Composition of dissolved organic matter within a lacustrine environment.

Margaret V. McCaul [A], David Sutton [C], André J. Simpson [B], Adrian Spence [A], David J. McNally [B], Brian. W. Moran [A], Alok Goel [D], Brendan O'Connor [D], Kris Hart [A], Brian P. Kelleher [A, E]

[A] *School of Chemical Sciences, Dublin City University, Glasnevin, Dublin 9, Ireland.*

[B] *Department of Chemistry, University of Toronto, Scarborough College, 1265 Military Trail, Toronto, Ontario, M1C1A4.*

[C] *School of Science, Limerick Institute of Technology, Limerick, Ireland.*

[D] *Center for Bioanalytical Sciences and School of Biotechnology, DCU, Glasnevin, Dublin 9, Ireland.*

[E] *Corresponding author. E-mail: brian.kelleher@dcu.ie*

Environmental context.

Freshwater dissolved organic matter (DOM) is a complex mixture of chemical components that are central to many environmental processes, including carbon and nitrogen cycling, but questions remain as to its chemical characteristics, sources and transformation mechanisms. We studied the nature of DOM in a lake system and found that it is influenced by anthropogenic activities and also by soil microbial biomass. Human activities can therefore influence the huge amounts of carbon sequestered as DOM.

Abstract

Freshwater dissolved organic matter (DOM) is a complex mixture of chemical components that are central to many environmental processes, including carbon and nitrogen cycling. However, questions remain as to its chemical characteristics, sources and transformation mechanisms. Here, we employ 1- and 2-D nuclear magnetic resonance (NMR) spectroscopy to investigate the structural components of lacustrine DOM from Ireland, and how it varies within a lake system, as well as to assess potential sources. Major components found, such as carboxyl-rich alicyclic molecules (CRAM) are consistent with those recently identified in marine and freshwater DOM. Lignin-type markers and protein/peptides were identified and vary spatially. Phenylalanine was detected in lake areas influenced by agriculture while it is not detectable where zebra mussels are prominent. The presence of peptidoglycan, lipoproteins, large polymeric carbohydrates and proteinaceous material supports the substantial contribution of material derived from microorganisms. Evidence is provided that peptidoglycan and silicate species may in part originate from soil microbes.

Key words: Dissolved organic matter, NMR, lake, variability.

1. Introduction

Dissolved organic matter (DOM), both marine and freshwater, comprises of the largest pool of exchangeable carbon on the Earth's surface and is derived from numerous sources that influence its relative reactivity and our ability to predict its storage capacity and turnover times.[1,2] Terrestrial and freshwater DOM, whose input to ocean waters is largely controlled by riverine sources, experiences an annual flux of ca. 0.4×10^{15} g C/year to the marine environment.[3] The cycling of DOM from fresh to marine water is not only important in the global carbon cycle but also plays an important role in the enhanced solubility, bioavailability and fate of chemical contaminants and their global transport.[4,5]

Despite this importance, there is still much to learn about the chemical composition of freshwater DOM and how chemical constituents vary worldwide, and between freshwater and marine environments.[6,7] The application of NMR to study structures and interactions in environmental chemistry is growing and is a powerful tool in helping unravel the key structural components in major global carbon pools.[8-11] In recent work, 1- and 2-D solution state NMR spectroscopy has shown that major structural components of lake freshwater include carboxyl-rich alicyclic molecules (CRAM), heteropolysaccharides and aromatic compounds.[12] These components were first reported, and are consistent with those identified, in marine DOM.[13] Furthermore, it has been tentatively suggested that CRAM may be derived from cyclic terpenoids.[12,13] However, it is not clear whether these precursors are of terrestrial or aquatic origin or whether transformations proceed via biological and/or photochemical processes.

Traditional methods of DOM isolation require large sample volumes to overcome the low concentration in natural waters[14,15] or are laborious and time consuming.[16] Sampling is often carried out over just one or two days, which is unlikely to be long enough to provide a representative sample of the area. The samplers employed in this study were deployed over a four week period and provide a more representative material that is less susceptible to specific daily fluxes. Another advantage of using passive samplers of this kind is that filtration is not required, reducing the possibility of contamination and loss of

material. It has also been shown that the material collected on the samplers is similar to that collected using conventional DEAE-cellulose batch extraction, indicating that the passive sampler approach isolates the same components.[17] The same study also reported that 72-89% of total DOM can be captured on the sampler, with the majority of lost material comprising low molecular weight sugars.

While recent studies have contributed greatly to our knowledge of the overall composition of DOM, less is known of its mechanisms of formation, compositional variation and the origin of the most refractory DOM. Here we use DEAE-cellulose passive samplers, as reported by Lam and Simpson,[17] to concentrate DOM from different areas in Lough Derg, a large lake system on the River Shannon in Ireland. The River Shannon is the largest catchment within Ireland and Britain, draining a land area of ca. 18,000km^2. Lough Derg, the third largest lake in Ireland, is located at the southern end of the Shannon and covers an area of 120 km^2. NMR (both 1- and 2-D) is employed to study DOM structure and how it varies within a lake system and assess anthropogenic influence on its composition. The potential of surrounding soil microbial biomass as a source of DOM is also investigated by comparison of the NMR spectra of degraded soil microbial biomass and leachate to the DOM spectra.

2. Materials and Methods

2.1. Sampling and sample preparation

Six sampling sites around Lough Derg (Fig. 1) were chosen to represent areas influenced by different aspects of the surrounding landscape. At each site, two passive samplers (containing six membranes) were placed and suspended (using a fishing line) ca. 100 cm below the surface of the water. Samples were removed from the lake after 28 days. Sampling at the same sites was carried out in August 2008 and January 2009, so as to assess temporal variation in DOM components. DOM was isolated using a passive sampler.[17] Water from the lough was prefiltered through 0.22 μm poly(vinylidene difluoride) (PVDF) filters (Spectrapor). DOM was isolated on diethylaminoethyl cellulose resin (Sigma Aldrich), a selective resin that adsorbs negatively charged species

at neutral pH. The cellulose resin is contained within the PVDF tubing and protected via a high density polyethylene (HDPE) casing with predrilled holes. Prior to use, DEAE-cellulose was precleaned using a cycle of acid, base and distilled water washings. Cleaned DEAE-cellulose (250 mg) was slurry packed with distilled water into 7 cm long (24 mmwide) PVDF porous membranes, which were pre-soaked in 0.1% sodium azide for a minimum of 48 h.

Extraction of bound DOM from the passive samplers was performed by cutting and removing the resin from the PVDF membranes. The resin was then placed in 50 ml Teflon centrifuge tubes and extracted using ca. 40 ml of 0.1 M NaOH. The tubes were centrifuged (10000 g, 10 min) to pellet the resin, and the supernatant was decanted. The pellet was re-suspended and the previous steps were repeated four times, or until the extracting solvent was colourless, to ensure complete extraction of DOM from the resin. The extracted DOM was ion-exchanged using Amberjet 1200H Plus resin (Aldrich) and freeze-dried. Duplicate samples were freeze-dried and samples were re-suspended in deuterium oxide (D_2O) for NMR analysis.

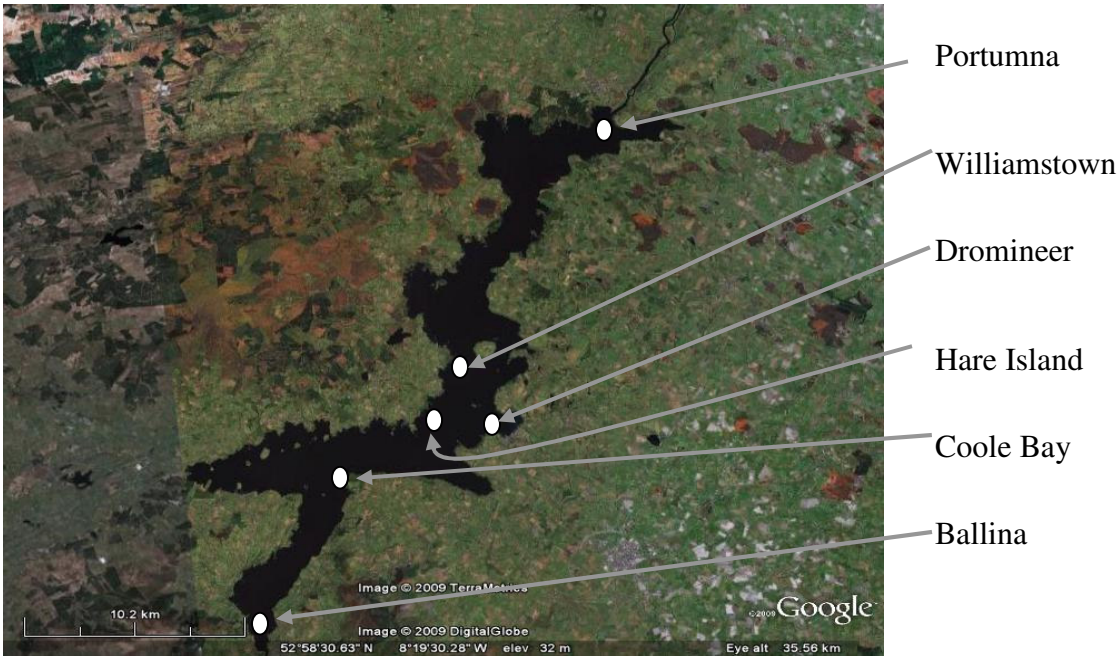

Fig. 1. Satellite image of Lough Derg (Ireland) and environs showing the six sampling sites.

2.2. NMR

Each sample (100 mg) was dissolved in 1 ml D_2O and titrated to pH 13.1 using NaOD (40% by wt) to ensure complete solubility. Samples were analyzed using a Bruker Avance 500 MHz NMR spectrometer equipped with a ^{1}H-BB-^{13}C 5 mm, triple resonance broadband inverse probe at 298 K. 1-D solution state ^{1}H NMR experiments were performed with 256 scans, a recycle delay of 3 s, 32768 time domain points, and an acquisition time of 1.6 s. Solvent suppression was achieved by presaturation utilizing relaxation gradients and echoes.[18] Spectra were apodized through multiplication with an exponential decay corresponding to 1 Hz line broadening, and a zero filling factor of 2. Diffusion-edited (DE) experiments were performed using a bipolar pulse longitudinal encode-decode sequence.[19] Scans (1024) were collected using a 2.5 ms, 49 gauss/cm, sine-shaped gradient pulse, a diffusion time of 100 ms, 8192 time domain points and 410 ms acquisition time Spectra were apodized through multiplication with an exponential decay corresponding to 10 Hz line broadening and zero filling factor of 2.

Total correlation spectroscopy (TOCSY) spectra were obtained in the phase sensitive mode, using time proportional phase incrimination (TPPI). TOCSY with presaturation of the solvent resonance was acquired using 2048 time domain points in the F2 dimension and 128 scans for each of the 128 slices in the F1 dimension. A mixing time of 60 ms was used with a relaxation delay of 1 s. Processing of both dimensions used a sine-squared function with a $\pi/2$ phase shift and a zero-filling factor of 2**.** TOCSY data was collected to help confirm the major assignments highlighted on the ^{1}H-^{13}C NMR correlations.

Heteronuclear multiple quantum coherence (HMQC) spectra were obtained in phase sensitive mode using Echo/Antiecho gradient selection. The HMQC experiments were carried out using 256 scans with 128 time domain points in the F1 dimension and 1024 time domain points in the F2 dimension. A relaxation delay of 1 s and ^{1}J ^{1}H-^{13}C of 145 Hz were used. F2 dimensions in HMQC experiments were processed using an exponential function corresponding to a 15 Hz line broadening. The F1 dimension was processed using a sine-squared function with a $\pi/2$ phase shift and a zero-filling factor of 2.

Spectral predictions were carried out using Advanced Chemistry Development's ACD/SpecManager and ACD/2D NMR Predictor using Neural Network Prediction algorithms (version 10.02). Parameters used for prediction including line shape, spectral resolution, sweep width and spectrometer frequency were set to match those of the real datasets as closely as possible. Please see accessory materials for an example.

2.3. Growth and degradation of soil microbial biomass

The soil used in this study is a light clay loam from a cultivated field near Lough Derg. Sampling was carried out according to a modified version of the protocol described processing Joseph et al. (2003). A 25-mm-diameter clean metal core was used to sample 100-mm long soil cores from the A horizon, which were transferred to sterile polyethylene bags and sealed at the collection site. Soil cores were transported at the ambient temperature and processed within 24 h of collection. The upper 30 mm of each core was discarded, and large pieces of roots and stones were removed from the remainder, which was sieved through a stainless steel sieve with a 2-mm aperture (IMPACT Laboratory Test Sieve, UK). Sieved samples were pooled, homogenized and stored at 4°C at its field moisture content for further analysis. A CHN combustion analyzer (Exeter Analytical CE440 elemental analyser) was used to determine the soil elemental composition, 4.25% C, 0.58% H, 0.15% N and 0.21% P.

Soil microbes were cultivated according to a modified version of the protocol described by Janssen *et al.*[20] Soil (1 g) was added to 100 ml aliquots of sterile distilled water and dispersed with a magnetic stirrer. Aliquots (1 ml) of soil suspension were added to 9 ml portions of dilute nutrient broth (DNB), containing gl^{-1}: Lab-Lemco' powder 1.0; yeast extract 2.0; peptone 5.0 and NaCl 5.0, at a concentration of 0.08 gl^{-1} distilled water (Oxoid Ltd., Hampshire, England). Diluted soil suspensions were mixed by vortexing at ca. 150 rpm for 10 s and used to prepare serial dilutions containing 10^{-2} to 10^{-4} g soil suspension. Aliquots (100 µl) of each dilution series was plated on duplicate LB agar plates containing 0.5% dripstone, 0.25% yeast extract, 0.1% *D*-glucose, 0.25% NaCl and 1.5% agar. Serially inoculated LB plates were incubated at room temperature for 2 days and all isolated colonies were selected from the 10^{-4} dilution of the soil and used to inoculate 3.0 ml LB broth. Cultures were incubated at for 48 h.

The degradation experiment was conducted according to a modified version of the protocol described by Kelleher *et al*.[21] The experimental design attempted to mimic *in situ* conditions and enable collection of transformed and leached organic matter (OM) for further analysis. Glass funnels with borosilicate sintered discs, with porosity grade 4 were submerged until flush with soil in a clay pot. The soil used was a native light clay-loam taken from fields surrounding Lough Derg. The cavity beneath the sintered disc was filled with the native soil and secured with glass wool and 0.4 g of the soil microbial biomass evenly distributed on the surface of the sintered disc. This set up enables microbes in the soil to access the microbial biomass. The biomass was sprinkled with water every second day to mimic rain and the runoff was collected in a vial attached to the end of the funnel. Moisture levels were kept constant throughout the experiment. Runoff and microbial biomass were collected at 6 and 14 weeks post degradation.

3. Results and discussion

3.1. General characterisation

Recent studies that have employed multidimensional NMR spectroscopy to study DOM show that marine and freshwater DOM share many structural similarities.[12,13] These major structural components are also present in all the DOM isolated from Lough Derg. For example, Fig. 2 shows the conventional ^{1}H (Fig. 2A) and diffusion edited (Fig. 2B) NMR spectra for the Ballina DOM sample and also show the diffusion edited ^{1}H spectrum of the Coole Bay sample area of lake (Fig. 2C).

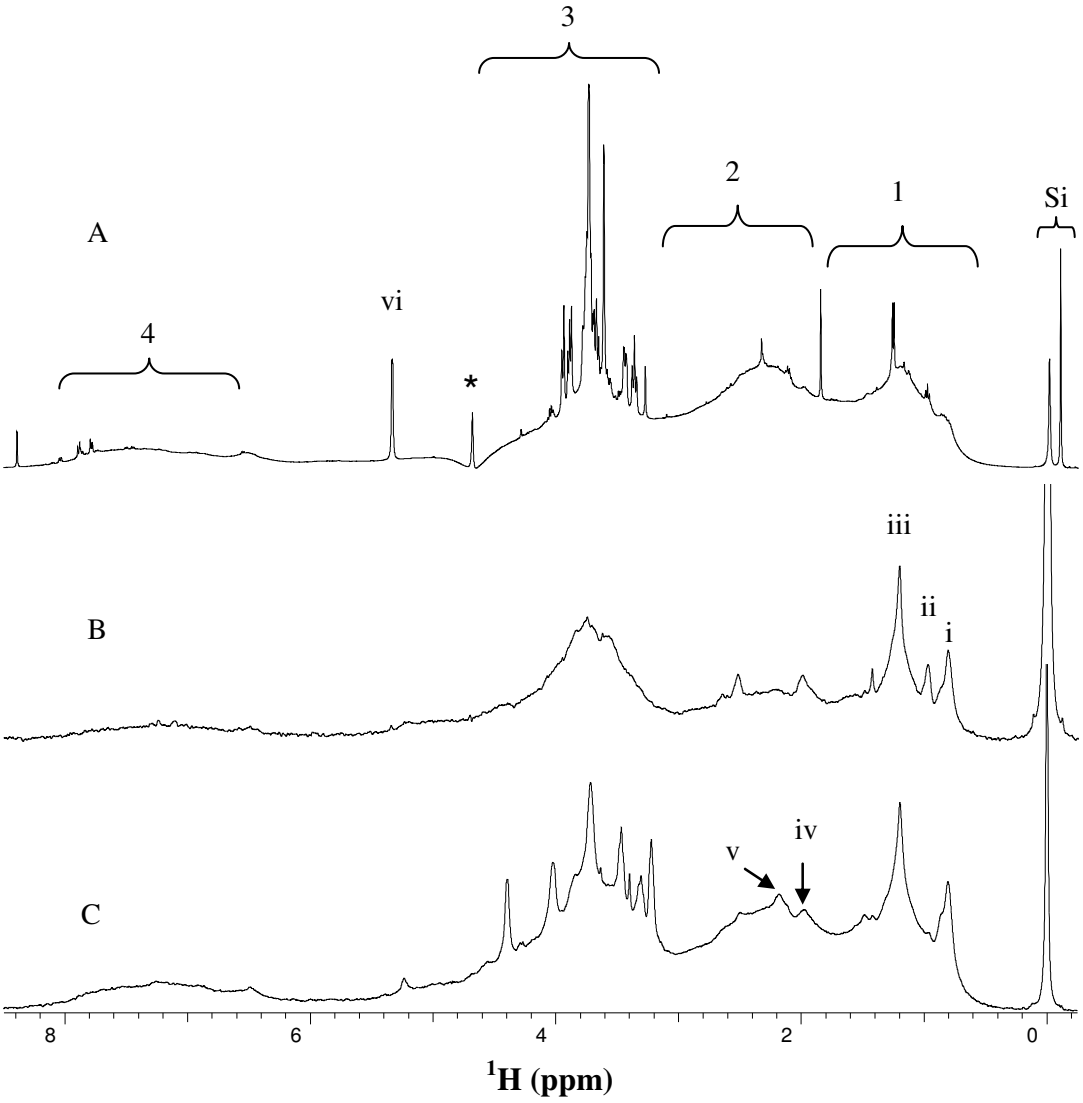

Fig. 2. ^{1}H NMR spectra for (A) Ballina DOM sample, (B) diffusion edited 1H spectrum of Ballina DOM sample, (C) diffusion edited ^{1}H spectrum of Coole Bay sample area of lake. *Indicates residual water signal. Designations 1-4 indicate general spectral regions: 1) linear terpenoids; 2) carboxyl-rich alicyclic molecules (CRAM); 3) carbohydrates and amino acids; 4) aromatics and amino acid side chains. Designations i-vi indicate specific assignments: i) aliphatic CH_3; ii) protein side chain residue; iii) aliphatic methylene $(CH_2)_n$; iv) *N*-acetyl group in

peptidoglycan or other constituents in lipids/waxes; v) aliphatic methylene units β to an acid or ester or double bond; vi) anomeric protons in carbohydrate. Si indicates a natural silicate and not TMS.

General assignments, consistent with those reported are: (1) aliphatics, including material derived from linear terpanoids; (2) carboxyl-rich alicyclic molecules (CRAM; see also Fig 5); (3) a mixture of carbohydrates and amino acids; (4) aromatics, including resonances from amino acid (AA) side chains.[12,13] More specific assignments refer to (i) CH_3, likely including resonances from aliphatic species and methylated amino acid side-chain residues in peptides/protein, (ii); consistent with a side chain residue also seen in the ^{1}H NMR spectrum for bovine serum albumin, (iii); aliphatic methylene $(CH_2)_n$, (iv); contributions from both *N*-acetyl group in peptidoglycans and other units lipids/waxes,[22,23] (v); mainly aliphatic methylene units β to an acid or ester i.e. R_2-OCO-CH_2-R_1 or double bond vi); anomeric protons in carbohydrate. 'Si' indicates a natural silicate species and not TMS (tetramethylsilane, $Si(CH_3)_4$), a commonly used NMR reference standard).[22]

Fig. 2A displays sharp peaks, especially in the carbohydrate region (3). Sharper lines observed in NMR are often characteristic of smaller structures [21], and this may indicate the breakdown of the carbohydrates from large polymeric structures into smaller fragments. To test this, diffusion edited (DE) NMR was performed on the Ballina sample. In diffusion edited NMR experiments, small molecules are essentially gated from the final spectrum but signals from macromolecules which display little translational diffusion are not gated and appear in the spectrum.[19,24] The diffusion edited spectrum of Ballina DOM is shown in Fig. 2B. Aliphatic chains are prominent, indicating that they have restricted diffusion, which suggests that they may be present in rigid domains or macromolecular structures. The relative intensity of the carbohydrate signals is much less in the diffusion edited spectrum vs. the conventional ^{1}H NMR spectrum, suggesting a large fraction of the carbohydrates in the DOM is present as relatively small mobile entities. However, there is still a considerable contribution from carbohydrate signals in the diffusion edited spectrum, supporting a second fraction of carbohydrate with greater molecular (or aggregate) size.

A characteristic resonance for CH_3 in methylated AA side chain residues (Fig. 2, signal i) is easily distinguishable in the diffusion edited NMR, suggesting the presence of protein/peptide.[22] Furthermore, the resonance at ca.1 ppm (Fig. 2, signal ii) is likely attributed to protein/peptide as this peak is also present in the ^1H NMR spectrum of bovine serum albumin.[25] Complimentary evidence for protein/peptide presence is provided by the emergence of α protons from AAs in Fig. 3. Proteinaceous compounds are viewed as labile in the environment [26] and their survival and occurrence have been explained through protection mechanisms such as encapsulation and formation of microbially resistant complexes with carbohydrates and lignin.[27-29] Lam *et al*.[12] detected weak protein/peptide contributions and was considered to be only a minor component in Lake Ontario DOM. However, the spectra generated indicate that the protein/peptide contribution may vary considerably between DOM from different sources in freshwater environments. It is estimated that plants often contain only 1–5% protein by weight and that protein structures are known to degrade rapidly in a soil environment.[30,31] It seems unlikely that the preservation of plant-derived peptide/protein structures can completely account for the contributions of proteins and peptides in DOM. It is therefore possible that a significant portion of peptide/protein in DOM arises from the cells of dead and living microbes of either aquatic or terrestrial origin.

Alternatively, microbially resistant ligno-protein complexes may also account for some of the protein present.[32] Lignin-type signatures were not found in the study of Lake Ontario DOM,[12] but the possibility of lignin contributions to Lough Derg DOM is highlighted by cross peaks that may represent lignin derived O-CH_3 units (Fig. 3), often the most intense signal in soil OM.[9,21,33] Methoxy cross peaks are clearly present in all the lake samples (overlapped with carbohydrate crosspeaks), especially Hare Island and Dromineer. Lignin is a strong indicator of terrestrial plant inputs and may be an indication of the age of DOM and/or the influence of the surrounding environment. Proteins originating from microbial cells may be encapsulated by, or sorbed to, lignin, making them less susceptible to degradation.

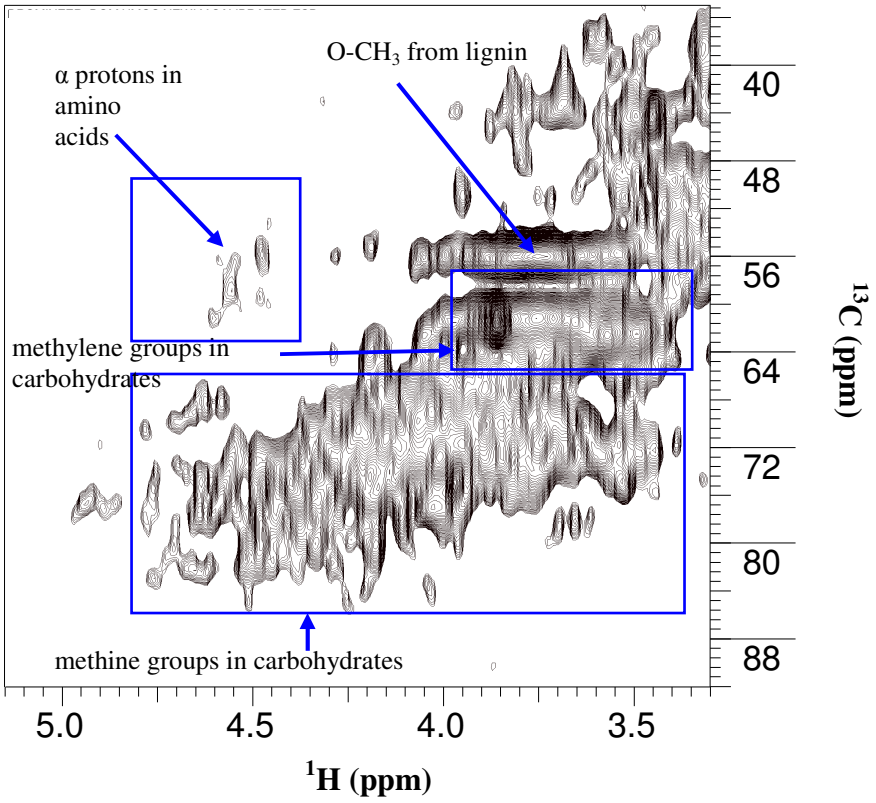

Fig. 3. Zoom region of DOM ^{1}H-^{13}C HMQC (Dromineer).

All Lough Derg DOM samples contain a contribution from carbohydrates that are not removed during diffusion editing (Fig. 2B, C) indicating that there is a polymeric carbohydrate component present that could potentially be associated with the cell walls of microorganisms.[12] Signals (iii) and (v) in Fig. 2B and 2C are consistent with aliphatic structures. The aliphatic $(CH_2)_n$ peak is dominant, indicating the presence of stable waxes and lipids.[34] Waxes and cutins derived from plants have been identified in abundance in humic extracts,[24] and are likely to be preserved because of their cross-linked structure and hydrophobicity.[35] Fig. 2C shows the DE ^{1}H NMR spectrum for the Coole Bay DOM sample. Signal (v) is particularly prominent and shows similarities to signals from lipoproteins observed in other natural samples.[22] Lipoprotein is a key component of bacterial cells (also plant, animal, yeast, fungal, algal and insect cells), is structurally diverse and is released during bacterial growth,[36] so its presence corroborates the importance of terrestrial microbes as sources of DOM. Microbial contributions are also supported by the presence of signal (iv) in Fig. 2C. This is consistent with peptidoglycan,

which comprises up to 90% by weight of Gram-positive bacteria and is the key structural component in all microbial cell walls. That peptidoglycan was found to accumulate is not unexpected since it is resistant (as are microbe cell walls) to many chemical and biological processes and has been found in the most refractory components of soil OM.[22]

3.2 Soil microbial contribution

Despite strong microbial signatures in Fig. 2 it is difficult to know from where the microbial residue originates. It has recently been shown that microbial presence in soil far exceeds presently accepted values and that considering the amounts of fresh cellular material in soil extracts, it is probable that the contributions of micro-organisms in the terrestrial environment are seriously underestimated.[25] Therefore, soil microbial biomass may also be an important source of freshwater DOM. The potential contribution of surrounding soil microbial biomass to Lough Derg DOM was studied by conducting a complementary laboratory experiment that monitored the degradation of soil microbial biomass cultured from soil sampled near the lake. Degradation occurred over 14 weeks, allowing NMR experiments to be conducted on degraded soil microbial biomass residue and leachate. Fig. 4 compares the DE ^{1}H NMR spectra of the 14 week leachate from degraded soil microbial biomass (A), to the "Dromineer" DOM sample from Lough Derg (B). Characteristic resonances, such as CH_3 in methylated AA side chain residues (signal i) and aliphatic methylene $(CH_2)_n$) (signal ii) that are present in the Dromineer DOM (Fig. 4B) are also present in the microbial leachate. These signals also persist in degraded plant matter, so it is not possible to say that they originate solely from soil microbial biomass.[26] However, peptidoglycan (PG, Fig. 4A, B) is present in both the DOM and the soil microbial biomass leachate, and this is confirmed in the HMQC spectra in Fig. 5. It should be noted that the slight shift in the proton axes of the PG microbial biomass is from the solvent (DMSO) used to swell the microbial biomass for analysis using HR-MAS NMR. The presence of peptidoglycan would suggest that complex biomaterials such as those from the cell walls of soil microorganisms can persist in the water environment and that it is possible that the peptidoglycan we see in DOM originally derived from microbes in soil. However, as peptidoglycan may also be produced by aquatic microbes it is not possible to definitively state the source of this material.

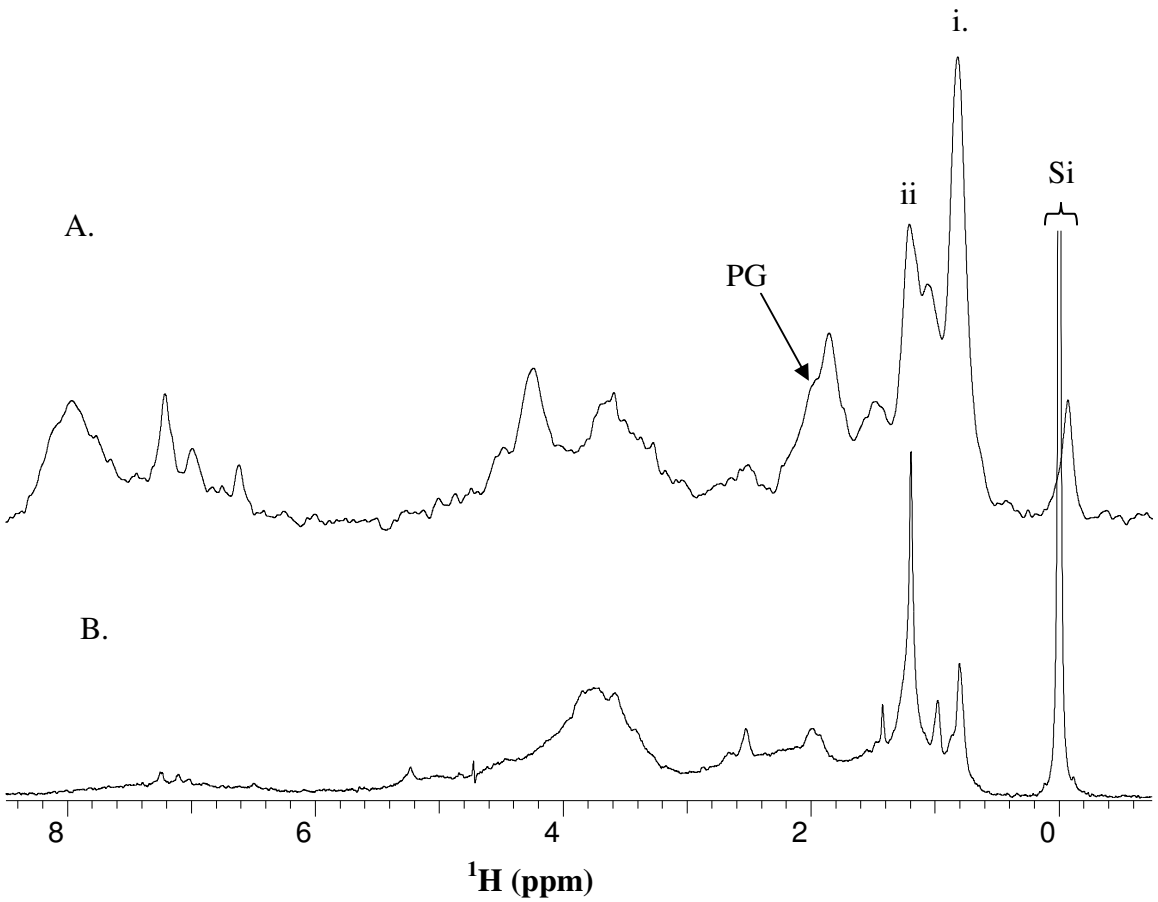

Fig. 4. (A) DE ^{1}H NMR of 14 week leachate from degraded soil microbial biomass, and (B), the DE ^{1}H NMR of the "Dromineer" DOM sample from Lough Derg. Specific assignments are: (i); CH_3, likely including resonances from aliphatic species and methylated amino acid side-chain residues in peptides/protein, (ii); aliphatic methylene (CH_2)n, (PG); peptidoglycan (Simpson *et al.*[22]) and Si indicates a natural silicate species and not TMS (a commonly used NMR reference standard).

Interestingly, natural silicate species (Si) present in DOM samples are also present in the microbial leachate spectrum. Carbon sequestration in the oceans is known to be coupled with the global cycle of silicon.[37-39] Rivers provide the conduit for 5 Tmol of silicon per year to the oceans, which is 80% of the total annual flux.[37,40] The remaining 20% comes from dust and submarine hydrothermal sources. It is thought that the ultimate source of

continental silicon flux to the oceans is weathering processes in terrestrial biogeosystems.[41,42] However, Sommer *et al*., have pointed out that silicon dynamics in terrestrial biogeosystems cannot be understood solely by way of mineral weathering.[43]

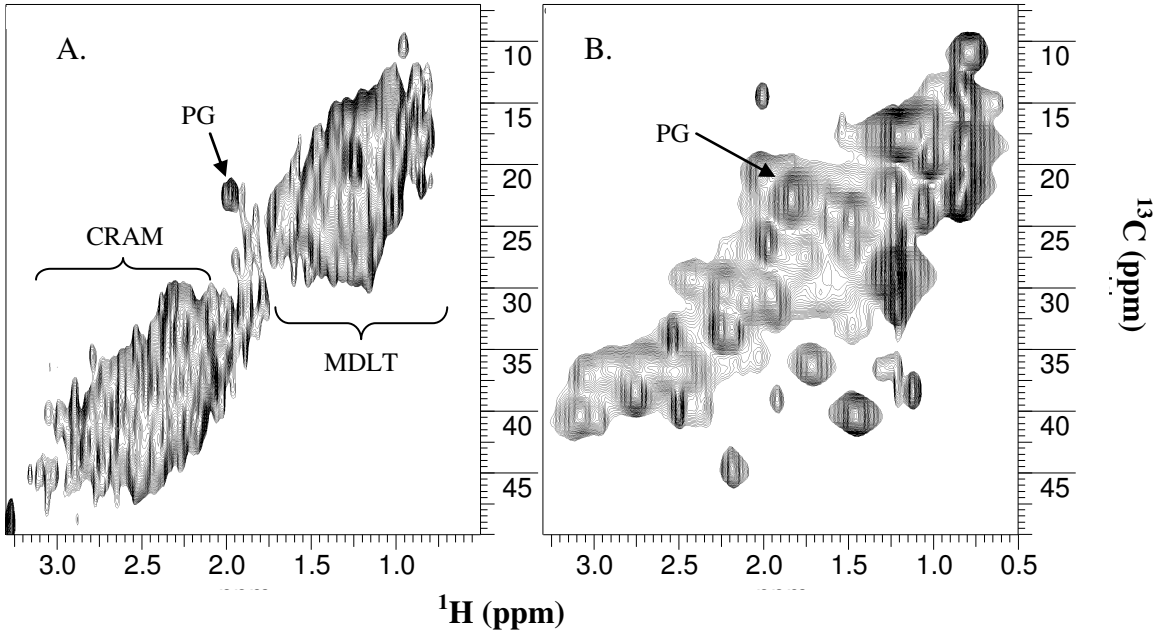

Fig. 5. HMQC of expanded aliphatic region of A. Dromineer and B. 14 week leachate. Abbreviations: CRAM, carboxyl-rich alicyclic molecules; PG, peptidoglycan; MDLT, material derived from linear terpenoids.

The silicate species in the NMR spectra are unusual and arise at around zero ppm (also present in HMQC data) and suggest methylated silica.[44] It is important to note that these signals are not from TMS, the commonly used internal standard for NMR. TMS is insoluble in water and no internal standards (of any kind) were used. Furthermore, similar signals are seen in all the natural water samples that have been analysed directly with NMR. In direct NMR, the water sample is studied "as-is", with no pre-concentration or pre-treatment of any type, indicating that these signals must be of natural origin.[45] Silicate species in the soil microbial leachate would therefore suggest that soil microorganisms accumulate their own stable silicon pools and may play a larger role in silicon cycling than presently thought.

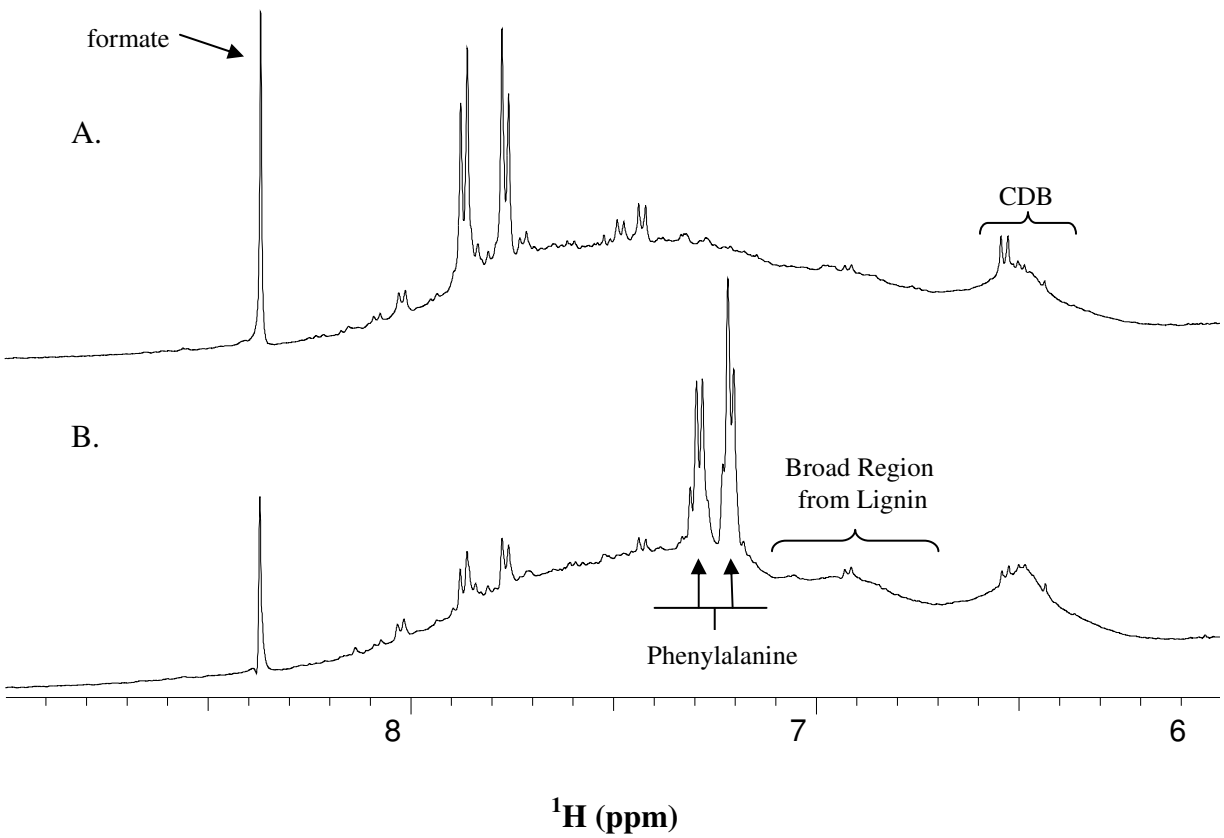

Fig. 6. ^{1}H NMR spectra for (A) Coole and (B) Dromineer samples. CDB = conjugated double bonds.

3.3. DOM variability

The spectra of samples from different sampling points in the lake that have been presented earlier have been generally similar in composition, although the presence of components such as proteins differ between samples. However, there are smaller differences between samples that may reflect how the surrounding terrestrial environment affects the distribution of DOM within the lacustrine environment. Fig. 6 displays the ^{1}H NMR spectra for the aromatic region of two sample sites in Lough Derg (Coole Bay and Dromineer). The samples display generally similar profiles and ratios of major chemical constituents. However, strong resonances that can be assigned to phenylalanine[22] in the Dromineer spectrum (and to a lesser extent Portumna and Williamstown) are not present in the Coole Bay sample. Phenylalanine is the most commonly found aromatic AA in

proteins and enzymes, is invariably present in any animal tissue and is also synthesised by common pathways in phytoplankton and bacteria. It is considered an easily degraded hydrolysable AA,[46] so its presence in some samples is of interest. Phenylalanine has been associated with increased concentrations in water of NH_4^+,[47] which in turn is a product and indicator of the presence of nitrogenous organic wastes. Dromineer is strongly influenced by the Nenagh River which passes through land utilized for agriculture and raising livestock, and also accommodates a sizable public marina. Higher phenylalanine concentrations may therefore be an indicator of elevated organic wastes from agriculture and industry. Interestingly, there appears to be little phenylalanine in the Coole Bay sample which is south of the Dromineer sampling site. This may be explained by the fact that the site is secluded, surrounded by forestry and is not fed or influenced directly by a river. However, during the sampling period from August to September; an exotic invasive species in Ireland, Zebra mussels (*Dreissena polymorpha*) were evident at highest concentrations on the eastern side of the lake at Coole Bay. The filtering activities of zebra mussels have been shown to have a large ecosystem-level influence on nitrogen cycling[48-50] and organic nitrogen concentrations decrease in water columns in microcosms with live zebra mussels.[51] It is therefore possible that the filtering activities of Zebra mussels result in recycling of larger organic nitrogen compounds such as phenylalanine. The presence of formate in both samples suggests a pathway of organic carbon degradation mainly reported for anoxic marine sediments[52] and indicates that anoxic breakdown by various microorganisms takes place in the lake. Formate and other volatile fatty acids (VFAs) are products of hydrolysis and anaerobic fermentation.[53]

A broad background hump from lignin often centered at 6.9-7.1 ppm is present in the aromatic regions of Fig. 6. The presence of lignin-type material is confirmed by the intense methoxy signal seen in the HMQC data (Fig. 3). In addition, the conjugated double bonds are likely the result of the presence of carotenoid structures known to be produced by aquatic species and present in freshwater DOM.[12] The fate of carotenoid structures is not well understood despite an estimated net annual production over 100 million tons from photosynthetic organisms alone.[54,55]

4. Conclusions

Given the influence of terrestrial organic matter on marine DOM and the similarity in the structures of both, it is challenging to assess the source of DOM and whether it is aquatic or terrestrial in origin. The findings here suggest a strong terrestrial input of recalcitrant material. Land management and human activities are important factors influencing the spatial distribution of DOM within the lacustrine environment. The input of plant material is confirmed by the presence of lignin-type signatures, while the influence of microbial biomass from either terrestrial or aquatic sources is highlighted by resonances for peptidoglycan and protein. Soil microbes may also contribute to silicon cycling through stable organo-silicon structures within the cells. The study also confirms the presence of CRAM in DOM from an Irish lake, which suggests that it may be globally ubiquitous.

Acknowledgements

The authors thank the Science Foundation of Ireland (GEOF509), the Irish Environmental Protection Agency (STRIVE program), the Geological Survey of Ireland, the Natural Science and Engineering Research Council of Canada (Discovery Grant, A.J.S) and the government of Ontario [Early Researcher Award (A.J.S)] for funding. Thank you also to the anonymous reviewers for the very helpful suggestions and criticisms.

References

[1] J. I. Hedges, R. G. Keil, R. Benner, What happens to terrestrial organic matter in the ocean? *Org*. *Geochem.* **1997**, *27*, 195.

[2] R. Benner, B. Benitez-Nelson, K. Kaiser, R. M. W. Amon, Export of young terrigenous dissolved organic carbon from rivers to the Arctic Ocean. *Geophys. Res. Lett*. **2004**, *31*, (L05305) 1.

[3] J. I. Hedges, Global biogeochemical cycles - progress and problems. *Mar. Chem.* **1992**, *39*, 67.

[4] C. T. Chiou, R. L. Malcolm, T. I. Brinton, D. E. Kile, Water solubility enhancement of some organic pollutants and pesticides by dissolved humic and fulvic-acids. *Environ*. *Sci*. *Technol.* **1986**, *20*, 502.

[5] D. Hansell, C. Carlson, *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, New York, **2002**, p. 774.

[6] T. Dittmar, J. A. Paeng, Heat-induced molecular signature in marine dissolved organic matter. *Nat*. *Geosci*. **2009,** *2*, 175.

[7] R. Stone, The invisible hand behind a vast carbon reservoir. *Science* **2010,** 328, *5985*, 1476.

[8] L. A. Cardoza, A. K. Korir, W. H. Otto, C. J. Wurrey, C. K. Larive, Applications of NMR spectroscopy in environmental science. *Prog. Nucl. Mag. Res. Sp*. **2004**, *45*, 209.

[9] A. J. Simpson, Multidimensional solution state NMR of humic substances: a practical guide and review. *Soil Sci.* **2001**, *166*, 795.

[10] N. Hertkorn, A. Kettrup, *Molecular level structural analysis of natural organic matter and of humic substances by multinuclear and higher dimensional NMR spectroscopy.* In: Perminova, I.V., Hertkom,N., Hatfield, K. (Eds.), *Use of Humates to Remediate Polluted Environments: From Theory to Practice.* Springer, Dordrecht, **2005,** pp. 391–435.

[11] B. P. Kelleher, A. J. Simpson Humic substances in soils: Are they really chemically distinct? *Environ. Sci. Technol.* **2006,** *40*, 4605.

[12] B. Lam, A. Baer, M. Alaee, B. Lefebvre, A. Moser, A. Williams, A. J. Simpson, Major structural components in freshwater dissolved organic matter. *Environ. Sci. Technol.* **2007**, *41*, 8240.

[13] N. Hertkorn, R. Benner, M. Frommberger, P. Schmitt-Kopplin, M. Witt, K. Kaiser, A. Kettrup, J. I. Hedges, Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **2006,** *70*, 2990.

[14] A. J. Leenheer, Comprehensive approach to preparative isolation and fractionation of dissolved organic carbon from natural waters and wastewaters. *Environ. Sci. Technol.* **1981,** *15*, 578.

[15] E. M. Thurman, R. L. Malcolm, Preparative isolation of aquatic humic substances. *Environ. Sci. Technol.* **1981,** *15*, 463.

[16] J. P. Simjouw, E. C. Minor, K. Mopper, Isolation and characterization of estuarine dissolved organic matter: comparison of ultrafiltration and C18 solid-phase extraction techniques. *Mar. Chem.* **2005,** *96*, 219.

[17] B. Lam, A. J. Simpson, Passive sampler for dissolved organic matter in freshwater environments. *Anal. Chem.* **2006,** *78*, 8194.

[18] A. J. Simpson, S. A. Brown, Purge NMR: Effective and easy solvent suppression. *J. Magn. Reson.* **2005,** *175*, 340.

[19] D. Wu, A. Chen, C. S. Johnson Jr., An improved diffusion ordered spectroscopy experiment incorporating bipolar-gradient pulses. *J. Magn. Reson. A.* **1995,** *115*, 260.

[20] P. H. Janssen, P. S. Yates, B. E. Grinton, P. M. Taylor, M. Sait, Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl. Environ. Microbiol.* **2002**, *68*, 2391.

[21] B. P. Kelleher, M. J. Simpson, A. J. Simpson, Assessing the fate and transformation of plant residues in the terrestrial environment using HR-MAS NMR spectroscopy. *Geochim. Cosmochim. Acta* **2006**, *70*, 4080.

[22] A. J. Simpson, G. Song, E. Smith, B. Lam, E. H. Novotny, M. H. B. Hayes, Unraveling the structural components of soil humin by use of solution-state Nuclear Magnetic Resonance Spectroscopy. *Environ. Sci. Technol.* **2007,** *41*, 876.

[23] B. G. Pautler, A. J. Simpson, D. J. McNally, S. F. Lamoureux, M. J. Simpson, Arctic Permafrost Active Layer Detachments Stimulate Microbial Activity and Degradation of Soil Organic Matter. *Environ. Sci. Technol.* **2010,** *44*, 4076.

[24] A. J. Simpson, W. L. Kingery, P. G. Hatcher, The identification of plant derived structures in humic materials using three dimensional NMR spectroscopy. *Environ. Sci. Technol.* **2003,** *37*, 337.

[25] A. J. Simpson, M. J. Simpson, E. Smith, B. P. Kelleher, Microbially derived inputs to soil organic matter: Are current estimates too low? *Environ. Sci. Technol.* **2007**, *41*, 8070.

[26] J. Fuhrman, Dissolved free amino acid cycling in an estuarine outflow plume. *Mar. Ecol.-Prog. Ser.* **1990,** *66*, 197.

[27] J. I. Hedges, G. Eglinton, P. G. Hatcher, D. L. Kirchman, C. Arnosti, S. Derenne, R. P. Evershed, I. Kogel-Knabner, J. W. de Leeuw, R. Littke, W. Michaelis, J. Rullkotter, The molecularly uncharacterized component of nonliving organic matter in natural environments. *Org. Geochem.* **2000,** *31*, 945.

[28] E. Tanoue, S. Nishiyama, M. Kamo, A. Tsugita, Bacterial membranes: possible source of a major dissolved protein in seawater. *Geochim. Cosmochim. Acta* **1995**, *59*, 2643.

[29] H. Ogawa, Y. Amagai, I. Koike, K. Kaiser, R. Benner, Production of refractory dissolved organic matter by bacteria. *Science* **2001,** *292*, 917.

[30] S. K. Park, N. S. Hettiarachchy, L. Were, Degradation behaviour of soy protein-wheat gluten films in simulated soil conditions. *J. Agr. Food Chem.* **2000,** *48*, 3027.

[31] R. A. Herman, J. D. Wolt, W. R. Halliday, Rapid degradation of the Cry1F insecticidal crystal protein in soil. *J. Agr. Food Chem.* **2002,** *50*, 7076.

[32] A. S. Waksman, K. R. N. Iyer, Contribution to our knowledge of the chemical nature and origin of humus. *Soil Sci.* **1932**, *36*, 69.

[33] K. M. Holtman, H.-M. Chang, H. Jameel, J. F. Kadla, Elucidation of Lignin Structure through Degradative Methods: Comparison of Modified DFRC and Thioacidolysis. *J. Agr. Food Chem.* **2003,** *51*, 3535.

[34] A. P. Deshmukh, A. J. Simpson, P. G. Hatcher, Evidence for cross-linking in tomato cutin using HR-MAS NMR spectroscopy. *Phytochemistry* **2003,** *64*, 1163.

[35] A. J. Simpson, M. J. Simpson, W. L. Kingery, B. A. Lefebvre, A. Moser, A. J. Williams, M. Kvasha, B. P. Kelleher, The application of ^{1}H high-resolution magic-angle spinning NMR for the study of clay-organic associations in natural and synthetic complexes. *Langmuir* **2006,** *22*, 4498.

[36] H. W. Zhang, D. W. Niesel, J. W. Peterson, G. R. Klimpel, Lipoprotein release by bacteria: Potential factor in bacterial pathogenesis. *Infect. Immun*. **1998,** *66*, 5196.

[37] P. Tréguer, D. M. Nelson, A. J. van Bennekorn, D. J. DeMaster, A. Leynaert, B. Quéguiner, The silica balance in the world ocean: a re-estimate. *Science* **1995,** *268*, 375.

[38] O. Ragueneau, P. Treguer, A. Leynaert, R. F. Anderson, M. A. Brzezinski, D. J. De Master, R. C. Dugdale, J. Dymond, G. Fische, R. Francois, C. Heinze, E. Maier-Reimer, V. Martin- Jezequel, D. M. Nelson, B. Queguiner, A review of the Si cycle in the modern ocean: recent progress and missing gaps in the application of biogenic opal as paleoproductivity proxy. *Global Planetary Change* **2000,** *26*, 317.

[39] A. Yool, T. Tyrrell, Role of diatoms in regulating the ocean's silicon cycle. *Global Biogeochem. Cy.* **2003,** *17*, 14.

[40] D. J. Conley, Terrestrial ecosystems and the global biogeochemical silica cycle. *Global Biogeochem. Cy*. **2002,** *16*, 1121.

[41] P. W. Birkeland, *Soils and Geomorphology.* 3rd Edn.. Oxford University Press, New York, **1999.**

[42] N. van Breemen, P. Buurman, *Soil Formation*. Kluwer Academic Press, Dordrecht, **2002.**

[43] M. Sommer, D. Kaczorek, Y. Kuzyakov, J. Breuer, Silicon pools and fluxes in soils and landscapes: a review. *J. Plant Nutr. Soil Sci.* **2006,** *169*, 310.

[44] R. Brindle, M. Punch, K. Albert, H MAS NMR spectroscopy of chemically modified silica gels: a fast method to characterize stationary interphases for chromatography. *Solid State Nucl. Mag.* **1996**, *6*, 251.

[45] B. Lam, A. J. Simpson, Direct ^{1}H NMR spectroscopy of dissolved organic matter in natural waters. *The Analyst* **2008,** *133*, 263.

[46] Y. Yamashita, E. Tanoue, Distribution and alteration of amino acids in bulk DOM along a transect from bay to oceanic waters. Mar. Chem. **2003,** *82*, 145.

[47] A. B. Jones, W. C. Dennison, G. R. Stmart, Macroalgal responses to nitrogen source and availability: Amino acid metabolic profiling as a bioindicator using Gracilaria edulis (Rhodophyta). *J. Phycol.* **1996,** *32*, 757.

[48] E. Mellina, J. B. Rasmussen, E. L. Mills, Impact of zebra mussel (Dreissena polymorpha) on phosphorus cycling and chlorophyll in lakes. *Can. J. Fish. Aquat. Sci*. **1995,** *52*, 2553.

[49] D. L. Arnott, M. J. Vanni, Nitrogen and phosphorus recycling by the zebra mussel (Dreissena polymorpha) in the western basin of Lake Erie. *Can. J. Fish. Aquat. Sci.* **1996**, *53*, 646.

[50] J. D. Conroy, W. J. Edwards, R. A. Pontius, D. D. Kane, H. Zhang, J. F. Shea, J. N. Richey, D. A. Culver, Soluble nitrogen and phosphorus excretion of exotic freshwater mussels (Dreissena spp.): potential impacts for nutrient remineralisation in western Lake Erie. *Freshwater Biol.* **2005,** *50*, 1146.

[51] O. Bykova, A. Laursen, V. Bostan, J. Bautista, L. McCarthy, Do zebra mussels (Dreissena polymorpha) alter lake water chemistry in a way that favours Microcystis growth? *Sci. Total Environ.* **2006**, *371*, 362.

[52] D. J. Burdige, Sediment pore waters. In: Hansell, D.A.,Carlson, C.A.(Eds.), *Biogeochemistry of Marine Dissolved Organic Matter*. Elsevier Science **2002,** *13*, pp. 612.

[53] K. Mopper, D. J. Kieber, Distribution and biological turnover of dissolved organic compounds in the water column of the Black Sea. *Deep-Sea Res.* **1991**, *38* (Suppl. 2), S1021.

[54] J. W. Louda, L. Liu, E. W. Baker, Senescence- and death-related alteration of chlorophylls and carotenoids in marine phytoplankton. *Org. Geochem*. **2002**, *33*, 12, 1635.

[55] T. Matsuno, New structures of carotenoids in marine animals. *Pure Appl. Chem.* **1985**, *57,* 5, 659.

Accessory material

Figure 1. Chemical shift prediction and 2D ^{1}H-^{13}C spectral simulation of the backbone structure of peptidoglycan. PEP indicates the peptide branches which are not included in the prediction/simulation. Note the chemical shifts of the carbohydrate units should be considered as rough approximations only as the chemical shifts of carbohydrates varies considerably with solution conditions (pH, concentration, salt background etc.). The main purpose of the simulation is to demonstrate the strong CH_3-(C=O)-N resonance which is characteristic of peptidoglycan and is highlighted with a red oval on the Figure. Spectral predictions were carried out using Advanced Chemistry Development's ACD/SpecManager and ACD/2D NMR Predictor using Neural Network Prediction algorithms (version 12.01) and water as the solvent. Parameters used for prediction including spectral resolution, and base frequency were chosen to match those of the real datasets as closely as possible.