Fragment-based approach to calculate hydrophobicity of anionic and nonionic surfactants from chromatographic retention on a C_{18} stationary phase


DOI
10.1002/etc.3564

Publication date
2017

Document Version
Final published version

Published in
Environmental Toxicology and Chemistry

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care)

Link to publication


General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
FRAGMENT-BASED APPROACH TO CALCULATE HYDROPHOBICITY OF ANIONIC AND NONIONIC SURFACTANTS DERIVED FROM CHROMATOGRAPHIC RETENTION ON A C\textsubscript{18} STATIONARY PHASE

JORI HAMMER,\textasteriskcentered J.-H. HAFTKA,\textasteriskcentered PETER SCHERPENSIE,\textasteriskcentered JOOP L.M. HERMENS, and PIM W.P. DE VOOGT\textasteriskcentered

\textasteriskcentered Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands
\textasteriskcentered Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
\textasteriskcentered KWR Watercycle Research Institute, Nieuwegein, The Netherlands

(Submitted 21 March 2016; Returned for Revision 9 May 2016; Accepted 25 July 2016)

Abstract: To predict the fate and potential effects of organic contaminants, information about their hydrophobicity is required. However, common parameters to describe the hydrophobicity of organic compounds (e.g., octanol–water partition constant \(K\text{\textsubscript{OW}}\)) proved to be inadequate for ionic and nonionic surfactants because of their surface-active properties. As an alternative approach to determine their hydrophobicity, the aim of the present study was therefore to measure the retention of a wide range of surfactants on a C\textsubscript{18} stationary phase. Capacity factors in pure water \(k'\text{\textsubscript{OW}}\) increased linearly with increasing number of carbon atoms in the surfactant structure. Fragment contribution values were determined for each structural unit with multilinear regression, and the results were consistent with the expected influence of these fragments on the hydrophobicity of surfactants. Capacity factors of reference compounds and \(K\text{\textsubscript{OW}}\) values from the literature were used to estimate \(\log K\text{\textsubscript{OW}}\) values for surfactants (\(\log K\text{\textsubscript{OW}}\)). These \(\log K\text{\textsubscript{OW}}\) values were also compared to \(K\text{\textsubscript{OW}}\) values calculated with 4 computational programs: KOWWIN, Marvin calculator, SPARC, and COSMOTHERM\textsubscript{X}. In conclusion, capacity factors from a C\textsubscript{18} stationary phase are found to better reflect hydrophobicity of surfactants than their \(K\text{\textsubscript{OW}}\) values.

Keywords: Surface-active property Chromatography Octanol–water partition coefficient Capacity factor

INTRODUCTION

Surfactants are high-production volume chemicals with applications in industry and numerous consumer products. Prediction of the fate and potential effects of surfactants in the environment is needed for regulation (e.g., Registration, Evaluation, Authorisation and Restriction of Chemicals [REACH] legislation) and requires information about the hydrophobicity of these chemicals. Hydrophobicity of organic compounds is often used in predictive models. The term “hydrophobicity” refers to the tendency of a compound to “escape” the aqueous phase and is illustrated by taking air–water partitioning as an example. In air–water partitioning, only the interactions in the water phase have to be considered, assuming air is regarded as an ideal gas phase without any interactions [1]. Partitioning between water and other phases is then influenced by hydrophobicity plus specific interactions between the (hydrophobic) compound and this other phase (e.g., hydrogen bonding or electrostatic interactions). Hydrophobicity depends on the energy needed for cavity formation in the aqueous phase, electrostatic interactions, van der Waals interactions, and hydrogen bonding interactions [1]. Cavity formation is defined by the energy required to form a cavity in the bulk water phase to accommodate the hydrophobic compound and is thereby dependent on the molecular size of the compound and the intermolecular forces between water–water and/or water–compound molecules [1]. Electrostatic interactions occur when the organic compound and the aqueous phase both contain charged sites that can attract each other. Van der Waals interactions are interactions between any kind of molecules in the form of dispersive interactions (induced dipole–induced dipole), Debye interactions (dipole–induced dipole), and Keesom interactions (dipole–dipole). The polarity of water itself further allows for the formation of hydrogen bonds between specific hydrogen bond donor and acceptor sites of the hydrophobic compound (and between water molecules). Electrostatic interactions can also be responsible for specific sorption to other phases (i.e., calcium bridging between anionic surfactants and negatively charged sorbents) and play an important role in the partition process of anionic surfactants. Note that sorption of ionized surfactants to (environmental) sorbent phases also depends on pH and ionic strength.

The \(n\)-octanol–water partition constant \(K\text{\textsubscript{OW}}\) is often used to describe the hydrophobicity of organic compounds and tends to be a good predictor of the sorption and fate of most neutral polar and nonpolar organic compounds [2]. The \(K\text{\textsubscript{OW}}\) value is usually determined using the slow-stirring (Organisation for Economic Co-operation and Development Test No. 123) [3] or shake flask (Organisation for Economic Co-operation and Development Test No. 107) method [4]. However, the determination of \(K\text{\textsubscript{OW}}\) values of (ionic) surfactants is technically very difficult because of their specific properties. The surface-active properties of surfactants will cause them to accumulate at the octanol–water interface and cause emulsification of both phases. The concentrations of surfactants in water and octanol measured in the slow-stirring or shake flask method are therefore not considered representative of the actual distribution.
of surfactants in the octanol–water system. Hence, alternative parameters to obtain a measure of the hydrophobicity of (ionized) surfactants in water are urgently required. Yet, currently, there is no robust approach or tool that offers reliable results for prediction of the environmental fate and behavior of surfactants. In a previous study, we presented sorption mechanisms of nonionic and anionic surfactants to polyacrylate-coated solid-phase microextraction fibers [5]. Alternatively, the present study aimed to measure a parameter that reflects the hydrophobicity of nonionic and anionic surfactants which can subsequently be applied to predict sorption, bioaccumulation, and toxicity of these compounds.

In Organisation for Economic Co-operation and Development method 117, capacity factors ($k'$) of analytes and reference compounds are measured with reversed-phase high-performance liquid chromatography (HPLC) using a C18 stationary phase and a mobile phase with a constant methanol–water composition (methanol fraction, $\varphi = 0.9$) to predict $K_{OW}$ values of "difficult" substances [6]. However, determining $K_{OW}$ values with Organisation for Economic Co-operation and Development method 117 is not suitable for surfactants because there are insufficient $K_{OW}$ values available for reference molecules (specifically for surfactants) to enable a valid calibration. Therefore, in the present study capacity factors are used to directly represent a measure of the hydrophobicity of surfactants. A linear log $k'$ versus $\varphi$ relationship is observed for many compounds when methanol is used as the organic modifier [7,8], even for surfactants [9], and allows for extrapolation of the capacity factor to 0% organic modifier (pure water). Extrapolation to a mobile phase which only contains water minimizes the effect of organic solvent on the nature of the aqueous phase. This approach has also been used in previous studies [7–10]. To study the influence of the hydrophobic carbon chain of surfactants, we determined the capacity factor in pure water for several groups of neutral and anionic surfactants with increasing carbon chain lengths. Until now, determining the relationship between log $K_{OW}$ and HPLC retention indices on a C18 column has only been performed for neutral nonpolar molecules [6,7,11–15].

**MATERIALS AND METHODS**

**Theory**

A stationary phase composed of C18 chains is relatively inert and should not contain any specific sorption sites. This would mean that the stationary phase ideally contains no free silanol groups. The only direct attraction between the test compounds and the C18 phase occurs in the form of van der Waals interactions. For a given molecule, these interactions are of similar strength for both the mobile and the stationary phases and cannot be responsible for most of the retention on the C18 phase [1]. Retention on a C18 phase in the presence of pure water as mobile phase relates to the interactions of the compound with water because of its hydrophobicity rather than specific interactions with the C18 phase. To measure the retention of test analytes on a C18 phase, logarithmic capacity factors (log $k'$) were calculated from the retention time of the target compound ($t_r$) and that of an unretained compound ($t_0$)

$$\log k' = \log \frac{t_r - t_0}{t_0}$$  

(1)

Because the retention of a hydrophobic compound on a C18 phase in pure water as eluent can increase up to several hours, an organic modifier is added to the mobile phase to promote elution of the compounds. When methanol is used as an organic modifier, the relationship between the logarithmic capacity factor and the methanol fraction in the mobile phase is linear for most neutral solutes [9,16,17] and can be expressed as

$$\log k' = a \times \varphi + \log k'_0$$  

(2)

where the slope $a$ represents the change in the logarithmic capacity factor of the target compound as a function of the methanol fraction ($\varphi$) and the intercept ($\log k'_0$) represents the logarithmic capacity factor at $\varphi_0$ (i.e., in pure water). A sufficient number of measurements have to be made at different modifier strengths before extrapolation to pure water can be applied. Differences in calculated log $k'_0$ values of the test compounds can be the result of either different alkyl chain lengths or changes in the type of head group. A multilinear regression analysis was subsequently applied to calculate the fragmental contribution of each structural unit to the total log $k'_0$ value of respective compounds within a class of surfactants

$$\log k'_0 = \Sigma (fn) + b$$  

(3)

where $\Sigma(fn)$ is the summation of the contribution of each fragment (f) multiplied by the number of fragments occurring in the molecular structure (n) and $b$ is the intercept. Fragments in the present study consisted of the following structural units: hydrocarbon (CH2), aromatic hydrocarbon (aromatic CH), fluorocarbon (CF2), aromatic hydroxyl (aromatic OH), aliphatic hydroxyl (aliphatic OH), methyl ester (COOC), sulfonate (SO3–), sulfite (SO42–), and carboxylate in hydrocarbon (CO2) or perfluorinated carbon structures (CO2F). For alcohol ethoxylates, $n$ refers to the number of ethoxylate units, (OC2H)n.

**Chemicals**

Linear alkyl benzenes (with alkyl chain lengths C2, C6, and C12), linear alkyl phenols (C9, C8, and C6), sodium salts of linear alkyl carboxylates (C11, C13, C14, C15, C17, and C19; dissociation constant $pK_a \approx 4.86$) [18], polycyclic aromatic hydrocarbons (PAHs), linear alkyl methyl esters (C9, C11, and C13), alcohol ethoxylates (C10, C12, C13, C14, and C16; with ethoxylate units ranging between 4 and 8 units), linear perfluoroalkyl carboxylic acids (C4–C13 and C15; $pK_a \approx -0.5$) [19], and uracil were obtained from Sigma-Aldrich. The compound C11-octaehtylene glycol was synthesized by Roy Sheppard (Unilever, Port Sunlight Research Laboratory, Port Sunlight, United Kingdom) [20]. The compound C5-phenol was obtained from Fisher Scientific. Sodium salts of linear alklyl sulfonates (C11, C13–C16, and C18; $pK_a \approx -1.9$) [21] and linear alkyl sulfates (C11, C13, C15–C18; $pK_a \approx 1.9$) [22] were obtained from Research Plus. Sodium salts of p-alkyl benzene sulfonates (C2–y-benzene sulfonates with C5, C10, C12, and C13 and $y = 1, 2, or 5$, where $y$ indicates the position of the benzene sulfonate group on the alkyl chain; $pK_a \approx 2.5$) [23] were synthesized by J. Tolls [24]. All organic compounds had purities higher than 98%, except for C10-tetraethylen glycol and C12-2-benzene sulphonate (both $\geq 97\%$). Each compound group consisted of 3 to 8 substances varying in carbon chain length. The chain lengths and numbers/positions of specific functional groups of individual compounds and their Chemical Abstracts Service numbers are given in Supplemental Data, Table S1. Ammonium acetate and sodium acetate were purchased from Sigma-Aldrich. Methanol was obtained from Biosolve. Ultrapure water was obtained from a
Millipore water purification system (resistivity >18 MΩ/cm; Merck Chemicals).

**HPLC method**

Retention times of the analytes were measured at methanol–water compositions ranging from 40:60 to 90:10 (v:v) with a Shimadzu Prominence HPLC system, consisting of 2 LC-20AD pumps, an SIL-20A autosampler, a CTO-20A oven, and a CBM-20A controller. A GraceSmart C18 column (150 × 2.1 mm, 5 μm particle size; Grace Discovery Sciences) was used at a temperature of 40 °C with a flow rate of 0.2 mL/min. Run times for HPLC varied between 5 min and 90 min. Prior to each run, the column was equilibrated with the mobile phase for at least 8 min. Mixtures with a maximum of 6 surfactants with similar head groups but different carbon chain lengths were prepared in a solvent mixture equivalent to the mobile phase composition of the respective run. The mixtures were injected in volumes between 10 μL and 40 μL. Analyte concentrations ranged between 0.1 mg/L and 2 mg/L, depending on the compound sensitivity of the detector. No precipitates were visually observed between 0.1 mg/L and 2 mg/L, depending on the compound fraction used during the analysis of the compound mixtures. The mobile phase consisted of methanol and Milli-Q water, with both containing 5 mM ammonium acetate to promote peak shape for all anionic compounds. In addition, ammonium acetate enabled the formation of ammonium adducts for alcohol ethoxylates in the mass spectrometer. Uricil was used to determine the column dead time.

The compound C12-5-benzene sulfonate was used as an internal standard to correct for variation in k' as a result of small changes in chromographic conditions, for example, by small changes in the methanol fraction caused by variability in the volume delivered by the HPLC pumps. The internal standard was added to all compound mixtures and analyzed in every HPLC run. Capacity factors of C12-5-benzene sulfonate were separately determined at exact methanol fractions by carefully weighing methanol and water volumes. In this way, methanol fractions used during the analysis of the compound mixtures were corrected with a calibration curve of the measured capacity factor of C12-5-benzene sulfonate in the same mixture.

**Detection**

Alkyl phenols, PAHs, and nonionic and anionic surfactants were detected with a triple-quadrupole mass spectrometer (SCIEX 4000 QTRAP MS/MS system; Applied Biosystems) with a turbo ion spray source operated at 250 °C to 500 °C. Alkyl benzenes and alkyl methyl esters were detected with fluorescence (Shimadzu; RF-20AXS) and ultraviolet (UV; Shimadzu; SPD-20AV) detectors, respectively. Mass spectrometer, fluorescence, and UV detector settings of the analytes are shown in Supplemental Data, Table S1. Chromatograms were integrated with Analyst 1.5.1 (Applied Biosystems) or LC solutions 1.25 (Shimadzu).

**RESULTS AND DISCUSSION**

**Sorption mechanisms of anionic surfactants**

First, to gain a better understanding of the retention mechanisms of anionic surfactants on a C18 stationary phase, the effect of pH and ionic strength on the determined k' values was investigated. Although a C18 stationary phase is relatively inert, several studies on a variety of C18 HPLC columns have indicated the presence of positively or negatively charged sites that can affect the retention of charged analytes [25–28]. The effect of negatively charged sites on the stationary phase is eliminated by changing the pH of the mobile phase to below the pK_a value of these sites [29]. However, this also affects the degree of ionization of some of the anionic surfactants. Alkyl carboxylates have a relatively high pK_a value (~4.86 [18]) compared with the other analytes. Consequently, log k' values of C13-carboxylate reach a maximum at a pH value of approximately 4.2 and below (Figure 1A). The subsequent pK_a value of C13-carboxylate (pK_a ≈ 5.2) derived from Figure 1A appeared to be somewhat higher than the pK_a value of alkyl carboxylates, which may have resulted from the presence of methanol in the mobile phase [30]. The other surfactant groups plotted in Figure 1A all show similar trends, with a steady increase of log k' values at pH values lower than 6 and sharply increasing log k' values at pH 3 and below. The similarity of the curves indicates that this likely results from changes in the silica surface of the stationary phase. The silica surface of a C18 stationary phase can contain several forms of acidic silanol groups that affect elution of anionic compounds by ion exclusion [31,32]. When these groups are protonated at low pH, the ion-exclusion effect disappears and the silica surface allows for more interaction of the anionic compound with the stationary phase, thereby increasing retention [33]. Another less common explanation is the protonation of basic groups present in the stationary phase. Méndez et al. [34] reported anion-exchange properties of several C18 columns when measuring at acidic pH values (pH 3.4) and related this observation to impurities originating from the production process of the silica core. Although the column used in the present study contains an end-capped C18 phase, free
Effect of molecular structure on retention of surfactants. It is therefore difficult to determine the role of charged silanol groups on the elution behavior of these anionic compounds.

Another way to suppress the effect of electrostatic interactions is to increase the ionic strength of the mobile phase [35,36]. Increasing ionic strength would be a better approach than decreasing the pH because the anionic surfactants remain in their ionic state, which is relevant when studying their interaction with the aqueous phase. Ammonium acetate was used to examine the effect of ionic strength on the elution of anionic surfactants (Figure 1B). Without ammonium acetate, retention of anionic surfactants decreased to approximately 0 mM to 40 mM, retention of anionic surfactants increased because negatively charged sites on the stationary phase were probably suppressed and the ion-exclusion mechanism became less prominent. Retention of the compounds increased until approximately 20 mM ammonium acetate, and retention remained constant up to 50 mM ammonium acetate (not shown in Figure 1B). Yet, to prevent salts from interfering with the ionization in the mass spectrometric system, log \( k' \) values for all other measurements were determined at a concentration of 5 mM ammonium acetate. The differences in \( k'_0 \) values at 5 mM ammonium acetate and maximum \( k'_0 \) values (>20 mM ammonium acetate) are small (Supplemental Data, Figure S1) and can be corrected by using a correction factor of 1.028 (not applied to the data in the present study).

Effect of modifier strength on determination of \( k' \)

For all compounds in the present study, capacity factors (\( k' \)) were determined at a range of methanol fractions (\( \varphi \)) between 0.40 and 0.85, as long as elution times did not exceed 90 min. The relationship of \( \log k' \) versus \( \varphi \) was linear for most of the tested methanol fractions (Supplemental Data, Figure S2). Within a surfactant group, the slope of the \( \log k' \) versus \( \varphi \) relationship increased slightly (negatively) with increasing carbon chain length. For each surfactant group, the convergence point of these plots was well above 90% of methanol. This resulted in a constant order of elution of the mixtures at most methanol fractions, which allowed for extrapolation of the regression lines to lower methanol fractions [37]. However, overall slope values differed between surfactant groups (Supplemental Data, Figure S3). At high fractions of methanol, nonlinearity was observed for most compounds as elution times approached \( t_0 \) and the error in the calculation of \( k' \) increased. Surfactants with carbon chain lengths above 10 carbon atoms were often found to be linear up to \( \varphi = 0.88 \), whereas chain lengths shorter than 10 carbon atoms were only linear at much lower methanol fractions (maximum \( \varphi = 0.50 \) for perfluoro-C12-carboxylate). Data points that were used in the extrapolation of capacity factors to obtain \( \log k'_0 \) in pure water were always chosen from the linear part of the \( \log k' \) versus \( \varphi \) relationship and when \( t_r \) was at least 50% higher than \( t_0 \).

Effect of molecular structure on retention of surfactants

The linear \( \log k' \) versus \( \varphi \) relationship was used to calculate capacity factors in pure water. A linear relationship between \( \log k'_0 \) and the number of carbon atoms in the alkyl chain was observed for all test compounds (Figure 2 and Supplemental Data, Figure S4). The slope in Figure 2 indicates a change in the behavior of the test compound in water after addition of a hydrocarbon or a fluorocarbon group. A relatively steep slope refers to a strong increase in affinity to the C18 stationary phase resulting from increasing chain lengths of the hydrophobic carbon chain with associated increases in the energy required for cavity formation. However, the head group structure also slightly influences the slope of a surfactant group. A distinction can be made between certain compound groups based on the slope of the relationship of \( \log k'_0 \) versus the number of carbon atoms. All nonionic compounds have higher slopes (0.50–0.56 log unit per carbon atom) compared to those of ionic compounds (0.39–0.46 log unit per carbon atom), except for perfluoroalkyl carboxylates (0.56 log unit per carbon atom; Supplemental Data, Table S2). Also, nonionic compounds have a higher capacity factor than anionic surfactants when compared at the same carbon chain length (Supplemental Data, Table S2 and Figure S4). The nonionic alcohol ethoxylates generally showed higher \( \log k'_0 \) values than those of anionic surfactants with similar carbon chain lengths (except for perfluoroalkyl carboxylates).

Eadsforth et al. [38] measured capacity factors of alcohol ethoxylates on a C18 stationary phase at \( \varphi = 0.9 \) and found a similar trend when studying the effect of carbon chain length on the retention of these compounds. However, an increase in the number of ethoxylate units (OCCy) resulted in a very slight decrease of the capacity factor (approximately –0.01 log unit per ethoxylate unit), whereas in the present study a stronger increase in capacity factor with increasing number of ethoxylate units was observed (~0.09 log unit per ethoxylate unit) [38]. This deviation results from a different experimental setup in the present study compared with that of Eadsforth et al. [38], where capacity factors were measured at a constant methanol fraction (\( \varphi = 0.9 \)). When studying the \( \log k' \) versus \( \varphi \) relationship for 2 C12-ethylene glycols, our data also showed an increase in the capacity factor per ethoxylate unit at \( \varphi = 0.9 \); but a convergence point was observed at \( \varphi = 0.88 \), resulting in reversal of the elution order when extrapolated to \( \varphi = 0 \) (Supplemental Data, Figure S5).

Significantly higher \( \log k'_0 \) values were found for perfluoroalkyl carboxylates compared with their hydrocarbon equivalents. This is the effect of 2 mechanisms: 1) the fluorine atoms in CFx moieties are larger than the hydrogen atoms in CHx moieties and thereby increase the cavity size and energy required for perfluoroalkyl carboxylates to dissolve in water.

![Figure 2. Logarithmic capacity factors in pure water (log \( k'_0 \)) versus the number of carbon atoms in the alkyl chain (#C atoms [x]) for different anionic surfactant groups: perfluoroalkyl carboxylates (PFC,CO2–), primary and secondary alkyl benzene sulfonates (C7-1-BzSO3–, C7-2-BzSO3–), alkyl carboxylates (C6,CO2–), alkyl sulfates (C6SO4–), and alkyl sulfonates (C6SO3–).

compared with that of alkyl carboxylates [39], and 2) the highly electronegative fluorine atoms in the perfluoroalkyl chain maintain an electron-withdrawing effect and tend to pull the negative charge away from the carboxylic head group, making it less hydrated and therefore reducing affinity with the aqueous phase [40]. De Voogt et al. [9] found similar results for perfluoroalkyl carboxylates, but log \( K_0 \) values increased more per carbon atom (0.66 vs 0.57 log unit per carbon atom in the present study; Supplemental Data, Table S2), resulting in lower log \( K' \) values for perfluoroalkyl carboxylates with shorter chain lengths and higher log \( K' \) values for longer chain lengths when compared with the present findings (Supplemental Data, Table S3).

The increased log \( K_0 \) values of alkyl benzene sulfonates compared with those of alkyl sulfonates are explained by the presence of a benzene group (Figure 2). A change in the position of the benzene sulfonate head group on the alkyl chain affects the capacity factor only slightly. When the benzene sulfonate head group shifts from the end of the carbon chain (\( y = 1 \)) to another position (\( y > 1 \)), the hydrophobic alkyl chain becomes branched. Branching of alkyl benzene sulfonates is known to have an effect on their properties (i.e., on solubility, surface tension, and critical micelle concentration values) [41,42]. Likewise, the capacity factor is slightly decreased as the position of the head group is changed from the end toward the middle of the carbon chain. For example, less retention was observed for C\(_{12}\)-5-benzene sulfonate (internal standard) compared to C\(_{12}\)-2-benzene sulfonate (Supplemental Data, Table S3). Irrespective of the head group (i.e., carboxylate, sulfate, or sulfonate), all remaining anionic surfactants with aliphatic carbon chains showed similar log \( K_0 \) values for equal carbon chain length; and this phenomenon may be caused by their similar monovalent negative charge. The slope of the relationship of log \( K_0 \) versus the number of carbon atoms for PAHs (0.20 log unit per carbon atom) was much lower than that of other compounds because of the presence of aromatic carbon atoms. Because of the aromatic carbon structure, PAHs contains fewer hydrogen atoms than an aliphatic carbon structure and therefore have a smaller molecular volume [43].

**Contribution of structural fragments to retention**

Because all plots of log \( K_0 \) values versus number of carbon atoms yield linear relationships, the fragmental contribution of each structural unit to the total log \( K_0 \) value was calculated for all surfactants and expressed in fragment values (i.e., \( f[n] \) in Theory, above). The prediction of capacity factors from these fragment values is divided into 3 different contributions: 1) length and structure of the alkyl chain, 2) structure of the head group, and 3) remaining structural variables like number of ethoxylate units on the alkyl chain. Structural fragments and their fragmental contribution are shown in Table 1.

Fragments of the hydrocarbon or perfluorocarbon alkyl chain are presented as CH\(_2\) or CF\(_2\) groups. However, we were unable to make a distinction between CH\(_2\) and CH\(_3\) or CF\(_2\) and CF\(_3\) fragments because both fragments were always mutually present. As expected, the highest \( f \) values were calculated for CH\(_2\) (0.47 log unit) and CF\(_2\) fragments (0.56 log unit) because these are known to contribute to the hydrophobicity of neutral nonpolar compounds [39,43]. These 2 fragments have the least interaction with the aqueous phase and therefore only add to the energy needed to create a cavity for solvation. Fluorocarbon units had a slightly higher \( f \) value than hydrocarbon units because of the difference in geometry (size) between the fluorine and hydrogen atoms [39].

Aromatic hydrocarbon fragments (aromatic CH) showed a much lower \( f \) value (0.21 log unit) than aliphatic hydrocarbons because of the more compact structure (fewer hydrogen atoms) and therefore smaller total cavity in the aqueous phase [43]. Both aromatic and aliphatic OH fragments contribute negatively to log \( K_0 \) values (−0.93 and −0.97 log unit, respectively), which is explained by the hydrogen bond donor/acceptor properties of the hydroxyl moiety. Hydrogen bonding interactions are also present in the ethoxylate fragments, but these interactions are probably countered by the 2 carbon atoms in this structural fragment, resulting in a slightly positive \( f \) value (0.10 log unit). However, the hydrogen bonding interactions of the double-bonded and single-bonded oxygen atoms in the methyl ester fragment (COOC) result in a negative contribution to log \( K_0 \) (−0.50 log unit).

In Figure 2, the log \( K_0 \) values of several anionic surfactant groups (alkyl sulfates, alkyl carboxylates, and alkyl sulfonates) show hardly any difference at equal chain lengths. As a result, the \( f \) values of the anionic head groups SO\(_4\), SO\(_3\), and CO\(_2\) (\( H \)) are also highly comparable (−1.94, −2.01, and −2.02 log units, respectively). The relatively high negative \( f \) values correspond to the affinity of these fragments for the water phase and are the result either of hydrogen bonding interactions or of electrostatic interactions with water. However, it is difficult to determine the exact extent of each interaction. First, the number of hydrogen bonding sites (oxygen atoms) does not correlate with the \( f \) value of these structural fragments, but this can also be an effect of the positioning of the oxygen atoms in each head group. Also, there is no relation between the electrostatic charge and the calculated \( f \) value when comparing the pK\(_{OW} \) values of these compounds. The distinction between hydrocarbon-attached and fluorocarbon-attached carboxylic acid—CO\(_2\) (\( H \)) and CO\(_2\) (\( F \)), respectively—can be explained by the electron-withdrawing effect of the perfluorinated carbon chain on the carboxylic head group [40]. This effect causes a lower net negative charge of the CO\(_2\) (\( F \)) moiety, which results in a less hydrated carboxylic acid, subsequently weakening the affinity of the compounds with the aqueous phase [40,44]. The corresponding \( f \) value is therefore more positive for the head group of perfluoroalkyl carboxylates compared to that of alkyl carboxylates: −0.65 log unit for CO\(_2\) (\( F \)) versus −2.02 log units for CO\(_2\) (\( H \)). Overall, when comparing all fragments and their calculated \( f \) values, the results are consistent with the possible interactions the fragments can have with the aqueous phase (such as hydrogen bonding, cavity formation). In summary, the calculated \( f \) values from the present study correctly correspond with the expected interactions between the structural fragments and the aqueous phase (such as hydrogen bonding, cavity formation).

**Comparison of fragment values**

The programs KOWWIN from EPI Suite [45] and Marvin calculator from ChemAxon [46] use fragment values to predict log \( K_{OW} \) values. We compared our fragment values derived from the measured capacity factors with fragment values used for calculation of log \( K_{OW} \) values by these programs (Supplemental Data, Table S4). An exact quantitative comparison cannot be made, however, because the affinity of organic compounds for octanol (\( K_{OW} \) models) or a C\(_{18}\) stationary phase (present study) is based on different interactions. Because octanol contains hydrogen bond donor and accepting properties, it is expected that, for molecular fragments which can form hydrogen bonds, fragment values of KOWWIN and Marvin calculator will deviate from calculated fragment values from.
the present study. However, for fragments lacking these interactions (e.g., CH₂ and aromatic CH in Supplemental Data, Table S4) the octanol and C₁₈ phases are considered similar organic phases, and fragment values should differ less. The changes in hydrophobicity as a function of the number of ethoxylate units are relatively small according to fragment values from the present study ($f = +0.10$ log unit) and those from Marvin calculator ($f = -0.04$ log unit). An exception is the fragment value of the number of ethoxylate units calculated by KOWWIN ($f = -0.27$ log unit), implying significantly less affinity of ethoxylate units for octanol than for the C₁₈ phase. For the anionic head groups CO₂⁻ (F), SO₄⁻, and SO₃⁻, fragment values calculated with Marvin calculator show a similar trend with $f$ values calculated from HPLC capacity factors. The KOWWIN program shows the same trend for SO₄⁻ and SO₃⁻ but a much lower aqueous affinity for CO₂⁻ (F). The 2 programs do not distinguish between hydrocarbon-attached and fluorocarbon-attached carboxylic acids. However, KOWWIN uses a correction factor for the number of CF₂ units (−0.3 log units per unit) in a perfluoroalkyl chain. The justification of this correction factor is unclear because no reference could be obtained [47]. Another notable difference with the fragment values calculated in the present study is the relatively low fragment value of the COOC fragment calculated by KOWWIN. Although it would be expected that the presence of hydrogen bonds in the octanol phase would increase partitioning of an alkyl methyl ester to octanol compared to retention on a C₁₈ phase, the opposite is observed. Marvin calculator, on the other hand, shows much lower affinity, comparable to those of ethoxylate fragments. Most differences in fragment values are explained by the different interactions that are at play in these 2 systems (Supplemental Data, Table S4). However, regarding the anionic head groups, it must be noted that KOWWIN cannot calculate octanol–water distribution constants (log $D_{OW}$) of ionic compounds. For fragments like CO₂⁻ (F), the program uses protonated carboxylate fragments instead. In this way, the prediction method ignores the greater portion of interactions that define dissolution of ionic compounds into octanol. Marvin calculator overcomes this by using the $pK_a$ value to calculate log $D_{OW}$ and selects different fragment values depending on their protonation state. Although this seems to be a better approach, errors in $pK_a$ calculations can still cause deviations in predictions of log $D_{OW}$ [48]. Probably the largest concern with both programs is that the greater part of the fragment values used in their calculations originate from the same database [49]. For most (anionic) surfactants, there are simply no experimental data available, and this lack of surfactant data in databases causes uncertainties in the prediction of log $K_{OW}$ values.

**Comparison of log $K_{OW}$ values**

Besides the previously discussed programs, Marvin calculator and KOWWIN (which use relatively simple quantitative structure–activity relationship [QSAR] methods), there are other programs, such as SPARC from ARChem [50] and COSMO ThermX from COSMOlogic [51] which use a more mechanistic approach to predict log $K_{OW}$ values. The SPARC program relies on a combination of linear free energy theory for thermodynamic properties and perturbation molecular orbital methods for quantum mechanical contributions [52]. The program uses molecular increments just like KOWWIN and Marvin calculator but uses intermolecular interaction properties to define partition constants. Therefore, SPARC is more robust because the increments can be calibrated for any partition data and are not bound to a specific partition process [52]. However, SPARC still has an application domain that is limited by its calibration with experimental data. On the other hand, COSMO ThermX uses quantum mechanical calculations and statistical thermodynamics in a solvation model to describe interactions between solvent and solute. Its approach uses much fewer empirical parameters, and the output is therefore not limited to any kind of calibration [52,53]. The module COSMOConfX (Ver 3.0; COSMOlogic) can account for different molecular conformations and gives COSMO ThermX a more robust approach compared with the other programs. However, this specific feature also significantly increases calculation time. We used KOWWIN, Marvin calculator, SPARC, and COSMO ThermX to calculate log $K_{OW}$ values and compared these with log $K_{OW}$ values obtained from regression of capacity factors in pure water with literature log $K_{OW}$ values for reference compounds.

**Capacity factors of reference compounds with known experimental log $K_{OW}$ values** (alkyl benzenes, alkyl phenols, and alkyl methyl esters) were used to calculate log $K_{HPLC}^{OW}$ values for all test compounds (Supplemental Data, Figure S6). The experimental log $K_{OW}$ data used for the reference compounds were obtained from the EpiSuite database, and it must be noted that some of these values were obtained by an HPLC method similar to that in the present study [54]. Because the reference compounds are nonionic and much less polar than the tested surfactants in the present study and other data are not available, we assumed that this was the most reliable approach to predict log $K_{OW}$ values for our test compounds. The log $K_{HPLC}^{OW}$ values are shown in Supplemental Data, Table S3, together with calculated log $K_{OW}$ values from the 4 computational programs. The log $K_{HPLC}^{OW}$ values were plotted against calculated log $K_{OW}$ for all programs in Figure 3 and in Supplemental Data, Figure S7. All estimated log $K_{OW}$ values were calculated for the neutral form of the substance because log $K_{OW}$ calculations of ionic species resulted in very unrealistic values. When comparing calculated values with log $K_{HPLC}^{OW}$ values, KOWWIN returns the lowest log $K_{OW}$ values and shows a significant increase in log $K_{OW}$ values with increasing number of ethoxylate units, whereas the opposite is seen with log $K_{HPLC}^{OW}$ values for alcohol ethoxylates. Marvin calculator generally produces log $K_{OW}$ values that correspond better with log $K_{HPLC}^{OW}$ than log $K_{OW}$ values from KOWWIN. For all hydrocarbon surfactants, log $K_{OW}$ values calculated with SPARC show the best fit with log $K_{HPLC}^{OW}$ values. However, SPARC greatly
Hydrophobicity of surfactants determined from HPLC retention

Figure 3. Logarithmic octanol–water partition constants (log \( K_{OW} \)) calculated with COSMOThermX versus logarithmic \( K_{OW} \) values derived from capacity factors (log \( K_{HPLC}^{OW} \)) measured in the present study for different surfactant groups: perfluorooalkyl carboxylates (PFC\(_x\)CO\(_y\)), alkyl benzene sulfonates (C\(_x\)y-BzSO\(_3^-\)), alcohol ethoxylates (C\(_x\)(OCC)\(_y\)OH), alkyl carboxylates (C\(_x\)CO\(_2^-\)), alkyl sulfonates (C\(_x\)SO\(_3^-\)), and alkyl sulfates (C\(_x\)SO\(_4^-\)). Dotted line represents a 1:1 line. HPLC = high-performance liquid chromatography.

overpredicts log \( K_{OW} \) values for perfluorinated surfactants. For these compounds, the simpler QSAR programs like Marvin calculator and KOWWIN show a better fit. The program COSMOThermX slightly overpredicts log \( K_{OW} \) values for all surfactants, especially for alkyl sulfates, but gives the best fit for perfluorooalkyl carboxylates. Overall, SPARC shows the best relationship with log \( K_{HPLC}^{OW} \) values for surfactants with a hydrocarbon tail, whereas perfluorinated surfactants show the best fit with values calculated with COSMOThermX. It is important to note that the calculated log \( K_{OW} \) and derived log \( K_{HPLC}^{OW} \) values remain arbitrary because no experimental reference values are present, especially for anionic surfactants.

CONCLUSIONS

In the present study, an alternative approach is presented to determine the hydrophobicity of nonionic and anionic surfactants. The capacity factors measured on a C\(_{18}\) stationary phase showed a linear relationship with organic modifier strength in the mobile phase and allowed for the extrapolation and determination of capacity factors in pure water. Changes in pH and ionic strength (i.e., variation of ammonium acetate concentration) affected the capacity factor of anionic surfactants, possibly by screening of surface charges on the silica base of the C\(_{18}\) column as well as by protonation of test compounds (e.g., alkyl carboxylates). The calculated contributions of structural fragments gave more insight into the interactions that influence the hydrophobicity of surfactants. Hydrocarbon and perfluorocarbon fragments (CH\(_2\) and CF\(_2\) units) had the largest positive fragment values and thereby contribute most to the hydrophobicity by increasing the energy needed for cavity formation in the water phase. Anionic head groups (CO\(_2^-\), SO\(_4^-\), and SO\(_3^-\)) had the largest negative fragment values and therefore contribute more to the hydrophilic character of the surfactant (e.g., by hydrogen bonding and/or electrostatic interactions with the water phase). Furthermore, octanol–water partition constants of surfactants that were derived from capacity factors of reference compounds (log \( K_{HPLC}^{OW} \)) show the best agreement with calculated log \( K_{OW} \) values from \( K_{HPLC}^{OW} \) COSMOThermX.

Although capacity factors on a reversed-phase C\(_{18}\) column are representative of the hydrophobicity of surfactants, it is not only hydrophobicity that determines the partitioning of surfactants between water and environmental matrices. Hydrophobicity combined with specific interactions (i.e., hydrogen bonding, electrostatic interactions, etc.) between surfactants and other phases will eventually determine their fate. The actual importance of log \( K_{OW} \) for surfactants therefore remains ambiguous, and one can argue whether predicting and comparing log \( K_{OW} \) values for anionic surfactants has any meaning. In that sense, retention on a C\(_{18}\) column better reflects hydrophobicity because any specific interactions do not occur with the C\(_{18}\) phase. Fragment values from the present study correctly correspond with the interactions that are expected to play a role in determining the hydrophobicity of nonionic and anionic surfactants and from a theoretical and practical viewpoint seem to be a better approach than octanol–water partitioning.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.3564.

Acknowledgment—The present study was supported by the Dutch Technology Foundation STW, which is part of The Netherlands Organization for Scientific Research and is partly funded by the Ministry of Economic Affairs. Additional funding was received from Deltares (Utrecht, The Netherlands) and Environmental Risk Assessment and Management, which is a partnership of European detergent and surfactant producers. We thank M. Kraak for critical reading of the manuscript and R. Helmus for technical assistance in the laboratory.

Data availability—Data are available on request from the corresponding author, J. Hammer (jorthammer@gmail.com).

REFERENCES


