occur at the same log P/log Koc value. The width of the molecular weight ranges found within a given log P/log Koc value were notable, i.e. 150.104-888.451 at log P = 2/log Koc = 1.6; 125.12-694.263 at log P = 3/log Koc = 2.29; 216.151-839.555 at log P = 4/log Koc = 2.98. There are similarly large ranges within a given log P/log Koc value for molar volumes and for CMRs, which are related to polarizability. The German Environment Agency (UBA) is considering the use of the Log Koc value as a proxy for mobility (M). The large degree of potential diversity in chemical structure and molecular parameters related to chemical behavior at a particular log P or log Koc value suggests that log Koc might not contain enough information to function as a standalone surrogate for the mobility (M) of a chemical as related to its ability to move from a drinking water resource through the water plant purification process.

Industrial chemicals and pesticides do not share similar use and release patterns into the environment. Development of regulatory schema based on persistent, mobile, toxic (PMT) and very persistent, very mobile (vPvM) criteria should account for the differences between these chemical classes. Application of synthetic pesticides and herbicides to agricultural fields is a precise, highly technical, expensive and time-consuming process. The number of applications required is specific to the soil type and fertility, rainfall, erosion and weathering, field slope and runoff pathways, potency toward the intended pests, and crop type. US EPA and the European Union (EU) promulgate regulations toward minimization of pesticide use via programs of Integrated Pest Management (IPM). Several factors influence the environmental fate of pesticides including rates of abiotic or biotic degradation and dissipation, bioconcentration and sorption (mobility). In turn, these factors vary with sunlight intensity, pH of the water or soil, hydroxyl radical concentration, number and type of microbial organisms in contact with the pesticide, soil composition, and qualitative characteristics of the organic carbons present.

In summary, the German Environment Agency (UBA) is considering additional steps to protect the integrity of the drinking water supply. The chemicals of particular concern to the UBA as potential drinking water contaminants represent a subset of chemicals classified under REACH as
PMT or vPvM substances.  During routine water sampling, detection and quantitation of polar substances with high mobility can be problematic. The UBA is considering the use of the Log Koc value as a proxy for mobility (M). Log Koc is related to Log P by the equation Log Koc = 0.69 Log P + 0.22. In this study, we demonstrate that chemicals with log P values at or very close to 2.0, 3.0 or 4.0 (and their concomitant log Koc values) can vary significantly in their chemical structures, molecular weights, molar volumes, and calculated molar refractivity.

Table 2. ANOVA of log koc (from log Ps at or near 2, 3 or 4) with MgVol, CMR, and molecular weight.

| A | ANOVA: Single Factor—MgVol |
|---|-----------------------------|
| Summary | Count | Sum | Average | Variance |
| logKoc_1 | 25 | 69.41 | 2.7764 | 1.323574 |
| logKoc_2 | 25 | 67.59 | 2.7036 | 0.784957 |
| logKoc_3 | 25 | 71.11 | 2.8444 | 1.274767 |

| Source of Variation | SS | df | MS | F | P-value | F crit |
|---------------------|----|----|----|---|---------|--------|
| Between Groups      | 0.247904 | 2 | 0.123952 | 0.109909 | 0.896065 | 3.123907 |
| Within Groups       | 81.19917 | 72 | 1.127766 |     |         |        |
| Total               | 81.44707 | 74 |     |     |         |        |

| B | ANOVA: Single Factor—CMR |
|---|-----------------------------|
| Summary | Count | Sum | Average | Variance |
| logKoc_1 | 25 | 259.84 | 10.3936 | 14.87308 |
| logKoc_2 | 25 | 243.17 | 9.7268 | 10.42047 |
| logKoc_3 | 25 | 259.84 | 10.3936 | 14.87308 |

| Source of Variation | SS | df | MS | F | P-value | F crit |
|---------------------|----|----|----|---|---------|--------|
| Between Groups      | 7.410371 | 2 | 3.705185 | 0.276736 | 0.759057 | 3.123907 |
| Within Groups       | 963.9993 | 72 | 13.38888 |     |         |        |
| Total               | 971.4097 | 74 |     |     |         |        |

| C | ANOVA: Single Factor—Molecular Weight |
|---|--------------------------------------|
| Summary | Count | Sum | Average | Variance |
| logKoc_1 | 25 | 9154.596 | 366.1838 | 24529.7 |
| logKoc_2 | 25 | 8878.437 | 355.1375 | 16553.83 |
| logKoc_3 | 25 | 9364.402 | 374.5761 | 19869.91 |

| Source of Variation | SS | df | MS | F | P-value | F crit |
|---------------------|----|----|----|---|---------|--------|
| Between Groups      | 4752.591 | 2 | 2376.296 | 0.116956 | 0.889793 | 3.123907 |
| Within Groups       | 1462883 | 72 | 20317.81 |     |         |        |
| Total               | 1467635 | 74 |     |     |         |        |
(CMR), which is related to the mean polarizability of a molecule. The large degree of potential diversity in chemical structure and molecular parameters related to chemical behavior at a particular log P or log Koc value suggests that log Koc might not contain enough information to function as a standalone surrogate for the mobility (M) of a chemical, i.e. as related to its ability to move from a drinking water resource through the water plant purification process.\footnote{46}

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Carr J Smith https://orcid.org/0000-0002-8708-5208
Gene M Ko https://orcid.org/0000-0003-1672-9202

References
1. Rüdel H, Körner W, Letzel T, et al. Commentary: persistent, mobile and toxic substances in the environment: a spotlight on current research and regulatory activities. Environ Sci Eur 2020; 32: 5.
2. Neumann M, Schwarz MA, Sättler D, et al. A proposal for a chemical assessment concept for the protection of raw water resources under REACH. In: Extended Abstract for the Oral Presentation at the 25th Annual Meeting of the Society of Environmental Toxicology and Chemistry (SETAC Europe), Barcelona, Spain, 3–7 May 2015.
3. Neumann M. Proposal for criteria and an assessment concept for the identification of persistent, mobile and toxic (PMT) substances to protect raw water for the production of drinking water under the EU regulation REACH [in German]. Zbl Geol Paläont Teil I 2017; 1: 10–91.
4. Neumann M and Schliebner I. Protecting the sources of our drinking water from mobile chemicals—A revised proposal for implementing criteria and an assessment procedure to identify persistent, mobile and toxic (PMT) and very persistent, very mobile (vPvM) substances registered under REACH. Dessau-Rossau: German Environmental Agency, p. 20. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/171027_uba_pos_pmt_substances_engl_2auf_bf.pdf. (2017, accessed 19 December 2019).
5. Neumann M and Schliebner I. Protecting the sources of our drinking water: The criteria for identifying Persistent, Mobile, and Toxic (PMT) substances and very Persistent, and very Mobile (vPoM) substances under the EU chemical legislation REACH. UBA Texte 127/2019. ISSN: 1862-4804. Dessau-Rossau: German Environmental Agency (UBA), http://files.chemicalwatch.com/20190617_UBA_PMT_vPoM_criteria.pdf http://files.chemicalwatch.com/20190617_UBA_PMT_vPvM_criteria.pdf (2019, accessed 20 December 2019).
6. Reemtsma T, Berger U, Arp HPH, et al. Mind the gap: persistent and mobile organic compounds—water contaminants that slip through. Environ Sci Technol 2016; 50: 10308–10315.
7. United States Environmental Protection Agency (US EPA). Guidance for reporting on the environmental fate and transport of the stressors of concern in problem formulations, https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/guidance-reporting-environmental-fate-and-transport (2010, accessed 21 December 2019).
8. McCall PJ, Laskowski DA, Swann RL, et al. Measurement of sorption coefficients of organic chemicals and their use, in environmental fate analysis. In: Test protocols for environmental fate and movement of toxicants. proceedings of AOAC Symposium, AOAC: Washington, DC, 1981, pp. 94–109.
9. Leo A, Hansch C and Elkins D. Partition coefficients and their uses. Chem Rev 1971; 6: 525–616.
10. Piwoni MD and Kellely JW. Basic concepts of contaminant sorption at hazardous waste sites. EPA/540/4-90/053 October 1990 Groundwater Issue, pp. 1–7. 1990, https://www.epa.gov/sites/production/files/201506/documents/basic_concepts_sorption_haz_site.pdf (accessed 22 December 2019).
11. Smith CJ, Perfetti TA, Ko GM, et al. Ames mutagenicity, structural alerts of carcinogenicity, Hansch QSAR parameters (ClogP, CMR, MgVol), tumor site concordance/multiplicity, and tumorigenicity rank in NTP 2-year rodent studies. Toxicol Res Appl 2018; 2: 1–14.
12. Banerjee P, Erehman J, Gohlke BO, et al. Super natural II: a database of natural products. Nucleic Acids Res 2015; 43(D1): D935–D939.
13. BioByte Corp. Biobyte—Bio-Loom. Published 2016. Website: http://biobyte.com/bb/prod/bioloom.html (accessed 9 February 2017).
14. Hansch C and Leo A. Exploring QSAR: Volume 1: fundamentals and applications in chemistry and biology. In: Heller SR (ed) ACS reference book. 1st edn. Washington: American Chemical Society, 1995.
15. Leo A. Calculating log pot from structures. Chem Rev 1993; 93: 1281–1306.
16. Abraham MH. Scales of solute hydrogen-bonding – their construction and application to physicochemical and biochemical processes. Chem Soc Rev 1993; 22: 73–83.
17. Abraham MH and McGowan JC. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. Chromatographia 1987; 23: 243–246.
18. Hansch C, Leo A and Hoekman D. Volume 2: exploring qsar: hydrophobic, electronic, and steric constants. (ACS professional reference book). 1st edn. Washington: American Chemical Society, 1995.
19. Mannhold R, Poda GI, Ostermann C, et al. Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. J Pharm Sci 2009; 98: 861–893.
20. Arnot JA and Gobas FAPC. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environ Res 2006; 14: 257–297.
21. Devillers J, Domine D, Bintein S, et al. Comparison of fish bioconcentration models. In: Devillers J (ed.) Comparative QSAR. Washington: Taylor & Francis, 1998, pp. 1–50.
22. Garg R, Gupta SP, Gao H, et al. Quantitative structure–activity relationship studies on anti-HIV drugs. Chem Rev 1999; 99: 3525–3602.
23. Hansch C, Kim D, Leo A, et al. Toward a quantitative comparative toxicology of organic compounds. Crit Rev Toxicol 1989; 19: 185–226.
24. Leo A and Hansch C. Role of hydrophobic effects in mechanistic QSAR. Pers Drug Discov Des 1999; 17: 1–25.
25. Muller M and Nendza M. Literature study: comparative analysis of estimated and measured BCF data (OECD 305) with a special focus on differential accumulation of (mixtures of) stereoisomers, 2009. Dessau-Roßlau. Germany: Federal Environment Agency (Umweltbundesamt), http://www.uba.de/uba-info-medien-e/4088.html (2009, accessed 30 December 2019).
26. Selassie CD, Garg R and Mekapati SB. A mechanism-based approach to the study of the toxicity of endocrine disruptive agents. Pure Appl Chem 2003; 75: 2363–2374.
27. Smith CJ, Perfetti TA, Morton MJ, et al. The relative toxicity of substituted phenols reported in cigarette mainstream smoke. Tox Sci 2002; 69: 265–278.
28. Smith CJ, Perfetti TA, Garg R, et al. IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. Food Chem Toxicol 2003; 41: 807–817.
29. Smith CJ, Perfetti TA, Garg R, et al. Percutaneous penetration enhancers in cigarette mainstream smoke. Food Chem Toxicol 2004; 42: 9–15.
30. Smith CJ, Perfetti TA, Garg R, et al. Utility of the mouse dermal promotion assay in comparing the tumorigenic potential of cigarette mainstream smoke. Food Chem Toxicol 2006; 44: 1699–1706.
31. Fisher RA. Statistical methods for research workers, fifth edition, revised and enlarged. Edinburgh: Oliver and Boyd Ltd, 1934, pp. 1–354.
32. Garg R and Smith CJ. Predicting the bioconcentration factor of highly hydrophobic organic chemicals. Food Chem Toxicol 2014; 69: 252–259.
33. Lichtenberg E. Economics of pesticide use and regulation, reference module in earth systems and environmental sciences, encyclopedia of energy, natural resource, and environmental economics. 2013; 2013: 86–97.
34. Souza ML, Lovato MB, Fagundes M, et al. Soil fertility and rainfall during specific phenological phases affect seed trait variation in a widely distributed neotropical tree. Copaif Am J Bot 2019; 106(8): 1096–1105.
35. National Research Council (NRC). Soil and water quality: an agenda for agriculture. Washington: The National Academies Press, 1993.
36. US EPA. National pollutant discharge elimination system (npdes) permit writers’ manual for concentrated animal feeding operations (cafos), appendix a, basic soil science and soil fertility. https://www3.epa.gov/npdes/pubs/cafo_permitmanual_appendixa.pdf. (2013, accessed 24 April 2020).
37. Brady NC and Weil RR. The nature and properties of soils. 13th edn. New Jersey: Prentice- Hall, 2002, 960 pp.
38. Smith RL and Smith TM. Ecology and field biology. Upper Saddle River: Pearson Education, 1990, 720 pp.
39. U.S. Department of Agriculture (USDA). Natural resources conservation service (USDA-NRCS) 2011. National Soil Survey Handbook, title 430-VI, 2011, http://soils.usda.gov/technical/handbook/ (2011, accessed 11 April 2020).
40. Bateman RP. Rational pesticide use: spatially and temporally targeted application of specific products. In: Wilson M (ed.) Optimising pesticide use. Chichester: John Wiley & Sons Ltd, 2003, pp. 129–157.
41. US EPA. Label review manual, chapter 11: directions for use 11-1through 11-31. https://www.epa.gov/sites/production/files/2018-04/documents/lrm-complete-mar-2018.pdf (2014, accessed 10 April 2020).
42. European Parliament. Council of the European Union (EP-CEU), directive 2009/128/ec of the european parliament and of the council of 21 october 2009 establishing a framework for community action to achieve the sustainable use of pesticides, OJ L 309, 24.11.2009, pp. 71–86, https://eur-lex.europa.eu/legal-content/EN//ALL/?uri=celex%3A32009 L0128 (2009, accessed 11 April 2020).
43. FAO. Appendix 2. parameters of pesticides that influence processes in the soil. In: FAO Information Division Editorial Group (ed.) Pesticide Disposal Series 8. Assessing Soil Contamination. A Reference Manual. Rome: Food & Agriculture Organization of the United Nations (FAO), http://www.fao. org/3/X2570E/X2570E06.htm (2000, accessed 20 May 2020).
44. Schulze S, Paschke H, Meier T, et al. A rapid method for quantification of persistent and mobile organic substances in water using supercritical fluid chromatography coupled to high-resolution mass spectrometry. Anal Bioanal Chem 2020; 412: 4941–4952.
45. Reemtsma T, Weiss S, Mueller J, et al. Polar pollutants entry into the water cycle by municipal wastewater: a European perspective. Environ Sci Technol 2006; 40(17): 5451–5458.
46. Arnot JA and Mackay D. Policies for chemical hazard and risk priority setting: can persistence, bioaccumulation, toxicity, and quantity information be combined? Env Sci Technol 2008; 42(13): 4648–4654.