Hydrophilic species are the most reactive components of freshwater dissolved organic matter

Charlotte Grasset (charlotte.grasset@ebc.uu.se)  
Uppsala University  https://orcid.org/0000-0002-3251-7974

Marloes Groeneveld (marloes.groeneveld@ebc.uu.se)  
Uppsala Universitet

Lars Tranvik (lars.tranvik@ebc.uu.se)  
Uppsala University  https://orcid.org/0000-0003-3509-8266

Luke Robertson (luke.robertson@farmbio.uu.se)  
Uppsala Universitet

Jeffrey Hawkes (jeffrey.hawkes@kemi.uu.se)  
Uppsala University  https://orcid.org/0000-0003-0664-2242

Article

Keywords:

DOI: https://doi.org/

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License

Additional Declarations: There is NO Competing Interest.
Hydrophilic species are the most reactive components of freshwater dissolved organic matter


*Contributed equally
§ Corresponding author: jeffrey.hawkes@kemi.uu.se

Abstract
Aquatic dissolved organic matter (DOM) is a crucial component of the global carbon cycle, and the extent to which DOM escapes mineralization is important for the transport of organic carbon from the continents to the ocean. DOM persistence strongly depends on its molecular properties, but little is known about which specific properties cause the continuum in reactivity among different dissolved molecules. We investigated how DOM fractions, separated according to their hydrophobicity, differ in reactivity across three different inland water systems. We found a strong negative relationship between hydrophobicity and reactivity, consistent for the three systems. The most hydrophilic fraction was poorly recovered by solid phase extraction (SPE) (3-28% DOC recovery) and was thus selectively missed by mass spectrometry analysis during SPE. The change in DOM composition after incubation was very low according to SPE-ESI (electrospray ionization)-mass spectrometry (14% change, while replicates had 11% change), revealing that this method is sub-optimal to assess DOM reactivity, regardless of fraction hydrophobicity. Our results demonstrate that SPE-ESI mass spectrometry does not detect the most hydrophilic and most reactive species. Hence, they question our current understanding on the relationships between DOM reactivity and its molecular composition, which is built on the use of this method.

Main
Inland waters are a significant source of atmospheric greenhouse gases and they export substantial amounts of organic matter to the sea. Dissolved organic matter (DOM) is a major precursor for the greenhouse gas emissions and the extent to which it escapes mineralization is important for the transport of organic carbon from the continents to the ocean. In addition to extrinsic factors that control DOM biological degradation (e.g., temperature, light), the persistence of DOM in inland waters is related to the intrinsic properties of its chemical constituents. Compounds may be difficult to degrade because of their chemical structure or because they are extremely diverse with slight structural variations, making each substrate of vanishingly low concentration and difficult to use for microbes. The latter is termed the ‘dilution hypothesis’, and the two theories are actively debated. Evidence for either theory requires investigation of DOM reactivity at a molecular level to examine the lability of individual compounds at varying concentrations and in different ecological contexts.

Since DOM is composed of countless compounds with different reactivities, the rate of decay of DOM in incubation studies is not exponential, as would be expected from a single substrate with one first order reaction rate, but is instead built up of a continuum of different exponential
Previous work has shown that when bulk DOM is separated into low and high apparent molecular weight fractions with ultrafiltration (LMW and HMW, respectively), the HMW fraction is more reactive\textsuperscript{12,13}. This result has led to the so-called ‘size-reactivity continuum model’, in which DOM is theorized to be degraded to progressively smaller and less reactive forms over time\textsuperscript{14}. However, contradicting relationships between size and reactivity have been found for both apparent\textsuperscript{15,16} and actual molecular weight\textsuperscript{17,18}. Experimental studies that relate other characteristics of bulk DOM than size to reactivity are consequently needed.

The degree of hydrophobicity is another major characteristic that varies among DOM components, but this has not been previously related to bulk DOM reactivity. The hydrophilic character of a molecule describes its affinity to water and is related to its polarity (i.e., referring to the spatial distribution of electron density within the molecule). DOM mixtures can be separated by polarity since more hydrophobic species retain better on a hydrophobic material such as C\textsubscript{18}-bonded silica or styrene/divinyl benzene (e.g., Agilent PPL). Many labile biomolecular classes (e.g., sugars, free amino acids and peptides) are hydrophilic but surprisingly, to our knowledge, no study has tested how DOM fractions of different hydrophobicity differ in reactivity.

Recent research into DOM reactivity has suggested that the presence of some constituents of DOM can promote (or suppress) degradation of others\textsuperscript{19,20}, sometimes referred to as a “priming effect”. Therefore, bulk reactivity of DOM when all individual molecules occur together may not correspond to the sum of the reactivity of individual molecules or fractions when incubated separately. Consequently, it is not clear if DOM fractions of different hydrophobicity would interact in ways that affect their reactivities when they are degrading together.

An increasing number of studies have investigated the reactivity of DOM at the molecular level using high resolution mass spectrometry (MS) techniques\textsuperscript{7,21-24}. Specifically, in inland waters, the exponential decays of a multitude of compounds composing DOM have been related to their characteristics (O:C, H:C, molecular weight) using MS\textsuperscript{17}. Problematically, in high resolution MS analyses, the most hydrophilic fraction is excluded during the preliminary extraction and concentration step and during ionization\textsuperscript{25,26}. Therefore, the results of MS studies and other techniques that use extraction isolates (e.g., with solid phase extraction (SPE)) or electrospray ionization (ESI) MS, are biased towards more hydrophobic DOM and may not represent the reactivity of bulk DOM.

In this study we hypothesised that the most hydrophilic fractions of DOM would be the most reactive. Accordingly, the most hydrophobic fractions would persist due to low microbial ability to degrade them and their limited accessibility when dissolved in water. We compared the biological DOC loss of four fractions of differing hydrophobicity separated from three contrasting inland water samples (humic stream, clearwater and eutrophic lake samples). In addition, we monitored the DOC loss of all fractions pooled together and compared it to a theoretical DOC loss, assuming that the different fractions did not interact during degradation.
Results and Discussion

Consistent highest reactivity of hydrophilic fractions

When DOM was separated into fractions of different hydrophobicity, differences in reactivity emerged that were common for all sites. In all three sites, the most hydrophilic fractions (U and 1) were the most labile (k >= 0.005, 12-25% DOC loss at day 150; Table 1), and the most hydrophobic fractions (2 and 3) the most refractory (k < 0.005, 6-15% DOC loss at day 150; Table 1, Fig 1), supporting our hypothesis. There was however an overlap in reactivity between the different fractions, since all fractions comprised reactive DOC that was lost quickly in the first days of the incubation and more refractory DOC that remained at the end of the incubation. Nevertheless, a strong average effect was observed, evidenced by a strong and negative correlation between the initial constant decay and the hydrophobicity of the different fractions (Fig 2). This result, consistent for DOM obtained from three substantially different water bodies, reveals that hydrophobicity significantly contributes to the DOM reactivity continuum, with hydrophilic species being the most reactive.

We expected higher reactivity for hydrophilic fractions because known labile biomolecular classes are hydrophilic (e.g., sugars and amino acids). It is, however, uncertain if the intrinsic character of hydrophilicity generally results in higher reactivity. One reason for the higher reactivity of hydrophilic molecules could be that chemical functional groups on which most biodegradation reactions are based (e.g., hydrolysis, oxidation), are often polar (e.g., O containing functional groups). Indeed, in our study, the more hydrophilic fractions correlated with a higher abundance of high O/C compounds (Fig 3b, Text S1). Additionally, more hydrophobic DOM fractions could be relatively less reactive because hydrophobic species aggregate to decrease their extent of surface contact with water, and thus indirectly their accessibility to microorganisms.

The most hydrophilic compounds are outside of the SPE-ESI-MS analytical window

Our findings show that hydrophilic species are on average the most reactive but such species are generally lost during SPE prior to MS analysis. The DOC percentage that was recovered after SPE and analyzed by MS (but not necessarily detected) was low for the most hydrophilic fraction (3-28% fraction U all sites combined, median 18%) but substantially higher for the most hydrophobic fractions (58-103% for fractions 2 and 3, median 85%, Fig 4). The percentage of DOC recovery after SPE extraction of bulk environmental water samples is generally around 60-70%, close to our recombined samples (43-80%, median 70% for sample C, Fig 4). SPE on hydrophobic sorbent retains hydrophobic compounds and excludes the most hydrophilic compounds. It is generally a necessity for MS approaches to use SPE to concentrate DOM and remove salts as a preliminary step (with a PPL cartridge, although a few studies have managed to analyze samples from freshwater environments without pre-concentration on PPL). The variability in DOC recovery between the fractions in this study was expected because these fractions were previously already separated using a non-polar stationary phase (C18). It also confirms that a significant part of the hydrophilic DOM is lost during SPE. Additionally, comparison of ESI mass spectra before and after SPE show little difference in the spectral results, indicating that not just extraction, but also ionization and
transfer to the gas phase in electrospray is inefficient for hydrophilic species\textsuperscript{25}. Mass spectrometry is consequently sub-optimal for investigating bulk DOM and its reactivity.

Consistent with the limited ability of the MS analysis to explain changes in DOM reactivity, we also found a limited measurable change in MS composition before and after incubation (Fig 3a). Differences in composition between fractions of different hydrophobicity, and between sites, were generally much more important than those before and after incubation (Text S2). The dissimilarity in MS spectra before and after incubation, as quantified with the Bray-Curtis metric was on average $14 \pm 14\%$ (all fractions and all sites combined), and was the highest ($28 \pm 20\%$) for the most hydrophilic fraction (fraction U all sites combined). This was low considering that incubation replicates had $11 \pm 13\%$ dissimilarity. Our quantitative approach, based on DOC concentration, describing bulk DOM reactivity in relation to hydrophobicity is thus more broadly inclusive than the mass spectrometry peak abundance approach. There is substantial recent progress in the understanding of how the reactivity of DOM is related to its composition. However, this progress builds heavily on MS analysis\textsuperscript{7,17,22}, which suffers from loss of material during extraction as well as biases due to incomplete ionization. Future studies should investigate if the current knowledge still holds when assessing bulk DOM reactivity.

**Challenging the established relationships between DOM reactivity and H:C ratio**

Despite general low extraction efficiency and ionization coverage of the reactive DOM, we were able to see a consistent change in DOM composition before and after bio-incubation (Fig S1). Inspection of the mass of the molecular formulas lost revealed that the higher molecular mass compounds, especially those with comparatively fewer double bond equivalents, were more labile (Fig S2), corresponding well with the ‘size-reactivity continuum’ theory\textsuperscript{13}. The DOM composition changes were reproducible across the three sites. In the most hydrophobic fraction, lipid-like compounds with H:C $> 1.5$, O:C $< 0.5$ were preferentially removed (fraction 3, Fig S1). In the other fractions the most oxygenated species (O:C $> 0.6$) were the most prone to removal (fractions U, 1 and 2, Fig S1) as also found by other studies\textsuperscript{18,32,33}. This result complicates the prevailing concept that DOM degradability or persistence are mainly driven by aromaticity or H:C ratio\textsuperscript{7,10,17,23,34-36}.

**No consistent interactive effect in reactivity when all fractions are degraded together**

The three sites showed different patterns when all fractions were combined and incubated together. The DOC loss after 150 days in incubations where all fractions were recombined (C) was similar to what was theoretically expected (M), between 14 and 17\% for C and M samples (Table 1). However, for all three sites the reactivity continuum had a different shape from what is theoretically expected without an interaction between fractions (i.e., no shared symbol between samples C and M in Table 1), although the difference was sometimes visibly low (clearwater lake, Fig 1). For one site, the recombined sample (C) had a significantly higher reactivity, while for the two other sites it had a lower reactivity than expected (M) (Table 1). This suggests that there is no consistent synergistic or antagonistic effect when all fractions are degraded together. The mechanisms that would enhance or limit the degradation of different organic fractions are complex (described in detail in Sanches, et al.\textsuperscript{19} Bengtsson, et al.\textsuperscript{20}) and could be limited in aquatic ecosystems in comparison to soils, explaining why we did not find
a consistent effect\textsuperscript{20}. Alternatively, interactive effects occurred within one or several of the broad fractions that we examined, i.e., the components of DOM that interacted were not resolved. In addition, the average difference in reactivities between the fractions (Table 1) were possibly insufficient to result in a detectable and consistent effect across all three sites. Indeed, such interactive effects have been observed following the addition of single compounds (e.g., glucose) or in situ-produced DOM\textsuperscript{37-39} that are likely to have a higher reactivity than the most hydrophilic fractions in our experiment.

Recent progress in our understanding of the connection between the composition and reactivity of DOM builds heavily on mass spectrometry, which resolves individual molecular formulae at very high resolution\textsuperscript{7,17,22}, and on fluorescence spectroscopy, where emission-excitation spectra provide fingerprints that give information on the dynamics of broad classes of fluorescent DOM\textsuperscript{40}. These methods selectively miss the more hydrophilic DOM during extraction and ionization, and DOM that is not fluorescent, respectively, and they both provide only limited functional information. Here, we focus on one of the most profound characteristics of DOM, i.e., its degree of hydrophobicity, and we analyse the organic carbon concentration, limiting analytical bias. We show a strong connection between hydrophilicity and reactivity towards microbial degradation, revealing that studies that omit hydrophilic DOM cannot assess bulk DOM reactivity.

**Online Methods**

**Sampling sites**

Three inland water sites of the Uppland region (Sweden), one humic stream and two lakes with contrasting nutrient status and watershed characteristics, were selected because of their expected differences in DOM quality (Table S1). Fiby is a humic stream and thus has a short water retention time and a high abundance of fresh terrestrial ly-derived DOM. Långsjön (theoretical water residence time: 3-8 years) and Alstasjön (theoretical water residence time: 6 days) are meso-eutrophic and hypereutrophic lakes, respectively, and thus are expected to have higher contributions of in situ produced DOM than the humic stream.

To obtain chemical water characteristics of the sites, TP, TN and pH were measured on 60 µm plankton net filtered samples and DOC was measured on Whatman GF/F filtered water samples collected in October and November 2021 and stored at 4 °C in the dark before analysis. Total phosphorus concentrations were measured colorimetrically with a UV/Vis spectrophotometer (Lambda 40; Perkin Elmer; Waltham, Massachusetts, USA) using the molybdenum-blue method (Menzel & Corwin, 1965). Total nitrogen (TN) concentrations were determined on a TOC/TN analyzer (Shimadzu TOC-L/TNM-L, Kyoto, Japan). DOC concentrations were determined using a Sievers M9 total organic carbon (TOC) analyzer (GE Analytical Instruments, Boulder, Colorado, USA). pH was determined with a Metrohm 826 pH Mobile meter.

**Sample concentration by reverse osmosis**
About 50-150 L of water from each site was 3 to 22 times concentrated to a final concentration of approximately 140 mg DOC L\(^{-1}\) by reverse osmosis (Real Soft PROS/2S unit) in October and November 2021. Prior to concentration by reverse osmosis, the water was sequentially filtered through 5, 0.5 and 0.2 µm pore size membrane filters with a submersible pump through 10-in. filter cartridges and passed through a strongly acidic cation exchange resin (Dowex 50W X8, Dow Chemical Company).

Additionally, a smaller water sample (ca. 1 L) was filtered through a 60 µm plankton net to remove large particles and serve as a microbial inoculum during the incubation. All water samples were kept at 4 °C in the dark before the fractionation or before the start of the incubation.

**Sample fractionation**

Water samples were filtered within 24 h before fractionation with pre-combusted GF/F filters. To approximately 1 L of concentrated lake water sample, methanol (MeOH, ~50 mL) and trifluoroacetic acid (TFA, ~1 mL) were added to bring each to 5% MeOH and 0.1% TFA. Each sample was then individually loaded onto a preconditioned flash column (Biotage Sfär Duo C\(_{18}\), 120 g, 100 Å, 30 µm). After each sample was fully loaded onto the column, the unretained material was eluted with the manufacturer listed dead volume (160 mL) of 5% MeOH (0.1% TFA). The entire unretained eluent (~1 L) was collected into a bottle and labelled as the “unretained” fraction. Retained material was then eluted with 250 mL of 95% acetonitrile (CH\(_3\)CN; 0.1% TFA) and collected into a second bottle (250 mL), which was labelled the “retained” fraction. Both the “unretained” and “retained” fractions were lyophilized. The “unretained” fraction was weighed and stored in the freezer. The “retained” fraction was then purified using preparative HPLC (Kinetex XB-C\(_{18}\), 150 x 21.2 mm, 100 Å, 5 µm) using a gradient consisting of isocratic 5% CH\(_3\)CN (0.1% TFA) for five mins, then to 95% CH\(_3\)CN (0.1% TFA) over the next 50 mins. The column was then eluted with 95% CH\(_3\)CN (0.1% TFA) for five mins. The flow rate was 9 mL/min and fractions were collected every 60 seconds into pre-weighed glass test tubes. The test tubes were evaporated overnight using a centrifugal evaporator (30 °C) and weighed again to reveal the weight of each fraction (Fig S3).

**Incubation preparation**

Six different water samples were incubated for each site. Four fractions of increasing hydrophobicity were separated from concentrated DOC samples: fraction U (‘unretained’) corresponds to the unretained, most hydrophilic fraction and fractions 1 to 3 are the retained fractions of increasing hydrophobicity. C (‘combined’) is the recombination of U, 1, 2 and 3 in their original abundances to recreate a sample that is close to the original sample. In addition, the original concentrated water sample (abbreviated O for ‘original’) was included in the incubations and compared to the recreated original sample in order to assess if the fractionation affected the reactivity of the samples (Text S3).
The weights of the test tubes obtained after fractionation were evaluated along with the gradient conditions to arbitrarily choose the fractions 1, 2 and 3 with enough material in all three sites to make incubations with sufficient carbon concentration for analysis (Fig S3). Since most material was eluted within the first 35 minutes (corresponding to the first 35 tubes), the test tube ranges selected were tubes 1-11 (fraction 1), tubes 12-18 (fraction 2) and tubes 19-35 (fraction 3). Taking into account the dead volume of the column, fractions 1, 2 and 3 were eluted with 5-10%, 10-22% and 22-53% CH$_3$CN, respectively.

Fractions U, 1, 2 and 3 were then diluted, filtered and recombined into C as described below, over two days in January 2022, just before the start of the incubation, during which all samples were stored at 4 °C and in the dark when not processed. All tubes containing the freeze-dried retained fractions (fractions 1 to 3) and U were dissolved in MilliQ (Millipore) water and pooled together for the tubes corresponding to fractions 1, 2 or 3. The tubes were sonicated at least 3 times for 15 minutes at ca. 25 °C to help the material to dissolve. The retained fractions, and U and O samples were then filtered with pre-combusted GF/F filters to remove aggregates that could not be dissolved, or that were formed during storage for O. After this, C was made by pooling U and the retained fractions in the same proportion as for O (i.e., by pooling together 20 mL of the fractions previously diluted in 200 mL). At day 0 of the incubation, concentrated artificial lake water containing nutrients and other macro and micro constituents was added to each sample (U, 1, 2, 3, O and C) to reach the concentrations given in Bastviken, et al. (Table 1; TP 3.4 µg L$^{-1}$ and TN 71 µg L$^{-1}$). In addition, a microbial inoculum from each respective site was added to each sample to constitute 2 % of the total volume. All samples were further diluted with MilliQ to reach an initial DOC concentration of ca. 10 mg L$^{-1}$ (9.9 to 11.2 mg L$^{-1}$), except for fraction 1 from the clearwater lake, that was diluted to 5.6 mg L$^{-1}$ because of a lack of material. Note that the study design involved isolating DOM and separating it into polarity fractions, then redissolving the material into a standardized artificial lake water (common to all sites), and inoculating the samples with native bacteria from each site. Due to this approach, not everything about the water chemistry and biological community can be matched to in-situ conditions. For example, in two samples of the eutrophic site, we observed a low pH that hindered degradation (Text S3), but this effect did not alter the overall results and conclusions of the study.

**DOC and pH measurements**

Each fraction as well as C and O samples were divided in 30 headspace-free vials for the DOC analysis of 2 replicate samples at 15 different time points (3 sites * 6 samples * 30 = 540 tubes in total). DOC was measured with the TOC Analyzer at days 0, 2, 5, 9, 14, weekly and then bi-weekly over 150 days. pH was measured at the start and at the end of the incubation experiment in a 40 mL vial. In addition, 40 mL vials were prepared for MS analysis (two vials extracted and analysed in duplicate each at the start and at the end of the experiment) and absorbance and fluorescence (two vials analysed in triplicates at the start and at the end of the experiment).

**Solid phase extraction**
Solid phase extraction (SPE) was performed with 100-mg Bond Elut PPL cartridges (Agilent Technologies) within 10 days of the start of the incubation and within 2 days of the end (day 145). The cartridges were rinsed with methanol (hypergrade for LC-MS, Supelco), soaked in methanol for at least 2 hr, and then rinsed with 0.1% formic acid. The samples (40 mL, duplicates) were acidified to pH ≈ 2 with 6 M high purity HCl (Suprapure, VWR; as 50% in Milli-Q, 2 μlml⁻¹) and allowed to drip through the cartridges by gravity. The cartridges were flushed with 3 mL 0.1% formic acid to remove salts and then dried using N₂. The samples were eluted with 2 mL methanol into pre-combusted 2-ml amber vials and stored at -20 °C until analysis.

To quantify how much DOC was recovered after SPE extraction, part of the SPE extracts (ca. 0.7 mL of MeOH) was dried down in a water bath, redissolved in Milli-Q water, sonicated for 15 minutes and analyzed with the Sievers M9 total organic carbon (TOC) analyzer, after which the extraction efficiency was calculated.

Characterisation by mass spectrometry

Mass spectrometry was performed on the other part of the SPE extracts. 1 mL of SPE extracts were dried in a vacuum centrifuge and redissolved in 5% CH₃CN (LCMS grade, Supelco, 200 μL). 30 μL of each sample was injected in a liquid chromatography method (Agilent 1100), which used 1 mL/min isocratic flow of 25 mM ammonium acetate in 20% MeOH as mobile phase on a size exclusion column (Tosoh TSK Gel G3000SW 300 x 7.5 mm, 10 μm pore size). Eluent was split and directed to a charged aerosol detector to measure material abundance, and heated electrospay ionisation mass spectrometer (LTQ-Velos Orbitrap, Thermo Fisher) operating in negative mode, to measure the mass spectrum at approximately 1 transient per second. In this study, all transients were averaged together into a single peak list. All .raw and .mzXML files are available on the MassIVE data repository.

Potential doubly charged interferences were removed, along with spectral noise, and then formulas were assigned to the remaining peak list after internal calibration first to mass 369.11911, and then to a series of masses that are common to all DOM samples. Combinations of C (4-50), H (4-100), O (2-40), N (0-2) and S (0-1) were allowed, along with up to one ¹³C. Allowed formulas had to be in mass range 150-800, H/C 0.3-2.2, O/C<1, double bond equivalence minus oxygen 10 to -10, and could contain no more than one of the elements/isotopes N, S and ¹³C.

Assigned sample peak lists were normalised and a Bray-Curtis dissimilarity matrix was calculated, which formed the basis of a principal coordinate analysis (PCoA). Finally, the sample-wise normalised intensity of each molecular formula was analysed for correlation with sample position on principal coordinate 1 and 2 using Pearson’s rho, to determine how the intensity of individual molecular formulas covaried with overall molecular composition and sample dissimilarity. The full MATLAB code used for assignment, distance matrix and PCoA and covariance testing is available in SI.
**Characterisation by excitation emission spectroscopy**

UV–Vis absorbance spectra (250 to 600 nm) were measured in a 1-cm quartz cuvette using a Lambda35 UV–Vis Spectrometer (PerkinElmer Lambda 25, Perkin Elmer, Waltham, USA). Fluorescence scans were obtained using a FluoroMax-4 Spectrofluorometer (FluoroMax-4, Jobin Yvon, Horiba, Kyoto, Japan), with excitation-emission matrices (EEMs) from excitation wavelengths 250 to 445 nm with 5-nm increments and emission wavelengths 300 to 600 nm with 4-nm increments. A Milli-Q water blank run on the same day was used to correct the spectra, instrument biases and inner filter effects were corrected for, and the spectra were normalized to Raman units\(^4^4,4^5\) using the FDOMcorr toolbox\(^4^6\) for MATLAB (Mathworks, Inc., Natick, MA). The main DOM fluorescence components that varied throughout the data set were identified using PARAFAC\(^4^7\). The analysis was conducted on a set of 114 samples (3 sites, 6 samples per site, 2 time points, triplicates, plus 6 EEMs from the pH test) using the drEEM toolbox for MATLAB (Mathworks, Inc., Natick, MA) following Murphy, et al.\(^4^8\).

Primary and secondary Rayleigh and Raman scattering were removed and smoothed over, and the data was normalized to total fluorescence intensity of each sample. Nonnegativity constraints were applied on all modes (excitation, emission, and sample). The appropriate number of components was identified considering the effect of adding more components on the model fit (expressed as the sum of square errors), by visual inspection of the residuals and random initialization with 10 iterations with a convergence criterion of \(1 \times 10^{-8}\) to find a stable model. The model was validated using random split-half analysis by splitting the data set into 3 subsets.

**Statistical analyses**

The fraction of remaining DOC at time \(t\) (\(\frac{DOC_t}{DOC_0}\), unitless) was described using the reactivity continuum model that has previously been used in several inland water studies (e.g.\(^1^0,1^1,4^9\)):

\[
\frac{DOC_t}{DOC_0} = \left(\frac{a}{a + t}\right)^v
\]

\(a\) (days) is a rate parameter, it is the average lifetime of the more reactive DOC components

\(v\) (unitless) is a shape parameter. It relates to the preponderance of refractory compounds, a low \(v\) suggests the prevalence of refractory compounds\(^1^0,5^0\)

\(k_0\) (day\(^{-1}\)), the initial apparent decay coefficient, was calculated as \(v/a\) as an indicator of DOC overall reactivity.

A theoretical remaining DOC, noted ‘M’ in the rest of the manuscript, was calculated as the sum of the remaining DOC (\(\frac{DOC_t}{DOC_0}\), unitless) of the separated fractions (U, 1, 2, 3), multiplied by relative abundance of the fractions in the original sample.

Remaining DOC was modelled for all samples (U, 1, 2, 3, C, M and O) and for each site (humic stream, eutrophic and clearwater lakes) using a non-linear model with fraction as factor (gnls...
function; package nlme\(^{51}\). For the first retained fraction of the humic site, there was high variability of replicates from day 90, resulting in poor model performance. We consequently only included remaining DOC until this day. The quality of the models was assessed by checking residuals and by plotting measured values against modelled values. The significance of the fixed effects on the model parameters was tested with ANOVA. In addition, we tested if the model parameters \((a\text{ and } v)\) significantly differed between the different samples within each site by comparing models with different sets of parameters with ANOVA\(^{52}\). More specifically, this was done by testing if sharing both \(a\) and \(v\) for different samples decreased the model performance.

The correlation between the hydrophobicity of the fractions (U, 1, 2 and 3) and \(k_0\) was tested with a Spearman correlation. The proportion of the strong mobile phase (CH\(_3\)CN) relates to the affinity of DOM for the hydrophobic stationary phase (i.e., the C\(_{18}\) column), and was used as a proxy of DOM’s hydrophobicity. The hydrophobicity of the fraction was consequently assessed by the weighted average proportion of CH\(_3\)CN used to elute the fraction and was set to 0 for the unretained most hydrophilic fraction.
Figure 1. Measured (dots) fraction of remaining DOC (DOC/DOC₀, unitless) over time and reactivity continuum model (lines). U, 1, 2 and 3 are the DOM fractions of increasing hydrophobicity, U is the unretained and most hydrophilic fraction, 3 is the most hydrophobic fraction. C is the recombination of U, 1, 2 and 3 in their original abundances. M (circles) is the theoretical remaining DOC of the recombined fractions, calculated assuming an additive effect.
Figure 2. Correlation between the hydrophobicity of the different fractions (U, 1, 2 and 3), assessed by the weighted average proportion of the strong mobile phase used on the C\textsubscript{18} column (CH\textsubscript{3}CN) to elute the fraction, and the initial decay constant (log $k_0$). The hydrophobicity of the fractions increases with the eluent concentration, the concentration of the eluent was set to 0 for the most hydrophilic fraction which was not retained by the column. The Spearman correlation coefficient was $r=-0.915$ (p-value = 2.9*10\(^{-5}\)).
Figure 3. Differences in DOM composition between the samples (U, 1, 2, 3 and C) of the three sites, and before/after incubation, as assessed by high resolution MS. Principal coordinate analysis (PCoA) plot of the different samples in the three sites before and after incubation (a) and correlations between the first (b) and second (c) PCoA axes and the abundance of the individual molecular formulae (n= 3695 for PCoA1 and n=1772 for PCoA2). In b) and c), the color scale indicates the significant Spearman correlations (|r| > 0.334 for p value < 0.01 and n=59 samples) and each molecular formula is represented by one dot according to its H:C vs. O:C ratio.
Figure 4. Percentage of DOC recovery after solid phase extraction (SPE efficiency in %) for each sample ‘U’ unretained and most hydrophilic fraction, 1-3 retained fractions of increasing hydrophobicity, and C all fractions recombined in their original abundances.
Table 1. Parameters of the reactivity continuum model of remaining DOC over time and predicted DOC loss.

<table>
<thead>
<tr>
<th>Site (R²)</th>
<th>Sample</th>
<th>α</th>
<th>ν</th>
<th>k₀</th>
<th>Modelled DOC loss at 150 days (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>clearwater lake</td>
<td>U</td>
<td>0.4 ± 0.1</td>
<td>0.048 ± 0.002</td>
<td>0.121</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2.3 ± 0.5</td>
<td>0.031 ± 0.002</td>
<td>0.014</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.1 ± 1.7</td>
<td>0.023 ± 0.003</td>
<td>0.004</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.1 ± 3.7</td>
<td>0.023 ± 0.004</td>
<td>0.002</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.4 ± 0.1</td>
<td>0.028 ± 0.001</td>
<td>0.071</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.4 ± 0.1</td>
<td>0.031 ± 0.002</td>
<td>0.072</td>
<td>17</td>
</tr>
<tr>
<td>eutrophic lake</td>
<td>U</td>
<td>1.8 ± 0.3</td>
<td>0.036 ± 0.002</td>
<td>0.020</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6.5 ± 0.8</td>
<td>0.069 ± 0.003</td>
<td>0.011</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>2*</td>
<td>17.7 ± 3</td>
<td>0.069 ± 0.006</td>
<td>0.004</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20.4 ± 4.4</td>
<td>0.057 ± 0.006</td>
<td>0.003</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.5 ± 0.9</td>
<td>0.049 ± 0.003</td>
<td>0.009</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>M*</td>
<td>10.2 ± 2.5</td>
<td>0.053 ± 0.006</td>
<td>0.005</td>
<td>14</td>
</tr>
<tr>
<td>humic stream</td>
<td>U†</td>
<td>13.6 ± 2.4</td>
<td>0.068 ± 0.005</td>
<td>0.005</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>10.6 ± 2.4</td>
<td>0.068 ± 0.007</td>
<td>0.006</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2§</td>
<td>33.8 ± 7</td>
<td>0.095 ± 0.011</td>
<td>0.003</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>29.2 ± 8.1</td>
<td>0.066 ± 0.01</td>
<td>0.002</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>C§</td>
<td>45 ± 9.7</td>
<td>0.113 ± 0.015</td>
<td>0.003</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>M†</td>
<td>7.6 ± 2.7</td>
<td>0.051 ± 0.008</td>
<td>0.007</td>
<td>14</td>
</tr>
</tbody>
</table>

Samples: U unretained, most hydrophilic fraction; 1-3 retained fractions of increasing hydrophobicity; C and M experimental and theoretical remaining DOC of all fractions recombined in their original abundances.

The model parameters that are not statistically different for the samples within each site share the same symbols (*†§). Fractions with no sign differ statistically from all other fractions within the site.

α and ν are the parameters given by the reactivity continuum model (given ± SE), α (rate parameter) relates more to the initial reactivity and ν (shape parameter) to the part where the curve levels off. A low α and a high ν indicate a higher reactivity of DOC.

k₀, the initial apparent decay coefficient, was calculated as ν/α as an indicator of DOC overall reactivity, it increases with the reactivity of DOC.

For each site, the R² of the model is given by the regression between measured and modelled data.
References


Dulaquais, G. *et al.* Size exclusion chromatography and stable carbon isotopes reveal the limitations of solid phase extraction with PPL to capture autochthonous DOM.


Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FractionpaperSMvfinal.pdf
- FractionpaperSMvfinal.pdf