Modelling of natural organic matter affinity for mackinawite, FeS, based on FTIR spectra by partial least squares regression (PLSR)

Alexandre Tétrault

A Thesis
In
The Department
of
Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Science (Chemistry) at

Concordia University
Montréal, Québec, Canada

December 2021

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Alexandre Tétrault

Entitled: Modelling of natural organic matter affinity for mackinawite, FeS, based on FTIR spectra by partial least squares regression (PLSR)

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Chemistry)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair

Dr. Xavier Ottenwaelder

_____ Examiner

Dr. Gregor Kos

_____ Examiner

Dr. Cameron Skinner

_____ Supervisor

Dr. Yves Gélinas

Approved by _____

Dr. Yves Gélinas, Graduate Program Director

_____2021 _____

Dr. Pascale Sicotte, Dean of Arts and Science

ABSTRACT

Modelling of natural organic matter affinity for mackinawite, FeS, based on FTIR spectra by partial least squares regression (PLSR)

Alexandre Tétrault

Marine sediments represent the most important sink for natural organic matter (NOM) across geological time spans, in which carbon-containing molecules are sequestered away and can escape remineralization to CO_2 by microbial degradation. Strong associations between iron oxide minerals and organic matter reaching the seafloor play a fundamental role in this preservation and have been known for some decades. Despite the importance of this protective mechanism in the balances of the global carbon budget, very little is known about the affinity of NOM for reduced iron species such as mackinawite (FeS) in the anoxic layers of sediment. In this study, equilibrium partition coefficients (K_d) for three types of NOM (soil leachate, corn leaf and plankton lysate) on FeS were determined through batch sorption experiments. Attenuated Total Reflectance Fourier Transform Infrared spectra (ATR-FTIR) of the organic matter was then used as model inputs to train a partial least squares regression model (PLSR) to quantitatively predict K_d values based on the FTIR spectra of NOM. The final model fit the training data with an R^2 of 0.97 (n = 17) and the validation data with a Q^2 of 0.98 (n = 5) and a RMSEP of 0.036. Inspection of the PLSR regression coefficients indicate that functional groups characteristic of polysaccharides are the greatest positive predictors of NOM sorption onto FeS at sediment pore water pH. To our knowledge, this research presents a novel machine-learning approach to the quantitative modelling of NOM sorption to minerals found in marine environments.

Acknowledgements

I would like to convey my heartfelt gratitude to my supervisor, Dr. Yves Gélinas, who guided me throughout this challenging process and helped me become both a better scientist and a better father.

I would also like to thank previous grad student Kathryn Balind for laying the groundwork for this project on FeS and initiating me to the ways of the glove box. Thank you to Dr. Gregor Kos for providing helpful feedback on computer modelling techniques, and to Peter Liu for running preliminary H-NMR analyses.

I would like to thank my partner, Jeanne Dagenais-Lespérance, as well as my family for their continuing emotional support, as well as thanking my lab colleagues Anic Imfeld, Frederic Leone, Léticia Dupont, Navjote Kooner, Frederic Séguin and Robert Harutyunyan for making the lab such an inspiring and enjoyable place to learn.

Finally I would like to thank the Fonds de recherche du Québec - Nature et technologie, and the Natural Sciences and Engineering Research Council, Canada, for helping to fund this research.

Contributions of Authors

This manuscript was written by Alexandre Tétrault and edited by Dr. Yves Gélinas. Extractions of soil, corn and plankton organic matter for use in stock solutions were preformed jointly by Alexandre Tétrault and Kathryn Balind. All remaining experimental work and analyses were carried out by Alexandre Tétrault.

Table of Contents

Table of Contents cont'd

List of Figures

List of Figures cont'd

List of Tables

List of Abbreviations

ATR-FTIR	Attenuated total reflectance Fourier transform infrared spectroscopy
DOC	Dissolved organic carbon
DOM	Dissolved organic matter
FT-ICR-MS	Fourier transform ion cyclotron resonance mass spectrometry
HMW	High molecular weight (>1 kDa)
K_d	Partition coefficient
LMW	Low molecular weight (< 1 kDa)
LV	Latent variable
MLR	Multiple linear regression
NOM	Natural organic matter
OC	Organic carbon
OM	Organic matter
PLSR	Partial least squares regression
PZC	Point of zero charge
RMSE/CV/P	Root mean square error/ of cross-validation/ of prediction
RSD	Relative standard deviation
SSA	Specific surface area
SWI	Sediment-water interface
TOC	Total organic carbon

Chapter 1: Introduction

1.1 Carbon sequestration and marine sediments

The longest chapter of the global carbon cycle begins at the ocean floor. Starting within the muds and clays of ocean sediments and the continental shelf, organic carbon can be sequestered away from microbial degradation and begin a transformation of millions - if not tens of millions - of years, into a rock-like form known as kerogen (Hedges and Oades, 1997). Carbon sequestration is of significant importance to our present epoch in which the burning of fossil fuels and the use of petrochemicals are considered a defining feature (Falkowski et al., 2000). Although it is not the most potent, carbon dioxide (CO_2) is by far the most important greenhouse gas owing to its sheer abundance and rate of release into our atmosphere from human activities. For photosynthesizers such as plants and green algae, CO_2 is the building block for complex organic molecules such as sugars, proteins and lipids. A simple linear molecule, CO_2 is also the stable end-result of both carbon combustion and biological respiration. It floats around contentedly in our atmosphere, blanketing our planet. Two of its vibrational modes - the asymmetric C-O stretch and the bend - are readily excited by the absorption of infrared (IR) radiation and are said to be IR active. When IR radiation bounces off the Earth's surface and encounters CO_2 molecules in our atmosphere, these molecules can absorb the radiation and then re-emit it back to Earth. This phenomenon is known as the greenhouse effect and is largely responsible for the climate change our planet is presently experiencing. The rate at which our climate is changing is proving to have disastrous consequences for life on Earth (IPCC, 2021), and climate change mitigation should be the number one priority of the 21st century. The modelling of organic matter (OM) sorption affinities for reduced sedimentary iron minerals, as proposed by this research, is an important tool for such mitigation. A better conceptual understanding of the processes involved in organic carbon burial will aid in our ability to model and predict environmental changes and prepare us for the future.

1.2 The global carbon cycle

Since CO_2, as well as methane (CH_4) are such critical components of climate change, it is important to understand the system in which they exist. The global carbon cycle describes the

trajectory of carbon: its sources, transformations and sinks, both spatially and temporally across the planet. Carbon can take on many forms: as the backbone of organic molecules where it provides the characteristic C-H bond, in inorganic molecules such as the abundant carbonate rocks in the Earth's crust and as several elemental allotropes (e.g. graphite and diamond). A brief introduction is given to better understand and appreciate the role of sediments in carbon sequestration (see Figure 1.1).

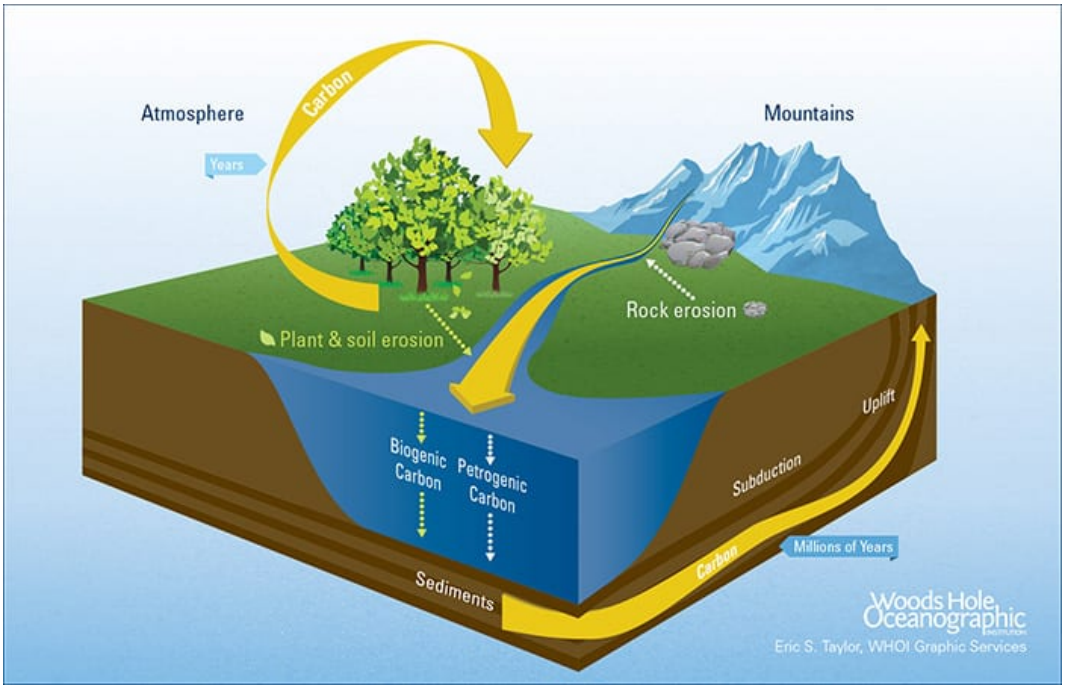

Figure 1.1. Major reservoirs in the global carbon cycle, with approximate turnover times. Figure from Woods Hole Oceanographic Institute (URL: https://www.whoi.edu/oceanus/feature/carbon-cycle/).

1.2.1 Inorganic carbon

Inorganic carbon represents by far the largest pool of carbon on Earth. Roughly 40 petagrams ($x10^{15}$ g) of carbon exists as dissolved CO_2 and bicarbonate in seawater, a reservoir that is 50 times greater than the atmosphere (Hedges and Keil, 1995; Falkowski et al., 2000). This reservoir is in equilibrium with the atmosphere and exchange occurs in both directions, following the reaction:

$$(1)\ CO_2 + H_2O \leftrightarrow H^+ + HCO_3^-$$

CO_2 is also withdrawn from the atmosphere via the weathering of silicate rocks, leading to the formation of solid calcium and magnesium carbonates which can persist for millions of years (Berner, 2003). Certain species of zooplankton build their shells from carbonates and thus also contribute to the

sequestration of carbon in inorganic forms in the sediment and, eventually, in the Earth's crust (Falkowski et al., 2000). Both these processes are reversible given the right conditions, such as thermal decomposition following deep burial and acidic dissolution (Berner, 2003). The increase in ocean acidity caused by atmospheric CO_2 is already driving the equilibrium towards carbonate dissolution with catastrophic consequences for marine life (Gao et al., 2019).

1.2.2 Organic carbon

The organic carbon (OC) cycle is characterized by the fixation of CO_2 into organic molecules via photosynthesis by primary producers (plants, phytoplankton), and the generation of CO_2 by aerobic respiration. The general formula for photosynthesis is given as:

$$(2)\ CO_2 + H_2O \rightarrow CH_2O + O_2$$

Anaerobic respiration by bacteria, not discussed in this work, is also an important consideration as it is the principal source of methane, an extremely potent greenhouse gas (IPCC, 2021). While the organic pool is smaller than that of inorganic carbon, it is extremely important for several reasons. The fossil fuels (oil, gas, coal) on which our societies are presently dependent are one result of the organic carbon cycle and originate from buried OM (Berner, 2003). Their extraction and combustion is responsible for the net increase in CO_2 in the Earth's atmosphere and the main driver of climate change.
Photosynthesis is an active means by which CO_2 is removed from the atmosphere and also sustains virtually all life on Earth. A fraction of the total organic matter into which CO_2 is fixed may wind up buried in marine sediments and ultimately removed from the active CO_2 pool, thus demonstrating the importance of marine sediments in long-term carbon sequestration (Hedges and Keil, 1995; Burdige, 2007).

1.3 Sources of organic carbon to marine sediments

1.3.1 Terrestrial organic carbon

Photosynthesis can occur in both terrestrial and aquatic ecosystems. Multicellular land plants require structural molecules to maintain their shape and support growth. Refractory macromolecules such as lignin, cutin and cellulose play these roles. These molecules contain stable C-C and ether

bonds (lignin, cellulose) and long aliphatic chains (cutin) which can prove very resistant to microbial degradation (Arndt et al., 2013). Much of this terrestrial biomass will remain on land and will provide short-term carbon sequestration, in forests for example. The total residence time of carbon sequestered in terrestrial sinks is only on the order of a few decades to at most a few centuries before being respired back into CO_2 by bacteria and fungi (Falkowski et al., 2000; Lützow et al., 2006) and is therefore not sufficient on its own to compensate for CO_2 accumulation in our atmosphere. However, it is estimated that 0.38-0.45 Pg (10^{15} g) of this terrestrially-derived OC is transported annually by rivers to marine environments (e.g. estuaries) before heading to continental margins (Meybeck, 1982; Bauer et al., 2013).

Terrestrial OM that gets transported to marine systems may be pre-aged in soil and during the transport process and can be significantly degraded, particularly in the dissolved organic carbon (DOC) fraction. Decades may pass between the initial fixation of this carbon in photosynthetic plants and its eventual export in rivers, which may include a stage of association to minerals in soils (Hedges et al., 2000). Also during this period, the more labile (easily-degraded) compounds will have been largely consumed by microorganisms leaving behind a component that is rich in large, refractory macromolecules to be carried to marine systems (Vannote et al., 1980; Hedges et al., 2000; Arndt et al., 2013). Importantly, a large fraction of the total organic carbon (TOC) in suspension at the mouths of rivers may already be associated to minerals (Hedges et al., 2000) before arriving to marine environments.

1.3.2 Marine organic carbon

Marine phytoplankton are responsible for primary production in the world's oceans (for a full review, see Sigman and Hain, 2012). They synthesize a higher proportion of proteins than terrestrial plants and are thus especially rich in amino acids, as well as lipids and carbohydrates (Burdige, 2007). The high nitrogen and phosphorous contents of phytoplankton also make them important food sources for aquatic life. Many species of phytoplankton are unicellular and therefore have no need for resistant structural molecules like cellulose, meanwhile the abundant peptide bonds in marine OM are easily degraded by consumers. Fresh phytoplankton biomass therefore represents the labile end-member of the OM spectrum in marine environments, while heavily-degraded humic substances arriving from land

represent the refractory end-member (Burdige, 2007). It should be noted that despite this generalization, certain species of algae also synthesize small proportions of highly refractory macromolecules called algeanans which contain hard-to-degrade aliphatic chains with C-C and ether linkages similar to terrestrial OM (e.g. Gelin et al., 1996). Organic matter from primary production in the ocean's photic zone eventually sinks to the ocean floor. En route, the OM is remineralized (broken down, in the aerobic case into CO_2 and anaerobically to small organic acids) by feeding bacteria (see Figure 1.2). As a result, only about 1% of the primary production reaches the ocean floor (Hedges and Keil, 1995), but the huge variability in this figure globally ranges from 0.3-30% (Lutz et al. cited in Arndt et al., 2013). Mineral particles can act as ballast for OM in the water column, decreasing residence time and therefore limiting the degradation that can occur before reaching the sediment-water interface (SWI) (Armstrong et al., 2001). As previously mentioned, much of the particulate OM reaching marine environments from land is already associated to minerals (Hedges et al., 2000), which may further decrease the reactivity of this fraction.

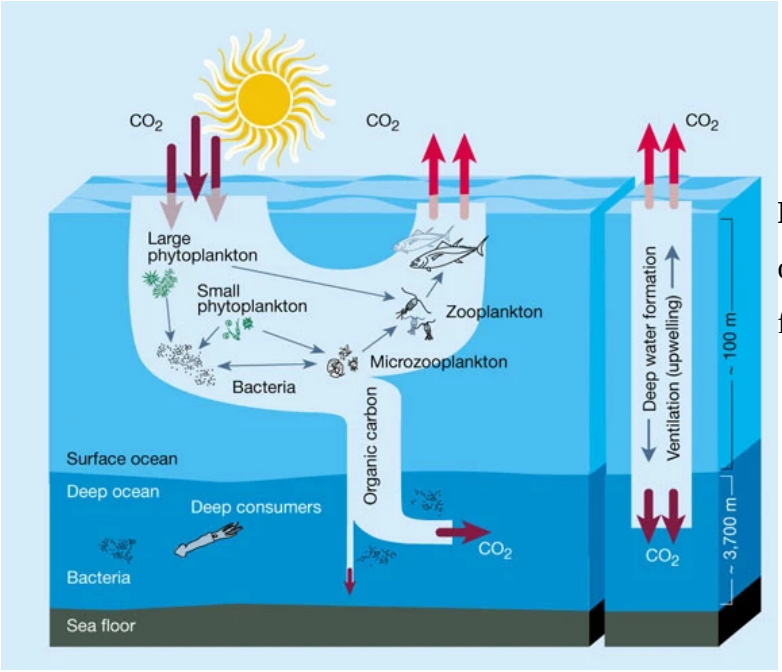

Figure 1.2. Transformations of OM and fluxes of CO_2 throughout the water column. Image from Chisholm (2000).

1.4 Marine sediments as carbon sink

Diagenesis is the process whereby sediments, under the influence of increasing pressure and temperature during burial, slowly transform into sedimentary rock in the Earth's crust. This process

equally affects the organic matter associated to sediments and is responsible for the formation of kerogen (also: oil, gas and coal), a fossilized form of OM that may reside in the Earth's crust for millions of years (Berner, 2003). One example of diagenesis is sulfurization, a process whereby sulfur atoms replace reactive functional groups in OM, making them less accessible to bacterial degradation (LaRowe et al., 2020 and sources therein). As a result of diagenesis, marine sediments are the fundamental sink for OM on Earth - the one location where burial of OM will ultimately remove CO_2 from the atmosphere and sequester carbon into the lithosphere on geological time scales (Garrels et al., 1976; Hedges and Keil, 1995). Surface sediments alone are estimated to account for the storage of 150 Pg of OC (Emerson and Hedges cited in Hedges and Keil, 1995).

1.4.1 Physical protection

Over the past three decades, it has been widely observed that the majority of organic matter in sediments (up to 90%) is directly bound to mineral phases (Keil et al., 1994; Hedges and Keil, 1995; Ransom et al., 1998) while similar associations between OM and minerals have been well-documented in soil systems (e.g. Kaiser and Guggenberger, 2000). Lalonde et al. (2012) determined that 21.5 ± 8.6% of OM in sediments is directly bound to reactive iron mineral phases. The persistence of OM sorbed to minerals indicates that minerals help to protect OM from microbial respiration, a phenomenon known as physical protection. Physical protection may impede microbial degradation via three principal mechanisms: steric hindrance (inaccessible target sites owing to molecules rigidly attached to mineral surfaces), encapsulation (within sheet clay minerals or within amorphous blebs as coprecipitated organic-inorganic material), and the shielding of labile functional groups through direct metal complexation (see Burdige, 2007). In contrast, OM desorbed from sediment minerals in a laboratory experiment was rapidly remineralized by microbes (Keil et al., 1994) despite having been otherwise stable for hundreds of years. Recently, researchers have proposed that proteinaceous material, a particularly labile OM pool, routinely escapes complete remineralization and accumulates in oxic marine sediments for millions of years due to sorption onto mineral surfaces (Estes et al., 2019). It appears that both the strength/type of mineral-OM interaction as well as the associated decrease in desorbability are important factors in promoting the stabilization of OM (Mikutta et al., 2007), more so even than the nature (labile vs. recalcitrant) of the organic matter itself.

1.5 Marine sediment chemistry

1.5.1 Redox depth profile

Marine sediments provide a unique chemical environment both in the pore waters and the solid phase, that is continually transforming with increasing depth below the SWI. One of the most important features of sediment chemistry is the changing redox profile with depth as electron acceptors are gradually used up by microbial respiration (Figure 1.3). At the top of the sediment column near the SWI, oxygen dissolved in the bottom waters penetrates into sediments and provides the strongest and most energetically-favourable electron acceptor. As sediment depth increases, however, oxygen is limited by diffusion and may quickly become consumed at the same rate as it is replenished. The transition to fully anoxic conditions (i.e. $[O_{2(aq)}] < 10^{-6}$ M) may occur as deep in the sediment as several meters in the case of deep ocean environments to a mere couple of millimetres below the SWI in high-productivity coastal margins (Rickard and Luther, 2007).

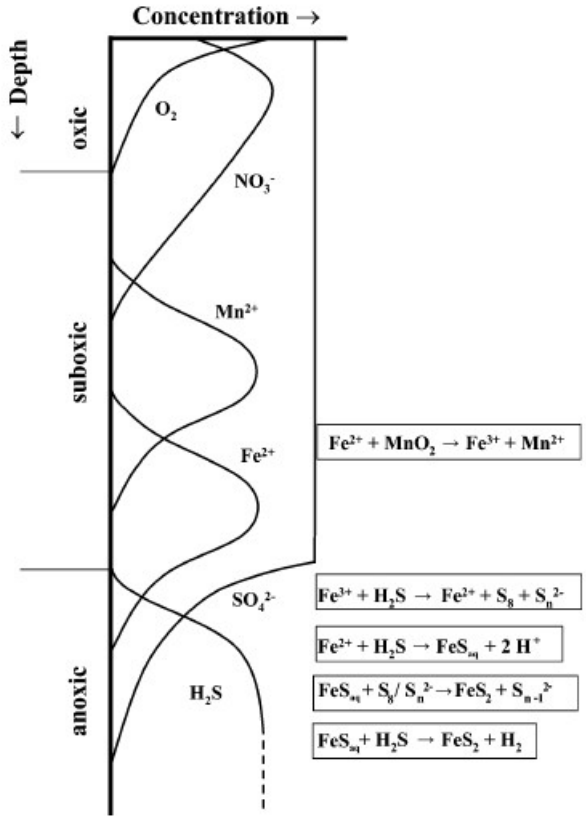

Figure 1.3. The order of redox-active dissolved species with depth in the sediment column. Higher-energy electron acceptors are depleted first. Figure from Rickard and Luther (2007).

The order of electron acceptors used for microbial respiration follows roughly from the decreasing yields in energy as the more desirable oxidizing agents are consumed first, along the lines of $O_2 > NO_3^- >$ Mn(IV) $>$ Fe(III) $> SO_4^{2-}$ (LaRowe et al., 2020). A list of the major redox half reactions found in groundwater and marine environments, along with associated reduction potentials, is given in Figure 1.4. It should be noted however that seasonal mixed redox cycles and bioturbation of sediments by benthic organisms may complicate this sequence considerably, for example

by the creation of heterogeneous micro-environments in the sediment where redox potential does not follow a simple diffusion gradient (Rickard and Luther, 2007). Nevertheless, this sequence of electron acceptors has a substantial impact on the pore water chemical profile as well as the minerology of sediments, including the speciation of iron and sulfur species. Climate change and human land-use are causing marked changes to the redox chemistry of many marine environments by way of oxygen depletion, causing the formation of O_2 "dead zones" and anoxia to set in earlier in the sediment column (Gilbert et al., 2005; Diaz and Rosenberg, 2008). These changes are likely to have significant consequences for the speciation of iron minerals within sediments and the OM associated to them.

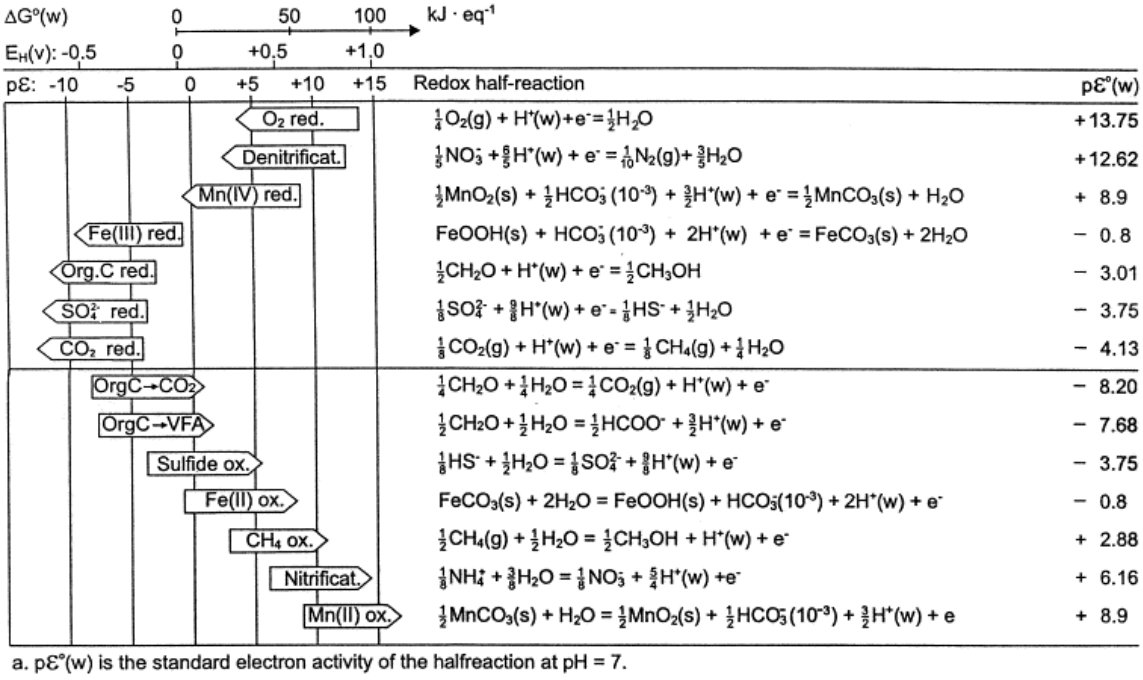

Figure 1.4. Major redox half reactions in groundwater and marine environments. Image from Christensen et al. (2000).

1.5.2 Major dissolved Fe species in soils and sediments

In soils and sediments, variations to pore water redox potentials (oxic vs. anoxic) cause iron minerals to cycle between ferric (Fe^{3+}) and ferrous (Fe^{2+}) states. This cycle has been variously called the iron shuttle (Barber et al., 2014) or redox-wheel (Li et al., 2012). Fe^{3+} is highly insoluble in oxic waters and rapidly precipitates out as iron oxides, initially as poorly-crystalline ferrihydrite $Fe(OH)_3$. When iron oxides come into contact with inorganic reducing agents such as microbially-produced

sulfide (S^{2-}) and Fe^{2+}, reduction of Fe^{3+} to soluble Fe^{2+} in the crystal structure causes dissolution of the mineral (Dos Santos Afonso and Stumm, 1992; Pedersen et al., 2005). A common phenomenon in marine sediments is the direct microbial reduction of solid-phase iron oxides by dissimilatory iron reducing bacteria such as *Shewanella* and *Geobacter* (Lovley and Anderson, 2000). This process also cycles Fe^{3+} to Fe^{2+}.

Fe^{2+} is considerably more soluble in anoxic waters than Fe^{3+}. In pore waters with total sulfide concentrations in the micromolar range, $[Fe^{2+}]_{Total}$ can reach concentrations of 1 µM - predominantly as $FeS_{(aq)}$ - up to three orders of magnitude higher than total dissolved iron concentrations in oxic pore waters (Rickard and Luther, 2007). The comparatively high solubility of Fe^{2+} in anoxic waters allows dissolved iron to recirculate back into oxic regions where contact with dissolved O_2 will cause oxidation and precipitation to occur once more. Multiple cycles of this dissolution/recrystallization event cause an overall increase in mineral crystallinity, likely through a process of Ostwald ripening, and shifts the dominant iron oxide forms away from ferrihydrite and towards more stable forms such as lepidocrocite (γ-FeOOH), goethite (α-FeOOH) and magnetite (Fe_3O_4) (Pedersen et al., 2005; Hansel et al., 2005).

1.5.3 Mackinawite

In alkaline seawater, the predominant dissolved sulfide species in anoxic sediment pore waters are $HS^-_{(aq)}$ and $H_2S_{(aq)}$,which build up due to the microbial reduction of sulfate (Rickard and Luther, 2007). Aqueous FeS^0 clusters form and rapidly precipitate out of solution as mackinawite or FeS_m (intrinsic solubility: $\log K_0 = -5.7$) (Rickard, 2006) . We are selecting mackinawite as our model reduced iron sulfide species as it is the first to precipitate out of solution in low-temperature environments and is commonly formed by sulfate-reducing prokaryotes (Rickard and Luther, 2007; Picard et al., 2016). Lennie et al. (1995) described mackinawite as having a tetragonal unit cell in which each iron atom is coordinated to four neighbouring sulfurs while also exhibiting Fe-Fe bonding in a square planar coordination in the (001) plane (Figure 1.5). The extended mackinawite crystal consists of sheets stacked along the axis normal to the (001) plane and held together by van der Waals forces (Rickard and Luther, 2007).

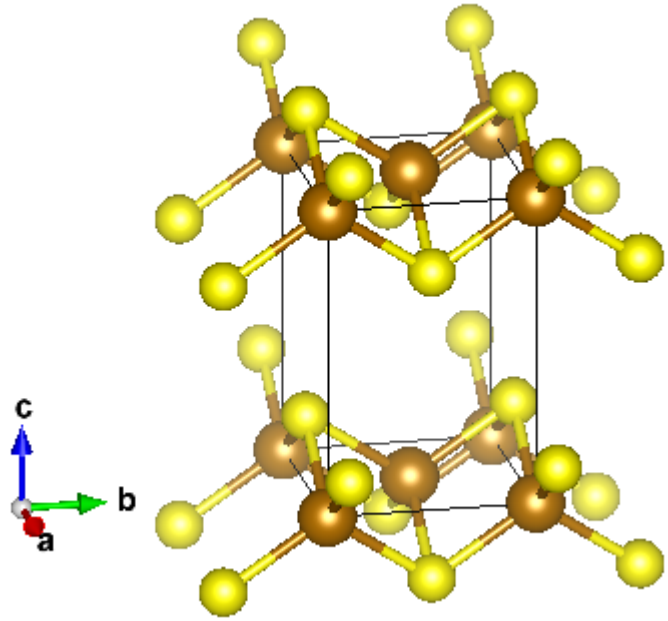

Figure 1.5. The tetragonal mackinawite unit cell. Yellow spheres: sulfur atoms, brown spheres: iron atoms. The *c*-axis is normal to the (001) plane.

Mackinawite produced in sediments is likely to be poorly ordered / nanocrystalline, showing only broad and poorly-defined peaks by X-ray diffraction associated to this stacking (Rickard and Luther, 2007), but is more amenable to study by Raman spectroscopy (Boughriet et al., 1997; Bourdoiseau et al., 2008). Density functional theory calculations have confirmed that the (001) surface is the most stable face of mackinawite and promotes the growth of thin, platy crystals that stack along the *c*-axis (Devey et al., 2008). Few studies of the surface chemistry of mackinawite in aqueous solution have been undertaken (Wittekindt and Marx, 2012), however an important study involving the potentiometric titration of mackinawite determined the point of zero charge (PZC) to be around 7.5 (Wolthers et al., 2005), slightly below the typical pH of 7.8 found in marine pore waters. Based on the higher affinity of aqueous Fe^{2+} for HS^- over OH^-, Wolthers et al. (2005) concluded that iron surface sites in FeS-saturated waters would be of the form $\equiv FeSH^0$ instead of $\equiv FeOH^0$ (the latter being the case for goethite). Crystal analysis led them to propose two surface sites: a weakly-acidic tricoordinated sulfur and a strongly-acidic monocoordinated sulfur (Figure 1.6).

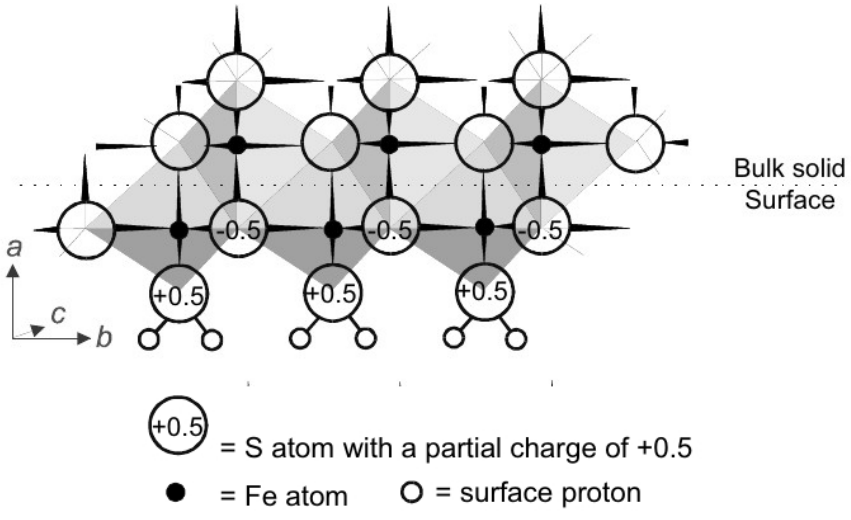

Figure 1.6. Depiction of two different sulfur surface sites of the mackinawite crystal. Figure from Wolthers et al. (2005).

1.6 Batch sorption experiment

To study the mobility or retention of chemical species through a system, researchers typically build simplified models of the environmental system of interest and measure adsorption parameters by way of flow-through or batch sorption experiments. This is the case for example in the study of contaminant transport for environmental assessment studies (Krupka et al., 1999). The same physico-chemical principles apply when studying the cycling of organic carbon. The batch sorption experiment has received extensive use as a method for probing sorption affinities and mechanisms for a variety of organic molecules onto mineral surfaces such as iron oxides in order to better understand carbon cycling through soil and marine environments (Gu et al., 1994; Boily et al., 2000; Chen et al., 2014; Coward et al., 2019). Single compounds have been employed, for example oxalate and malonate (Persson and Axe, 2005; Axe et al., 2006), phthalate (Tejedor-Tejedor et al., 1992) and salicylate (Yost et al., 1990; Thomas and Kelley, 2008; Yang et al., 2013), as representatives of typical functional groups found in the heterogeneous OM of pore waters. Through subsequent FTIR analyses, Yost et al. (1990) determined that inner-sphere ligand exchange between the carboxyl/hydroxyl functional groups of salicylate were the predominant mechanism of sorption onto the goethite surface. Other researchers have used natural OM (NOM) from various sources to arrive at similar conclusions regarding the importance of ligand exchange, for example on the surface of hematite (Fe_2O_3) (Gu et al., 1994). Arnarson & Keil (2000) performed batch sorption experiments on montmorillonite clay using NOM extracted from sediments. By their estimation, ligand exchange accounted for approximately 35% of sorption interactions while van der Waals forces and cation bridging contributed 60% and 5% respectively. Sorption interactions can also be probed by exploring the differential fractionation of dissolved OM (DOM) by mineral surfaces. Electrospray ionization Fourier-transform ion cyclotron resonance mass spectrometry has also been used to analyze the residual DOM following batch sorption experiments on iron minerals to determine which types of molecules adsorb preferentially (Lv et al., 2016). Coward et al. (2019) observed temporal fractionation occurring in DOM sorbing onto the surface of goethite. In this case, aromatic and polycyclic aromatic compounds with high O/C ratios demonstrated the greatest sorptive preference for the goethite surface, followed by lignin-like compounds.

In contrast to iron oxides, very little work has been undertaken to study the sorption of OM onto mackinawite. Wang et al. (2019) performed batch sorption experiments on four different iron species including mackinawite and pyrite using NOM extracted from peat. Their results showed that, when normalized by specific surface area, mackinawite had the highest sorption capacity, above that of even hematite and ferrihydrite. Their findings suggest, similarly to Coward et al.'s (2019) study of goethite, that aromatic and oxygen-rich compounds are preferentially sorbed by mackinawite. The difficulties posed by working with an oxygen-sensitive compound such as mackinawite may be partially responsible for the lack of sorption experiments on reduced metal sulfides.

1.7 Constant partition coefficient K_d

Batch sorption experiments that cover the full range of adsorption capacity of the mineral surface will typically deviate from linearity as sorbate molecules compete for available surface sites and interact with one another. Furthermore, since adsorption is a surface-mediated phenomenon, the heterogeneity of the mineral surface plays an important role in both shaping the sorption isotherm and affecting isotherm reproducibility. A mineral crystal possesses several different surfaces of varying energies. Typically the lowest-energy surface accounts for the largest faces of the crystal but also provides the least sorption energy. In the specific case of mackinawite, the (001) face is the predominent and lowest-energy surface (Section 1.5.3). The simple constant partition coefficient model employed in this research does not account for multi-site complexity. Isotherms such as the Langmuir and Freundlich models are mathematical fits to account for the deviations in sorption linearity due to competition, surface heterogeneity (i.e. multi-site sorption) and other effects. These fits can provide empirically-derived approximations of the equilibrium partition coefficient K_d which parameterizes the strength of bonding for dissolved species adsorbing onto a solid (Gu et al., 1994; Thimsen and Keil, 1998; Feng et al., 2005). In the limit of low total dissolved species when surface available sites are still well in excess of sorbate molecules, the sorption behaviour can nevertheless be approximated as a linear relationship (see Figure 1.7) known as the constant partition coefficient model, K_d.

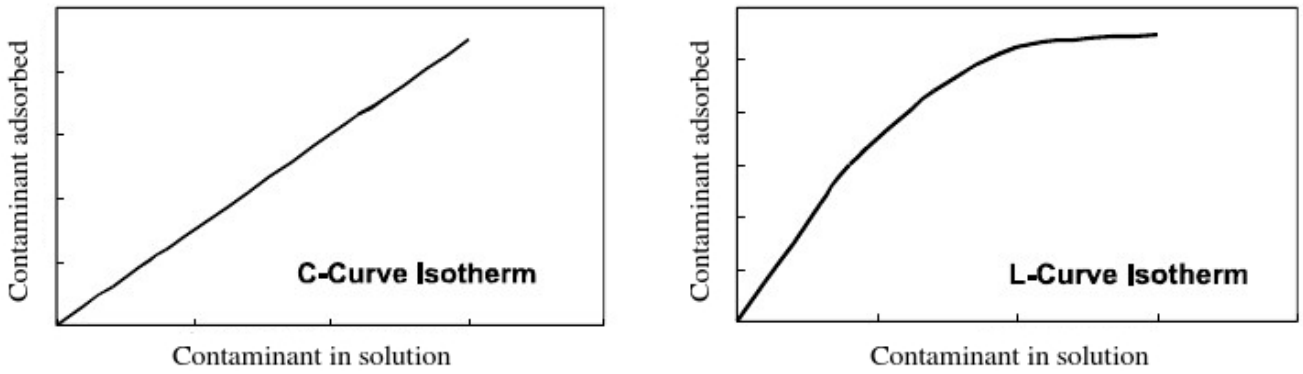

Figure 1.7. Left: Sorption is proportional to concentration and the relationship is modeled by the constant partition coefficient (K_d). Right: Sorption deviates from linearity; empirical relationship is fitted by Langmuir, Freundlich or other isotherm. Figure adapted from Krupka et al. (1999).

Krupka et al. (1999) give the equation for the conditional K_d, which deviates from the strict thermodynamic definition, as:

$$(3) \quad K_d = \frac{\sum_{n=1}^{\infty} A_i}{\sum_{n=1}^{\infty} C_i}$$

where A_i is the equilibrium concentration of OC of the i^{th} species sorbed (mg C kg^{-1} mineral) and C_i is the equilibrium concentration of DOC of the i^{th} species in solution (mg C L^{-1}). A and C are summed across all species in their respective phases and K_d takes units of L kg^{-1}. The partition coefficient K_d can be related to Freundlich constant via the equation:

$$(4) \; K_d = K_F \left(C_e\right)^{(1-n)/n}$$

where K_F is the empirically-fit Freundlich constant, C_e is the equilibrium DOC concentration, and *n* is a parameter describing the deviation from linearity of the Freundlich fit. When the fit is perfectly linear, $n = 1$ and the equation becomes independent of C_e.

Marine sediment pore water DOC concentrations for continental slopes and estuaries typically vary between 6-50 mg C L^{-1} (Alperin et al., 1999). At such low concentrations of DOC, K_d becomes an effective model for quantifying and comparing sorption affinities across diverse OM inputs and varying solid phases. A batch sorption experiment in this working range of DOC yields a linear relationship

between A and C at equilibrium where the slope of the best fit line gives the value of K_d. Numerous factors are known to influence sorption equilibria and therefore the values of K_d. A non-exhaustive list of these factors include solution pH, ionic strength and the associated activities of dissolved species, temperature and particle density (Thimsen and Keil, 1998; Krupka et al., 1999). The above parameters should therefore be well-documented in order to appropriately compare studies to discern trends. The vast majority of studies (this one included) are conducted under laboratory conditions (e.g. temperature and pressure), therefore the resulting equilibrium values pertain to those conditions and not the in-situ marine sediment pore water equilibrium values.

1.8 Fourier Transform Infrared Spectroscopy

Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR) is a simple, fast and effective method for studying the functional group distribution of OM. Molecular bonds can be excited into higher-energy vibrational modes by IR radiation if these modes possess a fluctuating dipole moment. In the harmonic oscillator model of molecular vibrations, the energy associated to these excitations is proportional to the frequency of vibration, given by:

$$(5)\ \nu = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}}$$

where ν is the frequency of the photon absorbed in wavenumbers (cm^{-1}), c is the speed of light, k is the force constant of the interatomic bond and μ is the reduced mass. Variations in the modes and bond types (stretching is more energetic than bending; double bonds more energetic than single bonds) and the nature of the atoms involved (lighter atoms will vibrate at higher frequencies) create distinct regions of absorbance for different functional groups and vibrational modes when absorbance is plotted against wavenumber. The resulting spectrum enables us to identify the functional groups that are present in complex OM mixtures and thus to get an indication of chemical reactivity. The ATR collection method permits fast and accurate measurement of liquids, solids and gasses. The sample is placed on a crystal surface and an IR beam is passed through the crystal, whereupon it is internally reflected against the crystal walls multiple times. At each reflection, a small evanescent wave escapes the crystal and is partially absorbed by the sample before reflecting back to the crystal (see Figure 1.8). This process subjects the specimen to multiple instances where excitational energy may be absorbed.

As with all FTIR techniques, it is important to ensure proper background correction for the ambient atmosphere and consistent sample homogenization prior to analysis.

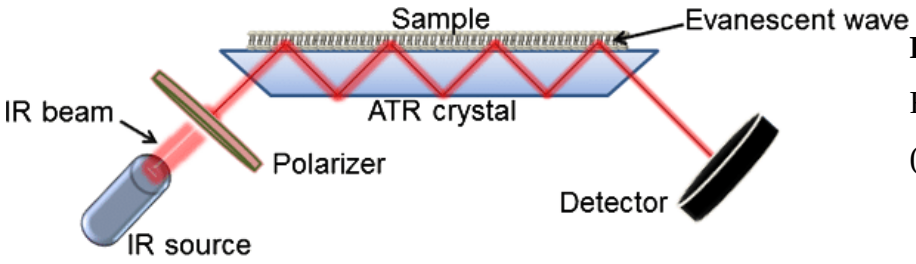

Figure 1.8. Schematic of ATR-FTIR. Image from Ausili et al. (2015).

1.9 Partial Least Squares Regression

In order to quantitatively predict sorption of NOM onto mackinawite, an appropriate numerical model must be used. Partial Least Squares (PLS) was invented by statistician Herman Wold in 1975 and then adapted for use as a multivariate regression technique (PLSR) in the field of chemometrics by his son Svante Wold, where it has since been fully embraced (Wold et al., 2001). In multivariate modelling, the normal equation gives the least squares solution to the linear equation $Y = BX$ that minimizes the error of fit, via:

$$(6)\ B=\left(X^T X\right)^{-1} X^T Y$$

where B represents the matrix containing intercept and slope of the best-fit line for a first order solution, and superscripts T and -1 denote matrix transpose and inverse respectively. Spectroscopic absorbance data as predictor (X) variables pose an issue for this approach as the X variables often contain significant noise, are usually correlated and far outnumber the instances of observations, leading to rank-deficiency (Wold et al., 2001). PLSR circumvents these issues by reducing the data X into underlying latent variables (LVs) described by scores T and loadings P such that:

$$(7)\ X=TP^T+E$$

where E is a minimized error term. The scores are weighted combinations of the original X variables calculated for each observation and limited to the number of components, i.e. LVs, needed to adequately describe the data. A variation of PLS called SIMPLS approaches the problem by first decomposing the cross-product $S = X^TY$ in order to immediately capture the greatest co-variance between predictors (X) and targets (Y) (de Jong, 1993). Benefits of the SIMPLS approach are its computational speed in comparison to other PLS algorithms, and the way in which deflation is carried

out. Scores from the a -1th component are projected out of S to ensure that subsequent components fit only the residual error, thus making them orthogonal to those preceding. By deflating the cross-product and not the X matrix directly, the weighted contributions of X to the scores, denoted by r_a, for the a^{th} component where $a \geq 1$, are directly interpretable from the original data set. The final regression coefficients are computed from the normal equation via the formula:

$$(8)\ B_{PLS} = R\left(T^T Y\right)$$

The SIMPLS approach will permit us to use FTIR absorbance data from various NOM compounds as predictors of sorption affinity onto the mackinawite surface. In this way, we will be able to determine which organic functional groups are the best predictors of sorption.

1.10 Objectives

The importance of iron minerals in OC sequestration in marine sediments cannot be overstated. However, our understanding of sorption affinities remains largely qualitative, that is to say a direct relationship between the chemical characteristics of OM inputs and a quantitative measure of sorption affinity onto sediment minerals has yet to be established, making cross-study comparisons and predictions difficult. In addition, very little research has been conducted on reduced iron sulfides despite their prevalence in anoxic marine sediments. Climate change and human land-use are increasing oxygen stress in many marine environments, bringing about a more rapid onset of anoxia in the sediment column. Our goal is to model the relationship between functional group distribution of NOM and sorption affinity for mackinawite in order to make quantitative predictions K_d. Such a tool would be of direct practical importance for the modelling of carbon cycling through marine sediments and provide a usable synthesis of the myriad studies on the characterization of OM sorbed to sediment minerals.

The following chapter contains the manuscript for this research thesis as prepared for submission to the academic journal *Geochimica et Cosmochimica Acta*. The title for this work is "Modelling of natural organic matter affinity for mackinawite, FeS, based on FTIR spectra by partial least squares regression (PLSR)", authored by Alexandre Tétrault and Yves Gélinas.

Chapter 2:
Modelling of natural organic matter affinity for mackinawite, FeS, based on FTIR spectra by partial least squares regression (PLSR)

2.1 Introduction

Marine sediments have long been established as a fundamental sink for natural organic matter (NOM), providing the one location where burial of NOM will ultimately remove CO_2 from the atmosphere via carbon sequestration into the lithosphere across geological time scales (Garrels et al., 1976; Hedges and Keil, 1995). Over the past three decades, it has been widely observed that the majority of organic matter in sediments (up to 90%) is directly bound to mineral phases (Keil and Hedges, 1993; Hedges and Keil, 1995; Ransom et al., 1998) while similar associations between NOM and minerals have been well-documented in soil systems (Kaiser and Guggenberger, 2000). The persistence of NOM sorbed to minerals indicates that minerals help to protect NOM from microbial respiration, a phenomenon known as physical protection (Keil et al., 1994). Clay minerals in marine sediments have been posited as the main drivers of organic carbon (OC) sequestration and responsible for increasing atmospheric O_2 concentrations beginning in the late Precambian (Kennedy et al., 2006). Lalonde et al. (2012) determined that 21.5 ± 8.6 % of NOM in sediments is directly bound to reactive iron mineral phases, leading to a global estimate of 19-45 petagrams (10^{15} g) of carbon sequestered in surface marine sediments via iron oxides. Emerson and Hedges (cited in Hedges and Keil, 1995) estimated that a global total of 150 petagrams of OC is preserved in marine sediments due in part to mineral interactions. Since these discoveries, the mechanisms through which sedimentary minerals sorb and protect NOM from remineralization have been extensively studied (McKnight et al., 1992; Gu et al., 1994; Filius et al., 1997; Meier et al., 1999; Riedel et al., 2013; Barber et al., 2017; Zhao et al., 2018; Coward et al., 2018).

A major difference between aluminosilicate clays and iron oxides is that the latter are redox sensitive. In soils and sediments, variations to pore water redox potentials (oxic vs. anoxic) cause iron minerals to cycle between ferric (Fe^{3+}) and ferrous (Fe^{2+}) states (Liu et al., 2012; Barber et al., 2014). At a variable depth below the sediment-water interface (SWI), oxygen diffusion from the bottom waters may fail to compensate for microbial aerobic respiration thus causing sequential switches to

microbial nitrate, ferric iron, manganese and then sulfate reduction when the sediment becomes anoxic (Rickard and Luther, 2007). The buildup of microbially-produced dissolved sulfide species in anoxic porewaters causes the concomitant precipitation of reduced iron(II) sulfide species such as mackinawite (FeS), greigeite (Fe_3S_4) and pyrite (FeS_2) - the most abundant mineral on the Earth's surface (Rickard and Luther, 2007).

Anthropogenic climate change is causing serious consequences for the Earth's systems. Human driven land-use changes and climate change are having significant impacts on carbon fluxes to marine environments and associated sedimentary carbon stocks (Atwood et al., 2020). Research from the past two decades has shown the onset of hypoxia in marine bottom waters due to increased water temperatures, changes in current circulations and eutrophication from nutrient runoff (Gilbert et al., 2005; Jutras et al., 2020). The average O_2 concentration in bottom waters and O_2 penetration depth in sediments are decreasing in many aquatic environments, causing oxygen-depleted dead zones (Diaz and Rosenberg, 2008). Given these changes, it is foreseeable that the abundance of reduced iron species in sediments is set to increase. In contrast to iron oxides, very little research has been conducted on the affinity of NOM for reduced iron sulfide species. Studying these interactions will improve our understanding of carbon sequestration in anoxic sediments and is critical to predicting the effects of hypoxia on carbon sequestration in marine and freshwater environments.

Mackinawite (FeS) is the first species formed by the rapid precipitation of dissolved Fe^{2+} and S^{2-} species (typically $HS^-_{(aq)}$) in low-temperature marine environments (Rickard and Luther, 2007) and constitutes an important mineral fraction of acid volatile sulfides (AVS) found in sediments (Rickard and Morse, 2005). For these reasons it can serve as a model reduced iron sulfide compound. According to Rickard & Morse (2005), mackinawite is metastable compared to greigite and pyrite but can nevertheless persist for geological timespans in sediments. The study of FeS and other metastable iron sulfides is rendered challenging due to their high susceptibility to O_2, making them unstable with respect to cryptic reoxidation during analyses. Furthermore, rapidly-precipitated mackinawite is typically nanocrystalline and possesses little to no defining features by powder X-ray diffraction (Rickard and Morse, 2005). As a black, light-absorbing mineral, it is not amenable to the same types of infrared studies that have been regularly performed on iron oxide-organic matter interactions (Tejedor-Tejedor et al., 1990; Gu et al., 1994; Axe et al., 2006; Yang et al., 2013; Chen et al., 2014). We believe

these challenges are in large part responsible for the relative paucity of data on interactions between NOM and reduced iron sulfides.

Batch sorption experiments provide useful data for comparative analyses of sorption affinities across changing redox regimes. This method has been used extensively for probing the sorption affinities of a variety of organic molecules onto mineral surfaces in oxic environments (Gu et al., 1994; Boily et al., 2000; Chen et al., 2014; Coward et al., 2019). The linear constant partition model is applicable at low dissolved organic carbon (DOC) concentrations (Krupka et al., 1999) and is parameterized by the coefficient K_d, given as:

$$(3)\quad K_d = \frac{\sum_{n=1}^{\infty} A_i}{\sum_{n=1}^{\infty} C_i}$$

where A_i is the equilibrium concentration of OC of the *i*th species sorbed (mg C kg^{-1} mineral) and C_i is the equilibrium concentration of DOC of the *i*th species in solution (mg C L^{-1}). A and C are summed across all species in their respective phases and K_d takes units of L kg^{-1}.

In order to better predict the effects of hypoxia on carbon cycling through soil and marine environments, we sought to arrive at a quantitative prediction of the partition coefficient K_d of NOM for surface sorption onto a model iron sulfide mineral. We hypothesize that different affinities are to be expected between iron oxides and iron sulfides due to the changing chemical natures of Fe^{3+} and Fe^{2+} species. We first experimentally determined the sorption affinities of three NOM sources and their mixtures for the surface of freshly-precipitated mackinawite through batch sorption experiments. We then built a predictive model using partial least squares regression (PLSR), taking as inputs the functional group characteristics of the varying NOM sources as determined by attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR). This approach mitigated the exposure of iron sulfides to oxygen commonly risked when directly measuring surface complexes, and did not require high-energy bombardment for surface characterization which has been known to alter the physical state of mackinawite (Rickard and Morse, 2005). Our model provided us with quantitative data on sorption affinity that can be compared to the existing studies from oxic systems.

2.2 Materials and Methods

2.2.1 Preparation of organic matter

Three different endmembers were selected to represent NOM inputs to marine sediments. Soil NOM was chosen to represent highly-degraded and stable terrestrial inputs to sediments. Soil NOM was extracted from aged compost soil (>10 years since last amendment) by sonicating the residues in milliQ water and retaining the aqueous fraction. Plankton, representing highly reactive marine NOM inputs to sediments, was extracted from a species of microalgae of genus *Nannochloropsis* (Reed Mariculture) by repeated freeze-thaw cycles in liquid N_2 and subsequent filtration (0.45 µm nominal pore size). Corn leaf extract, representing fresh and highly reactive terrestrially-derived NOM, was prepared by heating green corn leaves in milliQ water to 60°C and then macerating to a pulp in a blender. Solids were filtered off as above and the solution containing all dissolved NOM was retained for use. Each endmember solution was divided into high molecular weight (HMW > 1 kDa) and low molecular weight (LMW < 1 kDa) size fractions via dialysis membranes (Spectra/Por) and then re-concentrated by way of a rotovap connected to a vacuum pump. Final DOC concentrations of each endmember size fraction were determined with a total organic carbon analyzer (Shimadzu TOC-VCSH). All endmember solutions were stored without biocide in the fridge at 4ºC when not in use.

In addition to the endmembers, NOM was also extracted from two marine sediments collected from the St-Lawrence Estuary. One sediment (Station SAG-05; 48º24.75'N / 70º49.26W) is characterized by more terrestrial inputs while the other (Station 19; 49º29.50'N / 65º13.19W) receives a greater contribution of marine NOM. Extractions were performed by sonicating the sediments in milliQ water for 2 hours, then centrifuging and filtering off any suspended solids from the supernatant. No size fractionation was performed on these samples.

Stock solutions of either HMW or LMW NOM were prepared prior to each batch sorption experiment using a single endmember or their mixtures, for example 20% soil - 20% plankton - 60% corn, to a concentration of approximately 56 mg L^{-1} DOC. The pH of the stock solutions was adjusted to 7.8 ± 0.1 with concentrated HCl or NaOH, and a portion was subsampled and freeze-dried for subsequent analysis by ATR-FTIR.

2.2.2 Mineral phase preparation

Mackinawite was chosen as the model reduced iron sulfide in marine systems. Prior to each batch sorption experiment, mackinawite was freshly precipitated into each reaction vial by combining 0.5 M solutions of iron (II) chloride tetrahydrate (Sigma #44939) and sodium sulfide nonahydrate (Sigma #208043) to obtain a consistent mass of 43.95 mg FeS, with a 1% excess of sulfide in solution. The mackinawite was precipitated onto 4.59 g of acid-washed and combusted quartz sand (Ottawa sand 20-30 mesh, Fisher Chemical) in order to approximate a final particle density of 0.8 and then thoroughly mixed to ensure coating of the sand.

2.2.3 Batch sorption experiment

All precipitations and sorption experiments were conducted in a glovebox under an N_2 atmosphere in order to prevent the oxidation of mackinawite. An oxygen meter was used to ensure that O_2 levels were maintained below 0.05%. Ten individual sorption experiments were performed per NOM input across a range of 5-40 mg L^{-1} DOC in 15 mL polypropylene conical vials cleaned of trace iron and leachable organic carbon. The masses and working volumes were chosen to obtain 0.96% FeS by mass which reflects sedimentary AVS concentrations in our neighbouring marine environment of the St. Lawrence Estuary (Balind, 2019). Each vial contained a final volume of 7 mL, consisting of the predetermined concentration of DOC in a 0.7 M NaCl solution in order to approximate the ionic strength of seawater. Blank solutions containing FeS, 0.7 M NaCl and no NOM, as well as 0.7 M NaCl, NOM, and no FeS were also prepared.

Vials were sealed under a stream of N_2 and mixed on an end-over-end shaker for 3 hours to allow sorption to run to equilibrium. After equilibration, vials were centrifuged at 4000 rcf for 10 minutes and the supernatant then carefully extracted in the glovebox while verifying the pH for any substantial drift. Before DOC analysis, the supernatant samples were acidified to pH ~2 with concentrated HCl to prevent the oxidation and precipitation of any $FeS_{(aq)}$, which might artificially decrease the residual levels of DOC. Residual DOC in solution was measured by TOC-VCSH analyzer, and the concentration of organic carbon sorbed to FeS at equilibrium was calculated by mass-

balance (Krupka et al., 1999). Partition coefficients were obtained for each NOM input by measuring its equilibrium sorption onto mackinawite across 5 sequential concentrations of DOC, in duplicate.

2.2.4 ATR-FTIR analysis

Each of the NOM solutions used in the batch sorption experiments was subsampled prior to sorption and analyzed by ATR-FTIR for subsequent use as model inputs. The freeze-dried aliquots were homogenized and analyzed on a Cary 630 ATR-FTIR with diamond crystal (Agilent Technologies) at a resolution of 4 cm^{-1}. All spectra were preprocessed before use in the PLSR model. Each spectrum was baseline corrected in the software Spectragryph v1.2.15 (Menges, 2020) due to significant scattering at the lower end of the wavenumber scale. Subsequently in R v4.1.0 (R Core Team, 2021), each spectrum was converted to absorbance and averaged over 5 separate measurements to account for variation in the ATR measurement procedure and then scaled between 0 to 1 absorbance units. Finally, all sample spectra were smoothed using the R *signal* package v0.7-7 (Signal developers, 2021) by applying a first order Savitzky-Golay filter with 15 spectral points.

2.2.5 Partial least squares regression (PLSR)

The PLSR models to predict sorption of NOM onto mackinawite based on functional group distribution were built in R using the package *pls* v2.7-3 (Mevik et al., 2020). In order to normalize the variance in the data, the logarithm of the K_d values obtained from the batch sorption experiments were instead used as the target variables for model training and testing, and the predictor variables consisted of the preprocessed ATR-FTIR spectra.

2.3 Results

2.3.1 Precipitation of mackinawite

Immediately prior to the batch sorption experiments, mackinawite was freshly precipitated in each experimental vial. A black solid rapidly precipitated upon mixture of Fe^{2+} and S^{2-} solutions. The average BET surface area of the solid formed this way and dried under an N_2 stream was

59.7 ± 5.2 $m^2 g^{-1}$. No definitive powder X-ray diffraction was obtained, likely owing to the small crystal size of the mackinawite. ATR-FTIR analyses of the surface-NOM complexes were inconclusive since the dark, highly-absorbing compound yielded no signal. Contact with HCl caused its immediate dissolution with the evolution of sulfide gas. No further characterization was performed.

2.3.2 Elemental analysis of NOM endmembers

Subsamples of the NOM endmember solutions were freeze-dried and analyzed as solids for carbon content as well as C/N ratio with an elemental analyzer coupled to an isotope ratio mass spectrometer (EA-IRMS; Eurovector EA3000). The solid residues of the three HMW endmembers were each approximately 30% OC by mass (Appendix, Table A.1). The C/N ratios of the endmembers decreased in the order corn > soil > plankton as is to be expected given the higher proportion of polypeptides in plankton compared to the other two endmembers. In the LMW fractions, however, the residues of the corn and soil stock solutions were dominated by inorganic salts. The percentage of OC for the solids dropped from approximately 20.8% for plankton, to 4.3% for corn and 1.4% for soil. An extremely low C/N ratio of 0.2 for the soil endmember is indicative of an abundance of nitrate salts in its residues, which is further supported by the FTIR spectrum showing a predominant nitrate peak centred around 1370 cm^{-1} (data not shown).

2.3.3 FTIR Analysis of NOM endmembers

The endmember compounds (soil, plankton and corn) of the HMW fraction were analyzed by ATR-FTIR to determine which chemical functional groups were present. Due to the overlapping nature of the peaks, the 2nd derivative of the spectra were calculated to obtain peak centres using a Savitzky-Golay filter (Signal developers, 2021), and these peak centres were plotted over the original smoothed spectra in Figure 2.1. Two regions are of particular interest for modelling as they contain the most chemical information. The first region between 3600-2500 cm^{-1} contains information on O-H, N-H and C-H stretching. A broad region for O-H stretching (~3275 cm^{-1}) can be seen in all spectra. C-H stretching vibrations for aliphatic compounds are evident from the peaks at 2920 and 2850 cm^{-1} although these are diminutive in the soil spectrum. A shoulder at approximately 3060 cm^{-1} indicates alkenes/aromatic compounds due to C-H stretching for sp^2 hybridized carbon.

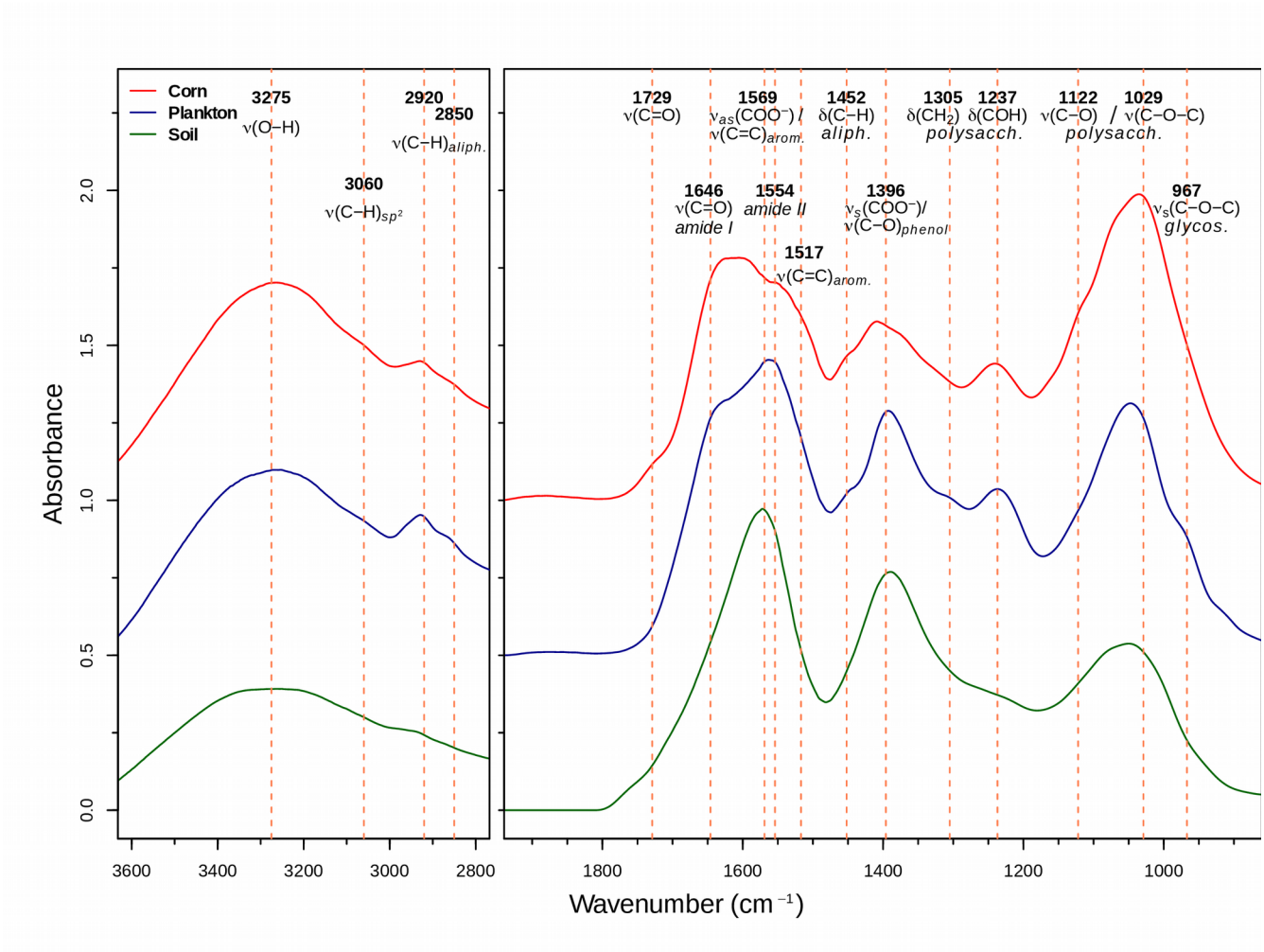

Figure 2.1. ATR-FTIR spectra of the HMW NOM endmembers along with major peak assignments.

The spectra demonstrate the greatest amount of variation in chemical composition in the second range of interest from 1800-900 cm^{-1} which includes the carbonyl stretching regions and fingerprint region. In general, three broad bands can be discerned: a carbonyl/amide/aromatic stretching region (approx. 1660-1500 cm^{-1}), a region associated to C-H bending, C-O phenol stretching and aromatic ring stretching (1470-1300 cm^{-1}), and a C-O stretching region indicative of alcohols, ethers and polysaccharides (1150-1000 cm^{-1}) (Parikh et al., 2014). Aside from small shoulders in the corn and soil spectra, very little absorbance occurs at 1700 cm^{-1} and above in the carbonyl C=O stretching region. The onset of the first spectral envelopes becomes evident at approximately 1650 cm^{-1}. The disappearance of carbonyl stretching is explained by Evangelou et al. (2002) and sources therein: deprotonated carboxylic acids become complexed to metal ions, causing the appearance of antisymmetric and symmetric COO^- stretching bands instead at approximately 1600 and 1380 cm^{-1}

respectively. These two regions correspond to major peaks in all three HMW endmembers and are the most prominent features of the soil OM. Along with carboxylate, regions assigned to aromatic C=C stretching are prominent in all three samples. This includes the peak centered at 1569 cm^{-1} in soil which overlaps with the two other endmembers, and the shoulders from 1517-1504 cm^{-1} (Jouraiphy et al., 2008; Hatcher et al., 2014; Boukir et al., 2019). Unique to the corn and plankton endmembers are peaks at approximately 1646 & 1554 cm^{-1}. We associate these to the C=O stretch of amide I (Howe et al., 2002; Abdulla et al., 2010; Hatcher et al., 2014) and δ(N-H) plus ν(C-N) modes of amide II (Howe et al., 2002; Jouraiphy et al., 2008; Mayers et al., 2013) indicative of polypeptides. Strong absorbance in these regions of the plankton spectrum is consistent with its low C/N ratio (Section 3.2). The corn endmember registered a much higher C/N ratio so we conclude that aromatic C=C stretching must also contribute to these peaks (Jouraiphy et al., 2008; Margenot et al., 2015).

In the adjacent spectral band spanning 1470-1300 cm^{-1}, we find peaks at 1452 cm^{-1} in the spectra of corn and plankton. These peaks likely correspond to $\delta_{as}(CH_3)$ and $\delta_{as}(CH_2)$ in aliphatics and peptides (Dean et al., 2010; Hatcher et al., 2014). The region centred at approximately 1396 cm^{-1} is most prominent in the soil OM fraction and decreases in plankton and corn. Along with symmetrical carboxylate stretching, this region includes phenol aromatic-OH stretching coupled with O-H deformations (Stone et al., 2001; Plaza et al., 2003; Roux-Michollet et al., 2010).

There is a unique band that is well-defined in the spectra of plankton and corn with peak centre at 1237 cm^{-1}. This band, along with a shoulder situated at 1305 cm^{-1} may be a feature of cellulose and other polysaccharides caused by COH bending and CH_2 rocking (Dai and Fan, 2011). These peaks also overlap with the phenol aromatic-OH stretching region (Abdulla et al., 2010; Boukir et al., 2019), but notably this pair is far less resolved in the soil spectrum indicating that polysaccharides are likely candidates for the corn and plankton spectra. Peaks at 1122 and 1029 cm^{-1} are signals commonly associated to cellulose and other polysaccharides (Howe et al., 2002; Plaza et al., 2003; Stehfest et al., 2005; Jouraiphy et al., 2008; Plácido et al., 2013). The shoulder at 967 cm^{-1} is assigned to ν_s(C-O-C) of the cellulose glycosidic bond and is considered characteristic (Kunov-Kruse et al., 2013). All peaks associated to polysaccharides are most prominent in the spectrum of corn, followed by plankton, and are nearly absent in the soil spectrum.

2.3.4 Batch sorption experiments

A total of 40 batch sorption experiments of endmember NOM and their mixtures were performed. Linear regressions were performed on the data and the fitted values of the slopes were retained as the estimates of K_d (Figure 2.2). K_d values for 25 HMW and 14 LMW samples are listed in Tables 2.1 and 2.2 along with their relative standard deviations (RSD) based on the fit of the line. Additionally, the averaged results of triplicate experiments on the two different sediment extracts are given in Table 2.3. One LMW sample was discarded due to the formation of precipitates in the OM stock solution prior to analysis, leading to an artificial decrease of the measured K_d value for that sample. Values of K_d among all samples, calculated as the slope of best fit by linear regression, spanned nearly two orders of magnitude through the range of 38 to 1130 L kg^{-1}. A comparison of $logK_d$ values obtained for the HMW and LMW fractions is given in Figure 2.3. The largest distribution coefficients were obtained among HMW compounds with a mean $logK_d$ value of 2.62, while the mean value for the LMW fraction was 1.85. The general trends observed among sorption affinities were corn > plankton > soil for HMW compounds and plankton > corn > soil for LMW. Due to the differences in magnitude and affinity order between the two fractions, HMW and LMW compounds were modelled separately.

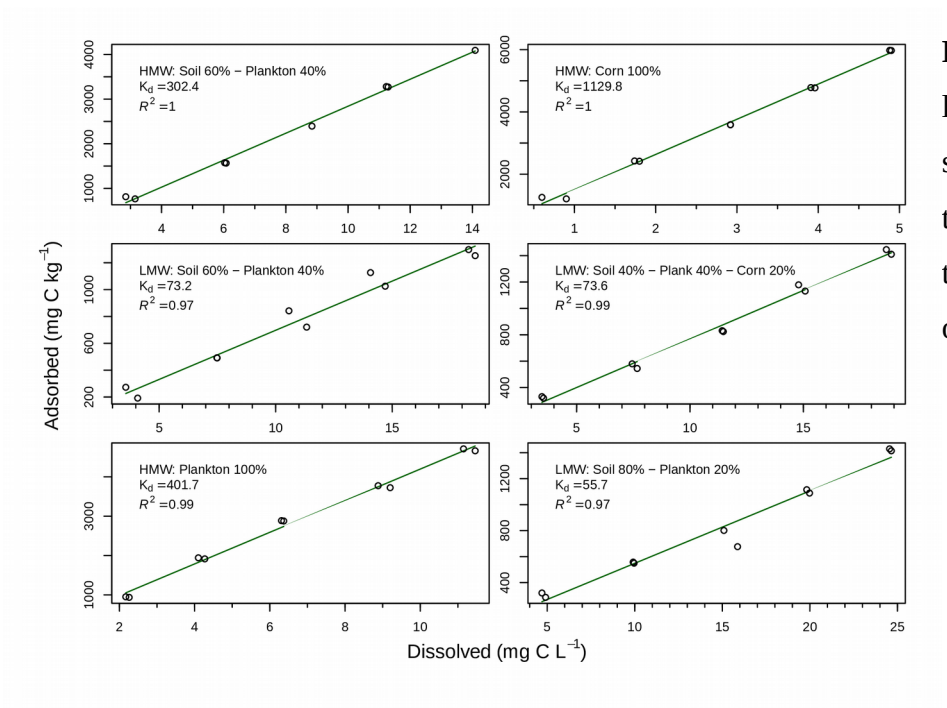

Figure 2.2. Examples of linear regressions on 6 batch sorption experiments using the NOM endmembers and their mixtures. K_d is in units of L kg^{-1}.

Table 2.1. Results of linear regressions on batch sorption experiments of HMW NOM. K_d from slope of best fit line.

HMW Samples	K_d	R^2	Relative SER (%)	Inter-Day RSD (%)	Inter-Day RSD w/o Outlier (%)
Soil 100%	188.3	0.97	5.8		
Soil 80% - Plankton 20%	218.8	0.98	5.6		
Soil 60% - Plankton 40% (n = 4)	239.0	0.99	4.0	18.6	
Soil 40% - Plankton 60%	305.7	0.98	4.7		
Soil 20% - Plankton 80%	391.7	0.96	7.5		
Plankton 100% (n = 4)	349.7	0.99	3.2	25.0	2.0
Plankton 80% - Corn 20%	398.4	0.99	3.3		
Plankton 40% - Corn 60%	553.6	0.91	12.7		
Plankton 20% - Corn 80%	816.7	0.97	7.6		
Corn 100% (n = 4)	997.6	0.98	5.1	12.2	
Soil 20% - Corn 80%	603.5	0.93	10.9		
Soil 20% - Plankton 20% - Corn 60%	401.4	0.99	4.4		
Soil 20% - Plankton 40% - Corn 40%	452.2	0.95	8.0		
Soil 20% - Plankton 60% - Corn 20%	261.0	0.96	7.6		
Soil 40% - Plankton 40% - Corn 20%	325.5	0.98	5.8		
Soil 60% - Plankton 20% - Corn 20%	285.0	0.98	4.9		
Average	**469.0**	**0.98**	**5.3**	**18.6**	**11.0**

Relative Standard Error of the Regression (SER) refers to the error on linear regression slope calculated from a single set of batch sorption experiments. Inter-day relative standard deviation (RSD) is the variation in K_d measured when replicate experiments were performed on separate days (number of replicates in parentheses). Samples in red and one replicate of the Plankton 100% were excluded from PLSR; K_d in units of L kg^{-1}.

Table 2.2. Results of linear regressions on batch sorption experiments of LMW NOM. K_d from slope of best fit line.

LMW Samples	K_d	R^2	Relative SER (%)	Inter-Day RSD (%)	Inter-Day RSD w/o Outlier(%)
Soil 100% (n = 3)	47.3	0.96	6.0	21.4	15.1
Soil 80% - Corn 20%	49.2	0.90	12.1		
Soil 80% - Plankton 20%	55.7	0.97	6.6		
Soil 60% - Plankton 40%	73.2	0.97	6.5		
Soil 40% - Plankton 60%	70.3	0.96	7.2		
Soil 20% - Plankton 80%	100.2	0.93	10.0		
Plankton 80% - Corn 20%	136.4	0.92	10.5		
Corn 100%	74.3	0.99	4.0		
Soil 20% - Plankton 60% - Corn 20%	100.2	0.99	9.4		
Soil 20% - Plankton 40% - Corn 40%	73.7	0.96	6.8		
Soil 40% - Plankton 40% - Corn 20%	73.6	0.99	3.5		
Soil 60% - Plankton 20% - Corn 20%	51.5	0.99	3.4		
Average	**75.5**	**0.96**	**6.9**		

Relative Standard Error of the Regression (SER) refers to the error on linear regression slope calculated from a single set of batch sorption experiments. Inter-day relative standard deviation (RSD) is the variation in K_d measured when replicate experiments were performed on separate days (number of replicates in parentheses). One replicate of the Soil 100% was excluded from PLSR; K_d in units of L kg^{-1}.

Table 2.3. Results of linear regressions on batch sorption experiments of sediment-extracted NOM. K_d from slope of best fit line.

Sediment Samples	K_d	R^2	Relative SER (%)	Inter-Day RSD (%)
SAG-05 (n = 3)	204.4	0.97	4.8	19.4
Station 19 (n = 3)	306.4	0.99	3.8	8.9
Average	**255.4**	**0.98**	**4.3**	**14.1**

Relative Standard Error of the Regression (SER) refers to the error on linear regression slope calculated from a single set of batch sorption experiments. Inter-day relative standard deviation (RSD) is the variation in K_d measured when replicate experiments were performed on separate days (number of replicates in parentheses); K_d in units of L kg^{-1}.

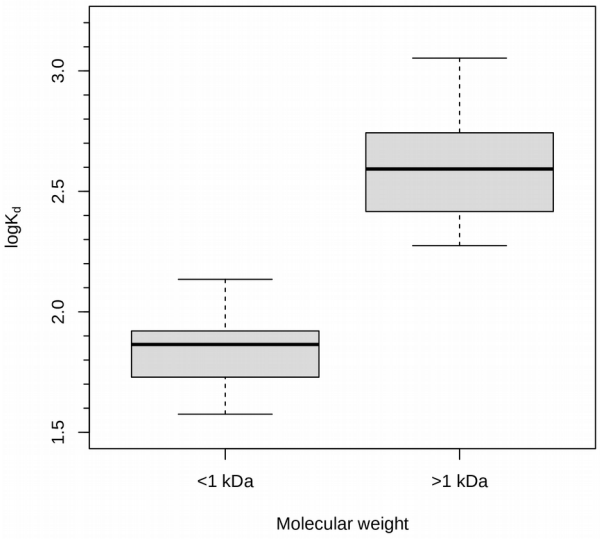

Figure 2.3. Boxplot distributions of $\log K_d$ values for two size fractions of NOM.

For both HMW and LMW fractions, a multiple linear regression (MLR) fit was performed in R to predict the experimentally-determined K_d based on the relative proportions of soil, plankton and corn NOM in each sample. The MLR fit for the HMW compounds along with a 95% confidence band for the fit is plotted in Figure 2.4. This preliminary fit yielded an R^2 of 0.89. Three samples appear to deviate significantly from linearity as their 95% confidence intervals do not overlap with the confidence band of the MLR prediction due to lower than predicted results. These points were therefore excluded from the subsequent PLSR model-building step as outliers (see Tables 2.1 and 2.2 for the identity of the dropped experiments). After their removal, the coefficient of determination for the MLR model increased to 0.96.

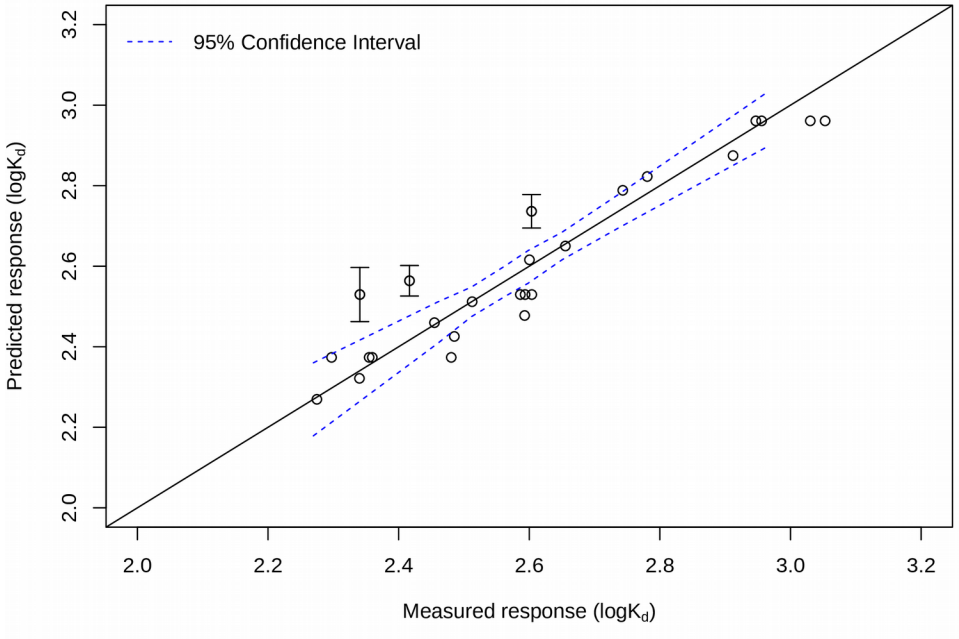

Figure 2.4. MLR model for HMW compounds with 95% confidence band. Three points with their 95% confidence intervals appear as clear outliers. R^2 for fit (minus outliers): 0.96.

A similar screening procedure was preformed on the LMW compounds. The results of the MLR fit on the 14 remaining LMW samples based on NOM composition are given in Figure 2.5. The R^2 of the fit is 0.86, showing more scatter than for the HMW compounds. In addition, one sample point appears to be an outlier at the 95% confidence level. Dropping this point improves the fit by MLR and increases R^2 to 0.92.

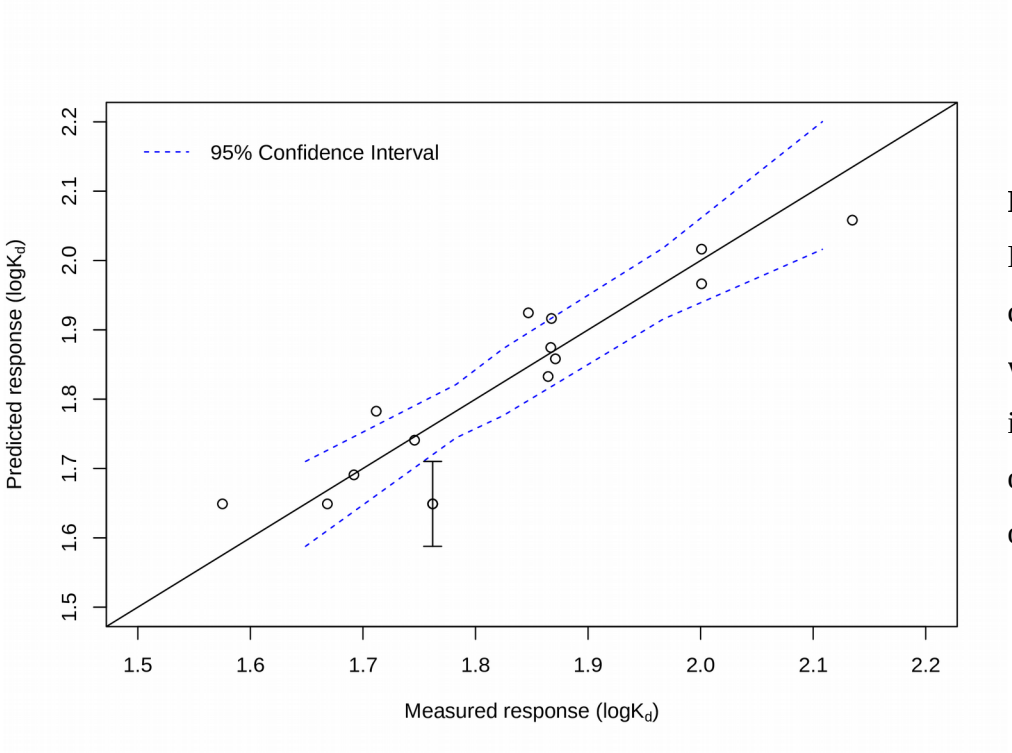

Figure 2.5. MLR model for LMW compounds with 95% confidence band. One point with its 95% confidence interval appears as a clear outlier. R^2 for fit (minus outlier): 0.92.

2.3.5 PLSR model

The final sample pool for HMW compounds consisted of 22 observations once outliers from the batch sorption experiments were removed. Of these, 5 were set aside as a test group to validate the predictive results of the PLSR fit, leaving 17 observations with which to train the model. The model was built using the FTIR spectral range from 3600-900 cm^{-1} since the region below 900 cm^{-1} was prone to fluctuations from the baseline correction. An additional region between 2500-1900 cm^{-1} was excluded due to the fact that it contained little useful spectroscopic information and was prone to interference from background correction of the ATR diamond. Repeated 5-fold cross-validation (5 repeats of 5 randomly generated folds) was performed using the R *caret* package v6.0-88 (Kuhn, 2021) to determine the optimal number of components/latent variables (LVs) without leading to overfitting.

The global minimum in root mean square error of cross-validation (RMSECV) obtained this way was 0.058 at 2 LVs. The predictive PLSR model was then built, yielding a fit of the training data with a root mean square error (RMSE) of 0.042 and an R^2 of 0.97. Two LVs were sufficient to describe 92.9% of the variance among HMW FTIR spectra (*X* variables) and 96.8% of the variance in $\log K_d$ (*Y* variable). The fit of the training data by PLSR is summarized in Figure 2.6**.**

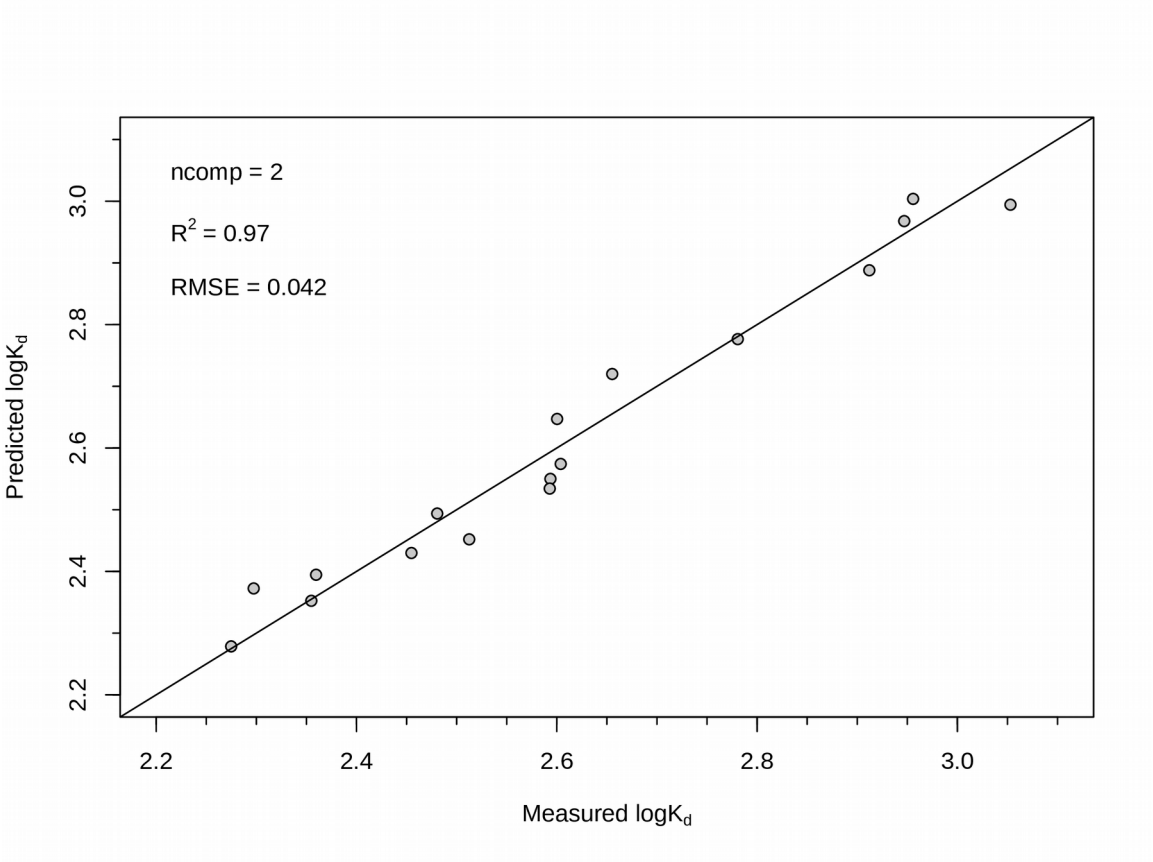

Figure 2.6. Final fit of HMW training samples by PLSR at 2 LVs/components.

The data were also inspected for outliers by computing Hotelling's T^2 value for each data point (Appendix, Figure A.1) using the *HotellingEllipse* package v0.1.1 in R (Goueguel, 2021). All samples' T^2 values fall well below the 95% cutoff of 7.85 at 2 LVs, indicating they can be retained. To validate the model's predictive error on unseen data as estimated by cross-validation, the model was then applied to the 5 hold-out samples, yielding a root mean square error of prediction (RMSEP) of 0.036 with a predictive Q^2 of 0.98 (Figure 2.7).

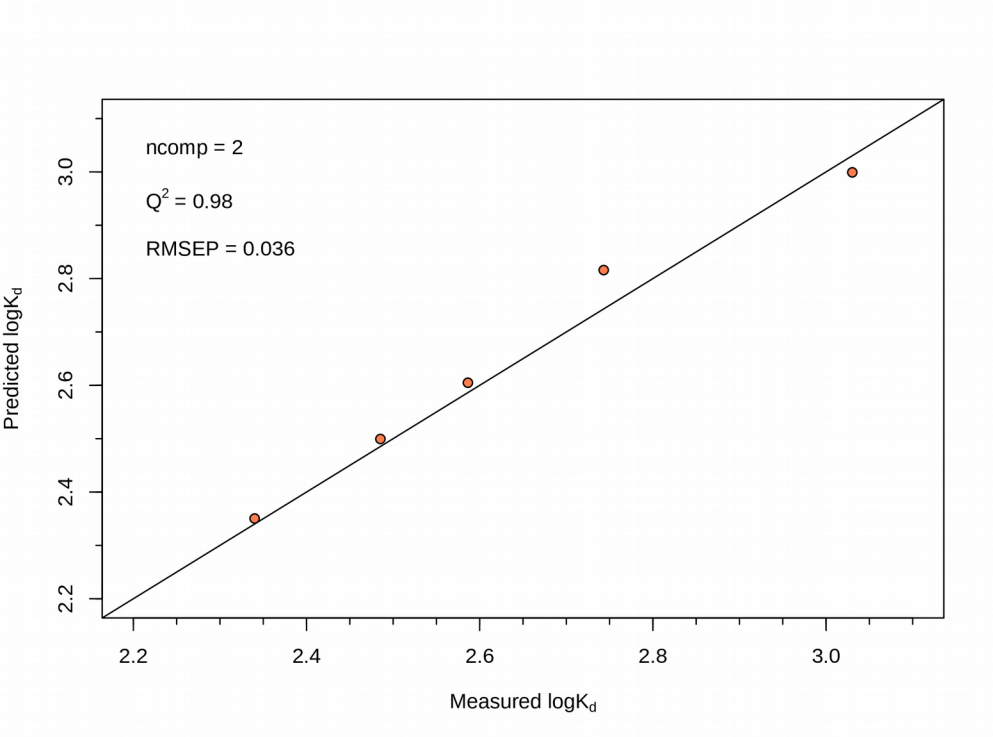

Figure 2.7. Validation of PLSR model at 2 LVs/components on 5 hold-out samples.

2.4 Discussion

2.4.1 Batch sorption experiments

DOC concentrations in marine sediment porewaters directly beneath the SWI can be an order of magnitude higher than in the water column and generally increase with depth. Alperin et al. (1999) reported ranges of approximately 6-56 mg C L^{-1} for sediment collected from continental slopes, while Thimsen & Keil (1998) measured estuary porewater concentrations of approximately 15-21 mg C L^{-1}, although estuaries receiving high amounts of terrestrial runoff can attain porewater concentrations comparable to the maximum as reported by Alperin et al. (1999) (Balind, 2019). The batch sorption experiments used to calculate K_d for the NOM mixtures were carried out with concentrations of DOC from 0 - 56 mg C L^{-1} thereby capturing the expected range for marine sediment porewaters. All sorption experiments displayed a high degree of linearity, with an average R^2 of 0.98 for the HMW (n = 25) and 0.96 for the LMW (n = 15) mixtures respectively (Tables 2.1 and 2.2), indicating that the constant partition coefficient is a suitable assumption and can be used to measure and compare sorption affinities. The R^2 values also indicate the level of intra-day reproducibility within a single experiment. Thimsen & Keil (1998) reported an average RSD of <10% for the measured K_d based on porewater

DOC sorption to montmorillonite and iron oxides and Arnarson & Keil (2000) reported an RSD of approximately 5.1% for sorption onto montmorillonite. Our predicted error on single K_d measurements averaged 6.3% RSD across 40 experiments.

A comparison of the mean K_d values obtained on replicate experiments provides a measure of inter-day variability. We observe among the 6 groups of replicates that inter-group RSD spans a range of 8.9-25.0% with a mean of 17.6%, almost triple the intra-day error (Tables 2.1 and 2.2). An increased error on inter-day replicates is not surprising for batch sorption experiments, especially given the sensitivity of mackinawite to unintentional oxidation. Serne and Relyea (cited in Krupka et al., 1999) reported that an interlaboratory study on batch sorption experiments performed by 9 different labs yielded K_d values varying between 1-3 orders of magnitude for the same procedure. Since the coefficients of determination were generally quite high between 0.95-1.0 within an individual experiment, variation in K_d was likely caused by some environmental parameter affecting all vials within an experiment and not due to fluctuations in sample pH, salinity or particle density of individual vials. The most likely candidate is cryptic oxidation of the air-sensitive mackinawite in rare cases, which could affect all samples in a batch accidentally during precipitation, mixing and centrifugation. The larger inter-day error demonstrates the sensitivity of the system to small parameter adjustments, however the variations in K_d are well below even one order of magnitude demonstrating that error was minimized throughout. This exercise does highlight the importance of detecting outliers as was done in the MLR step before moving forward with PLSR modelling.

2.4.2 Functional group characterization by ATR-FTIR

All spectra of HMW NOM were similar in profile to spectra obtained from the fulvic fraction of composting residues derived from activated sludge and plant matter obtained by Jouraiphy et al. (2008). The natural NOM maturation process brings about a loss of fine structure and homogenization resulting in operationally defined fractions of NOM including molecularly uncharacterized organic matter (MU-OM) (Lee et al., 2004; Burdige, 2007) and carboxylate-rich alicylic material (CRAM) (Hertkorn et al., 2006; LaRowe et al., 2020). The structure of NOM plays a role in determining how susceptible a particular fraction is to microbial degradation. The irregular crosslinked network of ether and C-C bonds typical of lignins proves far more recalcitrant to microbial degradation than simpler

repeating amide and glycosidic linkages of polypeptides and polysaccharides (LaRowe et al., 2020). As a result, protein and carbohydrate concentrations diminish far more rapidly with depth in both the water column and in sediments (Hertkorn et al., 2006; Burdige, 2007; Hatcher et al., 2014). Based on the varying maturities of our NOM extracts, it is expected to see the effects of degradation reflected in the FTIR spectra of the endmembers since corn and plankton represent much fresher and more reactive inputs to the sorption experiments. The dominant features in their HMW spectra are vibrations commonly associated to polysaccharides in the range of 1122-950 cm^{-1}. Similarly, corn and plankton are also richer in polypeptides as evidenced by the amide I peak at around 1646 cm^{-1} in the HMW spectra, which was notably absent in the soil. In contrast, the soil endmember represents the most degraded source of NOM. Its simplified FTIR spectrum consisted of a few broad, rounded peaks. Nevertheless, we infer the presence of phenolic and carboxylate groups due to an important region of absorbance centred at approximately 1390-1400 cm^{-1}. Additionally, the broad absorbance band at 1569 cm^{-1} in the HMW spectrum is indicative of aromatic and carboxylate groups. Similar spectral trends have been observed during the composting and associated humification of plant NOM and sewage: FTIR absorbance associated to polysaccharides and protein-like material were seen to decrease due to preferential microbial degradation while aromatic, oxygenated alkyl and carboxylic intensities increased (Jouraiphy et al., 2008; He et al., 2011). Thus the endmembers act as representatives not only of different sources of NOM to marine sediments but also represent variations of reactivity and degradation. Prolonged degradation is expected to render the spectra of plankton and corn more similar to that of soil.

2.4.3 Sorption prediction by MLR & PLSR

Initial attempts to model K_d based on the relative proportions of soil, plankton and corn endmembers present in each mixture were accomplished by MLR. This served as an important proof-of-concept before proceeding with PLSR model building on FTIR spectra. MLR confirmed a linear relationship between the NOM mixtures and their experimental sorption affinities for FeS within the three component (soil, plankton, corn) system, as shown in Figures 2.4 and 2.5. After outliers were identified and removed as discussed in Section 2.4.1, the coefficients of determination for the MLR models rose to 0.96 and 0.92 for the HMW and LMW fractions respectively, providing an approximate upper-bound to the accuracy expected from a PLSR model built on this data.

An appropriate level of model complexity is required to best capture the data variance without leading to overfitting, but can be difficult to assess in the case of LVs for PLSR. Since the sum of contributions of soil, plankton and corn NOM to each sample were constrained to unity, only two variables were required as inputs for MLR, the relative proportion of the third component being predetermined by the previous two. This information therefore proved valuable for the PLSR model building step. Cross-validation of the PLSR model also indicated that 2 LVs were sufficient, and furthermore the R^2 of 0.97 was almost exactly as predicted by MLR and in fact slightly higher. Due to the intrinsically labile nature of NOM, it is possible that modelling by FTIR spectra can be more accurate than modelling based on the calculated ratios of each NOM component. Slight day-to-day variations in the DOC concentrations of NOM inputs should be reflected in changes to their FTIR spectra even if they go unmeasured by other analytical means. The predictive stability of the PLSR model was confirmed by analyzing a hold-out group of HMW samples (Figure 2.7), demonstrating a robust model despite the limited number of samples available with which to train it. In general, the low predictive errors and high coefficients of correlation indicate an excellent fit for the data. Due to the high salt content of the LMW solid residues (see Section 2.3.2), it was not possible to reliably model this fraction by PLSR. Initial attempts to background correct for salts in the FTIR were unstable results due to the disproportionate ratios of salt-to-NOM signals.

2.4.4 Quantitative comparison of K_d

Partition coefficients serve as a quantitative measure of adsorption affinity that can in the ideal case be compared across studies. In practice, the sensitivity of batch sorption experiments (Section 2.4.1) to a multitude of parameters can present a major and often unrecognized source of variability. Nevertheless, a comparison of the results to partition coefficients on iron oxides and other sedimentary minerals from the literature is instructive for understanding carbon cycling across oxic and anoxic boundaries. A summary of cross-study K_d values, normalized to specific surface area of the mineral phase, is given for model NOM compounds and for sediment NOM extracts in Tables 2.4 and 2.5. Many studies opted to build full isotherms and fit them with, e.g., the Langmuir model, making a direct comparison challenging. However, it has been pointed out (Limousin et al., 2007; Salvestrini et al., 2014) that as $q_e \rightarrow 0$, the ratio of q_e/C_e can be approximated from the Langmuir fit as $K_L q_{max}$.

Since q_e/C_e is the partition coefficient K_d at the linear portion of the Langmuir isotherm, this calculation was performed on the data when only Langmuir fits were available.

Table 2.4. Comparison of normalized K_d values for model NOM compounds across several studies and mineral phases.

K_d (mL/m²)	Mineral phase	NOM source	pH	Study
227.1[a]	Hematite (α-Fe_2O_3)	Fulvic acid (Suwanee river)	6.5	Gu et al. (1994)
223.4[a]	Hematite (α-Fe_2O_3)	Fulvic acid (Suwanee river)	4	Gu et al. (1996)
133.8[a]	Hematite (α-Fe_2O_3)	Hydrophilic NOM (wetland pond)	6.8	Gu et al. (1995)
98.15[a]	Goethite (α-FeOOH)	Polymaleic acid	5.8	Wang et al. (1997)
83.40[a]	Goethite (α-FeOOH)	Humic acid	8.5	Safiur Rahman et al. (2003)
58.62[a,b]	Goethite (α-FeOOH)	Fulvic acid (Suwanee river)	5.8	Wang et al. (1997)
27.86[c]	Kaolinite	Humic acid (peat)	7	Feng et al. (2005)
16.18[c]	Mackinawite (FeS)	Corn > 1kDa	7.8	Present
6.58[c]	Mackinawite (FeS)	Plankton > 1kDa	7.8	Present
3.15[c]	Mackinawite (FeS)	Soil > 1kDa	7.8	Present
2.67[c]	Montmorillonite	Humic acid (peat)	7	Feng et al. (2005)
1.26[c]	Goethite/Kaolinite mix	Humic acid (soil)	8	Chen et al. (2019)
1.24[c]	Mackinawite (FeS)	Corn < 1kDa	7.8	Present
0.76[c]	Mackinawite (FeS)	Soil < 1kDa	7.8	Present

a: K_d calculated from Langmuir fit; *b*: q_{max} extrapolated in study; *c*: K_d measured directly.

Table 2.5. Comparison of normalized K_d values for sediment-extracted NOM across several studies and mineral phases.

K_d (mL/m²)	Mineral phase	NOM source	pH	Study
16.8	Hematite (α-Fe_2O_3)	Sediment extract	7.8	Thimsen & Keil (1998)
13.6	Bulk sediment	Sediment extract	7.8	Thimsen & Keil (1998)
12.8	Montmorillonite	Sediment extract	7.8	Thimsen & Keil (1998)
5.13	Mackinawite (FeS)	Sediment extract (STN-19)	7.8	Present
3.42	Mackinawite (FeS)	Sediment extract (SAG05)	7.8	Present
0.097[a]	Montmorillonite	Sediment extract > 1kDa	8	Arnarson & Keil (2000)
0.058[a]	Montmorillonite	Sediment extract < 1kDa	8	Arnarson & Keil (2000)

a: SSA of montmorillonite from Thimsen & Keil (1998).

From Table 2.4, it appears that mackinawite performs comparably to montmorillonite in terms of sorption affinity to model NOM compounds. Pure mackinawite may have a higher affinity for sediment-extracted NOM than montmorillonite, although partition coefficients for montmorillonite vary over 2 orders of magnitude across studies at similar pH (Table 2.5). Sorption affinities of iron oxides (hematite and goethite) for model and sedimentary NOM tended to be higher than for mackinawite, by as much as 2 orders of magnitude for the model NOM compounds. Care must be taken when interpreting these trends as many of the model NOM studies on iron oxides were conducted at a lower pH and sorption behaviour onto minerals has been proven to be highly pH dependent (Gu et al., 1994).

2.4.5 Spectral analysis of regression coefficients

Once the PLSR model was fitted, the regression coefficients could be extracted and plotted against the wavenumber scale to obtain a pseudo-FTIR spectrum. Positive peaks indicate regions that increase the final predicted value of $\log K_d$ whereas the negative peaks decrease the outcome, making apparent the spectral regions that are correlated with greater or lesser sorption affinity with respect to the mean. The plot for the HMW regression coefficients, along with identification of major peaks, are

given in Figure 2.8. The coefficient plot for the HMW fit at two components is readily interpretable. The FTIR region between 1150-950 cm^{-1} associated to C-O and ring stretching of polysaccharides is correlated with the highest sorption affinity among HMW compounds. Additional regions of positive correlation are associated to O-H stretching (broad peak centred at 3273 cm^{-1}) and aliphatic C-H stretching (2920-2850 cm^{-1}), vibrational modes which are consistent with polysaccharides and polypeptides. The positive coefficient values assigned to these wavenumbers are the result of the plankton and corn endmembers and their higher K_d values. The greatest positive modifier in the coefficient trace is by far the polysaccharide region. There is also a small positive peak located at 1651 cm^{-1} which is associated to the amide I C=O stretch.

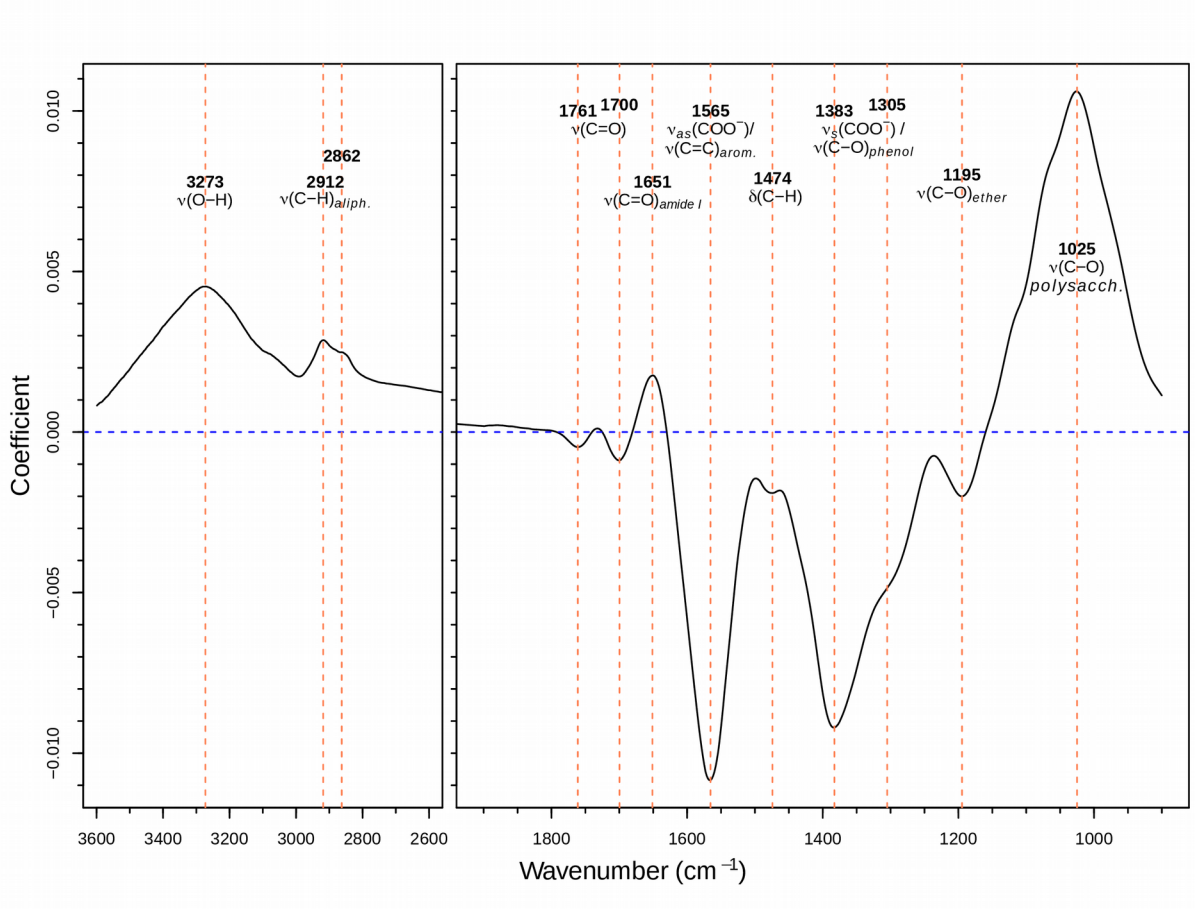

Figure 2.8. Regression coefficients by wavenumber as determined by PLSR at 2 LVs.

In contrast, a lower-than-average K_d value is correlated to strong absorbance at 1565 cm^{-1}, a region indicative of aromatic C=C stretching and carboxylate, which makes up the major absorbance

peak for the soil endmember, but also overlaps with absorbance in plankton and corn. In addition, a broad negative band has a maximum at 1383 cm^{-1} with a prominent shoulder at approximately 1305 cm^{-1}. This region includes phenol C-O stretching plus O-H bending and symmetrical carboxylate stretching. A small but distinct negative peak at 1195 cm^{-1} has been assigned to C-O stretching in both phenols and ethers (Ibarra et al., 1996). Additional valleys are found throughout the coefficient trace whose relative contributions to the modification of $logK_d$ are small but their presence is instructive nonetheless. These include valleys at 1761 and 1700 cm^{-1} which correspond to C=O stretching at wavenumbers consistent with lactones (Lyman et al., 2003) and carboxylic acid/ester (Artz et al., 2008), as well as a peak at 1474 cm^{-1} previously assigned to C-H deformations. An interesting consequence of PLSR is its ability to resolve important spectral regions that were not readily evident in the original spectra, including the ether and carbonyl peaks, as well as the resolution of the amide I peak. Oxidative degradation of lignin-like material involves a coincident carboxylation and demethoxylation of the aromatic material (Ergin et al., 1996). When all the negative regions of the PLSR coefficients are considered as a whole, we obtain the spectral profile of hydroxyl and carboxyl rich aromatic molecules possibly crosslinked by short aliphatic chains and ether linkages, consistent with the oxidized degradation products of lignin (Kirk, 1984; Arndt et al., 2013). Conversely, functional groups associated to polysaccharides and polypeptides were the greatest predictors of sorption.

2.4.6 Polysaccharide-mineral interactions

Polysaccharides have a long history of use in the mining industry as selective depressants for mineral separation (e.g. Laskowski et al., 2007). Hydrophilic polysaccharides such as dextrin and starch operate by binding to the metal centres of desirable ores, causing them to disaggregate and remain in solution while unwanted contaminant minerals are removed by flotation. Liu et al. (2000) proposed that the mechanism be viewed as an acid/base interaction, whereby hydroxyl groups in the polysaccharide ring act as Brønsted acids that interact with mineral surface metal-hydroxylated groups, causing the formation of five-membered rings (see Figure 2.9). The more basic the metal hydroxide surface, the higher the affinity for polysaccharides. Quartz (SiO_2) with a point of zero charge (PZC) of ~2 (Liu et al., 2000) is an example of an acidic mineral with minimal affinity for polysaccharides as compared to the more alkaline hematite (Laskowski et al., 2007). The maximum sorption of

polysaccharides onto metal centres often coincides with the PZC of the hydroxylated metal surface or in the case of certain metal sulfides such as nickel, to the PZC of the corresponding metal hydroxide. Liu et al. (2000) posit this is due to the inevitable oxidation of the metal sulfide surfaces prior to complexation, although another possibility is the hydroxylation of the metal sulfide surface at higher pH. Wolthers et al. (2005) calculated the PZC of the mackinawite surface to be approximately 7.5, which is indeed close to the experimental pH of 7.8 used in this experiment. If instead surface oxidation/hydroxylation does occur, PZC values are expected to vary between 6-8 as is typical of iron (oxy)hydroxides (Kosmulski et al., 2003). Whichever the dominant surface chemical groups may be, the mackinawite surface may have been close to neutral or carried a slight negative charge at this pH. These electronic effects could lead to repulsion with carboxylate functional groups, which were over-represented in the soil NOM, but only minimally affect the neutral polysaccharides. Notably, goethite has been measured to have a PZC as high as 9.3, a characteristic that causes sorption of carboxylated aromatic molecules to increase gradually with decreasing pH as attractive positive charge on goethite builds up (Filius et al., 1997). A high PZC and associated charge effects have been used to explain the greater affinity of goethite for fulvic and humic acids than other minerals with lower PZCs (Meier et al., 1999).

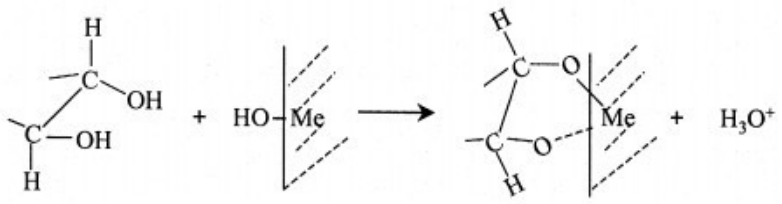

Figure 2.9. Proposed mechanism for polysaccharide binding to mineral surfaces from Liu et al. (2000).

2.4.7 Preferential sorption of NOM in varying environmental conditions

This study on iron sulfide provides unique data for comparison with the numerous studies of NOM sorption onto iron oxides. The majority of these studies were conducted in the context of soil systems but they remain pertinent in part due the high proportion of terrestrial NOM entering many marine systems (LaRowe et al., 2020). A general consensus emerged from these studies as to the molecular fractions with the highest sorption affinities for iron oxide surfaces. For example, Coward et al. (2018) used high resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) to study the fractionation of soil NOM onto goethite. They concluded that oxygenated lignin-like

and aromatic molecules were preferentially sorbed out of solution by short-range order iron mineral surfaces. Polysaccharide and peptide-like NOM was instead found to be retained by mechanisms of coprecipitation and colloidal aggregation with Fe^{3+}, two mechanisms that are not expected to be prevalent in our batch sorption experiments. In similar studies using FT-ICR-MS, Lv et al. (2016) also concluded that polyphenols (lignin-like), aromatic and carboxylic molecules were preferentially sorbed to the surfaces of goethite, lepidocrocite and ferrihydrite. Ohno et al. (2018) experimented with the sorption of water-extractable corn, wheat and soy crop residues onto an iron oxyhydroxide surface and determined that aromatic, lignin-like and N-containing aliphatic molecules were preferentially bound, while Riedel et al. (2013) demonstrated that reoxidation of anoxic porewaters selectively removed plant-derived polyphenols by coprecipitation with Fe^{3+}. In fact the strong affinity of carboxyl and hydroxyl-rich aromatic lignocellulose degradation products for metal oxide surfaces has been well-summarized in the review by Kaiser & Guggenberger (2000). These functional groups, which sorb strongly to iron oxides, dominated the FTIR spectral signature of the soil NOM endmember used in this experiment but instead yielded the lowest sorption affinities for iron sulfides.

Research conducted specifically in marine environments reveals a slightly more complicated picture with regards to preferentially sorption of NOM. Similarly to the soil environments, Barber et al. (2017) demonstrated that 25-62% of reactive ferric iron minerals in coastal sediments were directly complexed to NOM via C=C, C=O and C-OH functional groups. In contrast, a review by Burdige (2007) showed the identifiable NOM associated with five different sediment types consisted predominantly of amino acids (0-19%) and carbohydrates (1-22%) while lignins were in the range of 0-5%, although it should be noted that >60% of the NOM remained uncharacterized. Likewise, Sun et al. (2020) found pelagic sediment cores off the Chinese coast to consist of 80-90% amide/carboxylic, O-alkyl (polysaccharide) and aliphatic functionalities. They were unable, however, to identify any trends in selective preservation of marine vs. terrestrially-sourced NOM through associations to reactive iron minerals, highlighting the added complexity of site-specific redox cycling when predicting long-term preservation. With respect to more sulfidic environments, the NOM content of Black Sea sediments analyzed by Ergin et al. (1996) was found to consist of ligno-carbohydrate and proteinaceous matter. Notably the anoxic sediments had the highest proportions of protein and carbohydrates, which the authors ascribed to the premature halting of oxidative remineralization. Finally, Lalonde et al. (2012) observed that sediment NOM associated to iron-bearing minerals was generally enriched in ^{13}C and

nitrogen as compared to the bulk NOM, indicating stronger associations with carbohydrates and proteins. Samples from this study included sediments from sulfidic (Black Sea) and anoxic (Mexican and Indian margins) environments. These trends are consistent with the observations from our study, particularly where sulfidic and anoxic environmental conditions are explicitly present.

Given the results from this study, it is plausible that iron sulfides could help to promote the removal of proteinaceous and carbohydrate-like material from the water column at the expense of carboxylated polyphenols. This, coupled with decreased microbial degradation in the anoxic sediment zone, may contribute to the preservation of intrinsically labile organic matter, which is consistent with the findings from studies on marine environments. Although reduced iron species may be more selective in the retention of labile carbohydrate and protein-like material than their corresponding iron oxides, the fraction of total DOC removed from solution by iron sulfides may be lower. Aeration of previously anoxic pore waters has been shown to remove over 90% of dissolved iron and 27±7% of previously dissolved OC (Riedel et al., 2013), while trends of both increasing porewater DOC and increasing aromaticity of DOC below the oxic-anoxic interface have been reported elsewhere (Fu et al., 2006). These results are consistent with the preliminary comparative data between oxic and anoxic systems in Tables 2.4 and 2.5. The consequences of climate change for carbon sequestration in sediments are likely to be multiple and conflicting: since less DOC is retained by reduced iron species it is foreseeable that a positive feedback loop will occur due to increasing carbon release. On the other hand, the rates of microbial degradation of organic matter are diminished in anoxic environments. Furthermore, the present research indicates that reduced iron minerals may preferentially sorb to the more labile components of the DOC pool including polysaccharides and polypeptides, acting as a premature halt to microbial degradation. This may favour a negative feedback loop where a decrease in the oxic/anoxic transition depth may bring about a greater degree of NOM sequestration.

2.5 Conclusion

The results of this study support the notion that labile proteinaceous and polysaccharide-rich NOM are better retained by mackinawite than the carboxyl-rich aromatic material consistent with lignin degradation products. With increasing hypoxia/anoxia in sedimentary environments through the combined effects of climate change and human land-use, this change in sorptive preference from iron oxides to iron sulfides may have yet-to-be determined consequences for carbon sequestration in marine sediments. Preliminary results indicate that iron oxides sorb greater quantities of NOM while iron sulfides appear to preferentially target the more labile fractions. The predictive model-building approach by PLSR employed in this study yielded quantitative data which allowed us to compare results across numerous studies. We hope that quantitative modelling will permit researchers to take the step from a descriptive approach to a predictive approach in order to better understand the future of organic carbon dynamics in marine systems.

Chapter 3: Conclusions and Future Work

The present work demonstrates the successful modelling of NOM sorption onto the surface of mackinawite in conditions that mimic natural marine environments with respect to pH, sediment particle density and ionic strength. ATR-FTIR spectroscopy allowed us to determine which organic functional groups were the greatest predictors of sorption onto mackinawite. Furthermore, the measurement of the constant partition coefficient K_d for each batch sorption experiment allowed for a quantitative comparison with previous studies of NOM sorption onto iron oxides and other sediment minerals. We conclude that mackinawite tends to demonstrate a lower sorption affinity for NOM than iron oxides, but that the chemical characteristics of NOM favoured by mackinawite are consistent with more labile compounds such as proteins and polysaccharides. If the proportion of reduced iron species in marine sediments is set to increase with climate change, it remains to be seen which of these conflicting influences on carbon sequestration will predominate.

As mentioned in Section 2.3.2, the FTIR spectra of the LMW NOM fraction was dominated by inorganic salts. Unfortunately, despite our best efforts to background correct for these salts, we were unable to obtain reliable spectra of the LMW NOM. Going forward, the development of a method for isolating the LMW fraction without salt interference would be beneficial for modelling and prediction purposes. However, due to the relatively low K_d values of the LMW compounds, we feel that it is sufficient for this work to concentrate on the HMW fraction which has the largest impact on carbon burial.

Sorption affinities from several studies were compiled in Tables 2.4 and 2.5, providing to our knowledge the first cross-study quantitative comparison of sorption affinities for sedimentary minerals. Furthermore, it is an initial attempt to make use of the myriad isotherm fits for NOM sorption onto minerals that exist in the literature. Due to the high variability in sorption experiments and sensitivity to physico-chemical parameters, in the future a larger dataset would be useful for a better identification of carbon sequestration trends across differing environments. This would involve obtaining additional K_d values for various pHs and mineral types. It must be remembered that batch sorption experiments, both in this research and in the literature, are conducted in a laboratory setting and therefore accurate extrapolation to real marine environments must be done with caution. The non-exhaustive list of variables which affect equilibrium parameters, such as temperature and ionic strength as outlined in

Section 1.7, would have to be directly probed in order to make accurate predictions of equilibria in marine pore waters. The present research is an initial step towards a quantitative comparison of equilibria across laboratory studies, with implications for natural environments.

Finally, we have determined which functional groups serve as predictors of increased K_d for the sorption of NOM onto mackinawite. The causal link between functional group and sorption, however, must be explored more directly. This will likely involve a more controlled pool of sorbates, model compounds possessing limited functional groups of either hydroxyls, carboxyls or amides only. In this way, the binding mechanism as proposed in Section 2.4.6 can be interrogated as well. This work presents a novel step towards understanding the role of reduced iron sulfides in carbon sequestration. The consequences of climate change for sediment redox chemistry already being observed indicate that reduced iron species are likely to play a greater role in the future. Further studies on iron sulfides, both in controlled laboratory environments and from natural sediments, are required to better predict the consequences of climate change on carbon cycling through marine environments.

Bibliography

Abdulla H. A. N., Minor E. C., Dias R. F. and Hatcher P. G. (2010) Changes in the compound classes of dissolved organic matter along an estuarine transect: A study using FTIR and 13C NMR. *Geochim. Cosmochim. Acta* **74**, 3815–3838.

Alperin M. J., Martens C. S., Albert D. B., Suayah I. B., Benninger L. K., Blair N. E. and Jahnke R. A. (1999) Benthic fluxes and porewater concentration profiles of dissolved organic carbon in sediments from the North Carolina continental slope. *Geochim. Cosmochim. Acta* **63**, 427–448.

Armstrong R. A., Lee C., Hedges J. I., Honjo S. and Wakeham S. G. (2001) A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **49**, 219–236.

Arnarson T. S. and Keil R. G. (2000) Mechanisms of pore water organic matter adsorption to montmorillonite. *Mar. Chem.* **71**, 309–320.

Arndt S., Jørgensen B. B., LaRowe D. E., Middelburg J. J., Pancost R. D. and Regnier P. (2013) Quantifying the degradation of organic matter in marine sediments: A review and synthesis. *Earth-Sci. Rev.* **123**, 53–86.

Artz R. R. E., Chapman S. J., Jean Robertson A. H., Potts J. M., Laggoun-Défarge F., Gogo S., Comont L., Disnar J.-R. and Francez A.-J. (2008) FTIR spectroscopy can be used as a screening tool for organic matter quality in regenerating cutover peatlands. *Soil Biol. Biochem.* **40**, 515–527.

Atwood T. B., Witt A., Mayorga J., Hammill E. and Sala E. (2020) Global patterns in marine sediment carbon stocks. *Front. Mar. Sci.* **7**, 165.

Ausili A., Sánchez M. and Gómez-Fernández J. C. (2015) Attenuated total reflectance infrared spectroscopy: A powerful method for the simultaneous study of structure and spatial orientation of lipids and membrane proteins. *Biomed. Spectrosc. Imaging* **4**, 159–170.

Axe K., Vejgården M. and Persson P. (2006) An ATR-FTIR spectroscopic study of the competitive adsorption between oxalate and malonate at the water–goethite interface. *J. Colloid Interface Sci.* **294**, 31–37.

Balind K. (2019) Exploring the Affinity and Selectivity of Sedimentary Mackinawite (FeS) Towards Natural Organic Matter. Doctoral dissertation, Concordia University.

Barber A., Brandes J., Leri A., Lalonde K., Balind K., Wirick S., Wang J. and Gélinas Y. (2017) Preservation of organic matter in marine sediments by inner-sphere interactions with reactive iron. *Sci. Rep.* **7**, 1–10.

Barber A., Lalonde K., Mucci A. and Gélinas Y. (2014) The role of iron in the diagenesis of organic carbon and nitrogen in sediments: A long-term incubation experiment. *Mar. Chem.* **162**, 1–9.

Bauer J. E., Cai W.-J., Raymond P. A., Bianchi T. S., Hopkinson C. S. and Regnier P. A. G. (2013) The

changing carbon cycle of the coastal ocean. *Nature* **504**, 61–70.

Berner R. A. (2003) The long-term carbon cycle, fossil fuels and atmospheric composition. *Nature* **426**, 323–326.

Boily J.-F., Persson P. and Sjöberg S. (2000) Benzenecarboxylate surface complexation at the goethite (α-FeOOH)/water interface: II. Linking IR spectroscopic observations to mechanistic surface complexation models for phthalate, trimellitate, and pyromellitate. *Geochim. Cosmochim. Acta* **64**, 3453–3470.

Boughriet A., Figueiredo R. S., Laureyns J. and Recourt P. (1997) Identification of newly generated iron phases in recent anoxic sediments: 57Fe Mössbauer and microRaman spectroscopic studies. *J. Chem. Soc. Faraday Trans.* **93**, 3209–3215.

Boukir A., Fellak S. and Doumenq P. (2019) Structural characterization of Argania spinosa Moroccan wooden artifacts during natural degradation progress using infrared spectroscopy (ATR-FTIR) and X-Ray diffraction (XRD). *Heliyon* **5**, e02477.

Bourdoiseau J.-A., Jeannin M., Sabot R., Rémazeilles C. and Refait Ph. (2008) Characterisation of mackinawite by Raman spectroscopy: Effects of crystallisation, drying and oxidation. *Corros. Sci.* **50**, 3247–3255.

Burdige D. J. (2007) Preservation of organic matter in marine sediments: controls, mechanisms, and an imbalance in sediment organic carbon budgets? *Chem. Rev.* **107**, 467–485.

Chen C., Dynes J. J., Wang J. and Sparks D. L. (2014) Properties of Fe-organic matter associations via coprecipitation versus adsorption. *Environ. Sci. Technol.* **48**, 13751–13759.

Chisholm S. W. (2000) Stirring times in the Southern Ocean. *Nature* **407**, 685–686.

Christensen T. H., Bjerg P. L., Banwart S. A., Jakobsen R., Heron G. and Albrechtsen H.-J. (2000) Characterization of redox conditions in groundwater contaminant plumes. *J. Contam. Hydrol.* **45**, 165–241.

Coward E. K., Ohno T. and Plante A. F. (2018) Adsorption and molecular fractionation of dissolved organic matter on iron-bearing mineral matrices of varying crystallinity. *Environ. Sci. Technol.* **52**, 1036–1044.

Coward E. K., Ohno T. and Sparks D. L. (2019) Direct evidence for temporal molecular fractionation of dissolved organic matter at the iron oxyhydroxide interface. *Environ. Sci. Technol.* **53**, 642–650.

Dai D. and Fan M. (2011) Investigation of the dislocation of natural fibres by Fourier-transform infrared spectroscopy. *Vib. Spectrosc.* **55**, 300–306.

Dean A. P., Sigee D. C., Estrada B. and Pittman J. K. (2010) Using FTIR spectroscopy for rapid determination of lipid accumulation in response to nitrogen limitation in freshwater microalgae. *Bioresour. Technol.* **101**, 4499–4507.

Devey A. J., Grau-Crespo R. and de Leeuw N. H. (2008) Combined density functional theory and interatomic potential study of the bulk and surface structures and properties of the iron sulfide mackinawite (FeS). *J. Phys. Chem. C* **112**, 10960–10967.

Diaz R. J. and Rosenberg R. (2008) Spreading dead zones and consequences for marine ecosystems. *Science* **321**, 926–929.

Dos Santos Afonso M. and Stumm W. (1992) Reductive dissolution of iron(III) (hydr)oxides by hydrogen sulfide. *Langmuir* **8**, 1671–1675.

Ergin M., Gaines A., Galletti G. C., Chiavari G., Fabbri D. and Yücesoy-Eryilmaz F. (1996) Early diagenesis of organic matter in recent Black Sea sediments: characterization and source assessment. *Appl. Geochem.* **11**, 711–720.

Estes E. R., Pockalny R., D'Hondt S., Inagaki F., Morono Y., Murray R. W., Nordlund D., Spivack A. J., Wankel S. D., Xiao N. and Hansel C. M. (2019) Persistent organic matter in oxic subseafloor sediment. *Nat. Geosci.* **12**, 126–131.

Evangelou V. P., Marsi M. and Chappell M. A. (2002) Potentiometric–spectroscopic evaluation of metal-ion complexes by humic fractions extracted from corn tissue. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **58**, 2159–2175.

Falkowski P., Scholes R. J., Boyle E., Canadell J., Canfield D., Elser J., Gruber N., Hibbard K., Högberg P., Linder S., Mackenzie F. T., Moore III B., Pedersen T., Rosenthal Y., Seitzinger S., Smetacek V. and Steffen W. (2000) The global carbon cycle: a test of our knowledge of earth as a system. *Science* **290**, 291–296.

Feng X., Simpson A. J. and Simpson M. J. (2005) Chemical and mineralogical controls on humic acid sorption to clay mineral surfaces. *Org. Geochem.* **36**, 1553–1566.

Filius J. D., Hiemstra T. and Van Riemsdijk W. H. (1997) Adsorption of small weak organic acids on goethite: modeling of mechanisms. *J. Colloid Interface Sci.* **195**, 368–380.

Fu P., Wu F., Liu C.-Q., Wei Z., Bai Y. and Liao H. (2006) Spectroscopic characterization and molecular weight distribution of dissolved organic matter in sediment porewaters from Lake Erhai, Southwest China. *Biogeochemistry* **81**, 179–189.

Gao K., Beardall J., Häder D.-P., Hall-Spencer J. M., Gao G. and Hutchins D. A. (2019) Effects of ocean acidification on marine photosynthetic organisms under the concurrent influences of warming, UV radiation, and deoxygenation. *Front. Mar. Sci.* **6**, 322.

Garrels R., Lerman A. and Mackenzie F. (1976) Controls of atmospheric O2 and CO2: past, present, and future: geochemical models of the earth's surface environment, focusing on O2 and CO2 cycles, suggest that a dynamic steady-state system exists, maintained over time by effective feedback mechanisms. *Am. Sci.*, 306–315.

Gelin F., Boogers I., Noordeloos A. A. M., Damsté J. S. S., Hatcher P. G. and Leeuw J. W. de (1996) Novel, resistant microalgal polyethers: An important sink of organic carbon in the marine

environment? *Geochim. Cosmochim. Acta* **60**, 1275–1280.

Gilbert D., Sundby B., Gobeil C., Mucci A. and Tremblay G.-H. (2005) A seventy-two-year record of diminishing deep-water oxygen in the St. Lawrence estuary: The northwest Atlantic connection. *Limnol. Oceanogr.* **50**, 1654–1666.

Goueguel C. L. (2021) *HotellingEllipse: Hotelling T-Square and Confidence Ellipse.*,

Gu B., Schmitt J., Chen Z., Liang L. and McCarthy J. F. (1994) Adsorption and desorption of natural organic matter on iron oxide: mechanisms and models. *Environ. Sci. Technol.* **28**, 38–46.

Hansel C. M., Benner S. G. and Fendorf S. (2005) Competing Fe(II)-induced mineralization pathways of ferrihydrite. *Environ. Sci. Technol.* **39**, 7147–7153.

Hatcher P. G., Ravin A., Behar F. and Baudin F. (2014) Diagenesis of organic matter in a 400 m organic rich sediment core from offshore Namibia using solid state 13C NMR and FTIR. *Org. Geochem.* **75**, 8–23.

He X., Xi B., Wei Z., Guo X., Li M., An D. and Liu H. (2011) Spectroscopic characterization of water extractable organic matter during composting of municipal solid waste. *Chemosphere* **82**, 541–548.

Hedges J. I. and Keil R. G. (1995) Sedimentary organic matter preservation: an assessment and speculative synthesis. *Mar. Chem.* **49**, 81–115.

Hedges J. I., Mayorga E., Tsamakis E., McClain M. E., Aufdenkampe A., Quay P., Richey J. E., Benner R., Opsahl S., Black B., Pimentel T., Quintanilla J. and Maurice L. (2000) Organic matter in Bolivian tributaries of the Amazon River: A comparison to the lower mainstream. *Limnol. Oceanogr.* **45**, 1449–1466.

Hedges J. I. and Oades J. M. (1997) Comparative organic geochemistries of soils and marine sediments. *Org. Geochem.* **27**, 319–361.

Hertkorn N., Benner R., Frommberger M., Schmitt-Kopplin P., Witt M., Kaiser K., Kettrup A. and Hedges J. I. (2006) Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **70**, 2990–3010.

Howe K. J., Ishida K. P. and Clark M. M. (2002) Use of ATR/FTIR spectrometry to study fouling of microfiltration membranes by natural waters. *Desalination* **147**, 251–255.

Ibarra J. V., Muñoz E. and Moliner R. (1996) FTIR study of the evolution of coal structure during the coalification process. *Org. Geochem.* **24**, 725–735.

IPCC (2021) *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. eds. V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, Cambridge University Press, In Press.

de Jong S. (1993) SIMPLS: An alternative approach to partial least squares regression. *Chemom. Intell. Lab. Syst.* **18**, 251–263.

Jouraiphy A., Amir S., Winterton P., El Gharous M., Revel J.-C. and Hafidi M. (2008) Structural study of the fulvic fraction during composting of activated sludge–plant matter: Elemental analysis, FTIR and 13C NMR. *Bioresour. Technol.* **99**, 1066–1072.

Jutras M., Dufour C. O., Mucci A., Cyr F. and Gilbert D. (2020) Temporal changes in the causes of the observed oxygen decline in the St. Lawrence estuary. *J. Geophys. Res. Oceans* **125**, e2020JC016577.

Kaiser K. and Guggenberger G. (2000) The role of DOM sorption to mineral surfaces in the preservation of organic matter in soils. *Org. Geochem.* **31**, 711–725.

Keil R. G. and Hedges J. I. (1993) Sorption of organic matter to mineral surfaces and the preservation of organic matter in coastal marine sediments. *Chem. Geol.* **107**, 385–388.

Keil R. G., Montluçon D. B., Prahl F. G. and Hedges J. I. (1994) Sorptive preservation of labile organic matter in marine sediments. *Nature* **370**, 549–552.

Kennedy M., Droser M., Mayer L. M., Pevear D. and Mrofka D. (2006) Late precambrian oxygenation; inception of the clay mineral factory. *Science* **311**, 1446–1449.

Kirk K. T. (1984) Degradation of Lignin. In *Biochemistry of Microbial Degradation* (ed. D. T. Gibson). Microbiology Series. Marcel Dekker Inc., New York. pp. 399–437.

Kosmulski M., Maczka E., Jartych E. and Rosenholm Jarl. B. (2003) Synthesis and characterization of goethite and goethite–hematite composite: experimental study and literature survey. *Adv. Colloid Interface Sci.* **103**, 57–76.

Krupka K. M., Kaplan D. I., Whelan G., Serne R. J. and Mattigod S. V. (1999) Understanding variation in partition coefficient, Kd, values. Volume 1: The Kd model, methods of measurement, and application of chemical reaction codes. *Environ. Prot. Agency Wash. DC USA*.

Kuhn M. (2021) *caret: Classification and Regression Training*.,

Kunov-Kruse A. J., Riisager A., Saravanamurugan S., Berg R. W., Kristensen S. B. and Fehrmann R. (2013) Revisiting the Brønsted acid catalysed hydrolysis kinetics of polymeric carbohydrates in ionic liquids by in situ ATR-FTIR spectroscopy. *Green Chem.* **15**, 2843–2848.

Lalonde K., Mucci A., Ouellet A. and Gélinas Y. (2012) Preservation of organic matter in sediments promoted by iron. *Nature* **483**, 198–200.

LaRowe D. E., Arndt S., Bradley J. A., Estes E. R., Hoarfrost A., Lang S. Q., Lloyd K. G., Mahmoudi N., Orsi W. D., Shah Walter S. R., Steen A. D. and Zhao R. (2020) The fate of organic carbon in marine sediments - New insights from recent data and analysis. *Earth-Sci. Rev.* **204**, 103146.

Laskowski J. S., Liu Q. and O'Connor C. T. (2007) Current understanding of the mechanism of

polysaccharide adsorption at the mineral/aqueous solution interface. *Int. J. Miner. Process.* **84**, 59–68.

Lee C., Wakeham S. and Arnosti C. (2004) Particulate organic matter in the sea: the composition conundrum. *AMBIO J. Hum. Environ.* **33**, 565–575.

Lennie A. R., Redfern S. A. T., Schofield P. F. and Vaughan D. J. (1995) Synthesis and Rietveld crystal structure refinement of mackinawite, tetragonal FeS. *Mineral. Mag.* **59**, 677–683.

Li Y., Yu S., Strong J. and Wang H. (2012) Are the biogeochemical cycles of carbon, nitrogen, sulfur, and phosphorus driven by the "FeIII–FeII redox wheel" in dynamic redox environments? *J. Soils Sediments* **12**, 683–693.

Limousin G., Gaudet J.-P., Charlet L., Szenknect S., Barthès V. and Krimissa M. (2007) Sorption isotherms: A review on physical bases, modeling and measurement. *Appl. Geochem.* **22**, 249–275.

Liu J., Wang Z., Belchik S. M., Edwards M. J., Liu C., Kennedy D. W., Merkley E. D., Lipton M. S., Butt J. N., Richardson D. J., Zachara J. M., Fredrickson J. K., Rosso K. M. and Shi L. (2012) Identification and characterization of MtoA: a decaheme c-type cytochrome of the neutrophilic Fe(II)-oxidizing bacterium Sideroxydans lithotrophicus ES-1. *Front. Microbiol.* **3**, 37.

Liu Q., Zhang Y. and Laskowski J. S. (2000) The adsorption of polysaccharides onto mineral surfaces: an acid/base interaction. *Int. J. Miner. Process.* **60**, 229–245.

Lovley D. R. and Anderson R. T. (2000) Influence of dissimilatory metal reduction on fate of organic and metal contaminants in the subsurface. *Hydrogeol. J.* **8**, 77–88.

Lützow M. v, Kögel Knabner I., Ekschmitt K., Matzner E., Guggenberger G., Marschner B. and Flessa H. (2006) Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions – a review. *Eur. J. Soil Sci.* **57**, 426–445.

Lv J., Zhang S., Wang S., Luo L., Cao D. and Christie P. (2016) Molecular-scale investigation with ESI-FT-ICR-MS on fractionation of dissolved organic matter induced by adsorption on iron oxyhydroxides. *Environ. Sci. Technol.* **50**, 2328–2336.

Lyman D. J., Benck R., Dell S., Merle S. and Murray-Wijelath J. (2003) FTIR-ATR analysis of brewed coffee: effect of roasting conditions. *J. Agric. Food Chem.* **51**, 3268–3272.

Margenot A. J., Calderón F. J., Bowles T. M., Parikh S. J. and Jackson L. E. (2015) Soil organic matter functional group composition in relation to organic carbon, nitrogen, and phosphorus fractions in organically managed tomato fields. *Soil Sci. Soc. Am. J.* **79**, 772–782.

Mayers J. J., Flynn K. J. and Shields R. J. (2013) Rapid determination of bulk microalgal biochemical composition by Fourier-Transform Infrared spectroscopy. *Bioresour. Technol.* **148**, 215–220.

McKnight D. M., Bencala K. E., Zellweger G. W., Aiken G. R., Feder G. L. and Thorn K. A. (1992) Sorption of dissolved organic carbon by hydrous aluminum and iron oxides occurring at the

confluence of Deer Creek with the Snake River, Summit County, Colorado. *Environ. Sci. Technol.* **26**, 1388–1396.

Meier M., Namjesnik-Dejanovic K., Maurice P. A., Chin Y.-P. and Aiken G. R. (1999) Fractionation of aquatic natural organic matter upon sorption to goethite and kaolinite. *Chem. Geol.* **157**, 275–284.

Menges F. (2020) *Spectragryph–optical spectroscopy software*.,

Mevik B.-H., Wehrens R. and Liland K. H. (2020) *pls: Partial Least Squares and Principal Component Regression*.,

Meybeck M. (1982) Carbon, nitrogen, and phosphorus transport by world rivers. *Am. J. Sci.* **282**, 401–450.

Mikutta R., Mikutta C., Kalbitz K., Scheel T., Kaiser K. and Jahn R. (2007) Biodegradation of forest floor organic matter bound to minerals via different binding mechanisms. *Geochim. Cosmochim. Acta* **71**, 2569–2590.

Ohno T., Sleighter R. L. and Hatcher P. G. (2018) Adsorptive fractionation of corn, wheat, and soybean crop residue derived water-extractable organic matter on iron (oxy)hydroxide. *Geoderma* **326**, 156–163.

Parikh S. J., Goyne K. W., Margenot A. J., Mukome F. N. D. and Calderón F. J. (2014) Soil Chemical Insights Provided through Vibrational Spectroscopy. In *Advances in Agronomy* Elsevier, New York, NY. pp. 1–148.

Pedersen H. D., Postma D., Jakobsen R. and Larsen O. (2005) Fast transformation of iron oxyhydroxides by the catalytic action of aqueous Fe(II). *Geochim. Cosmochim. Acta* **69**, 3967–3977.

Persson P. and Axe K. (2005) Adsorption of oxalate and malonate at the water-goethite interface: Molecular surface speciation from IR spectroscopy. *Geochim. Cosmochim. Acta* **69**, 541–552.

Picard A., Gartman A. and Girguis P. R. (2016) What do we really know about the role of microorganisms in iron sulfide mineral formation? *Front. Earth Sci.* **4**, doi: 10.3389/feart.2016.00068.

Plácido J., Imam T. and Capareda S. (2013) Evaluation of ligninolytic enzymes, ultrasonication and liquid hot water as pretreatments for bioethanol production from cotton gin trash. *Bioresour. Technol.* **139**, 203–208.

Plaza C., Senesi N., Polo A., Brunetti G., García-Gil J. C. and D'Orazio V. (2003) Soil fulvic acid properties as a means to assess the use of pig slurry amendment. *Soil Tillage Res.* **74**, 179–190.

R Core Team (2021) *R: A Language and Environment for Statistical Computing*., R Foundation for Statistical Computing, Vienna, Austria.

Ransom B., Kim D., Kastner M. and Wainwright S. (1998) Organic matter preservation on continental slopes: importance of mineralogy and surface area. *Geochim. Cosmochim. Acta* **62**, 1329–1345.

Rickard D. (2006) The solubility of FeS. *Geochim. Cosmochim. Acta* **70**, 5779–5789.

Rickard D. and Luther G. W. (2007) Chemistry of iron sulfides. *Chem. Rev.* **107**, 514–562.

Rickard D. and Morse J. W. (2005) Acid volatile sulfide (AVS). *Mar. Chem.* **97**, 141–197.

Riedel T., Zak D., Biester H. and Dittmar T. (2013) Iron traps terrestrially derived dissolved organic matter at redox interfaces. *Proc. Natl. Acad. Sci.* **110**, 10101–10105.

Roux-Michollet D., Dudal Y., Jocteur-Monrozier L. and Czarnes S. (2010) Steam treatment of surface soil: how does it affect water-soluble organic matter, C mineralization, and bacterial community composition? *Biol. Fertil. Soils* **46**, 607–616.

Salvestrini S., Leone V., Iovino P., Canzano S. and Capasso S. (2014) Considerations about the correct evaluation of sorption thermodynamic parameters from equilibrium isotherms. *J. Chem. Thermodyn.* **68**, 310–316.

Sigman D. M. and Hain M. P. (2012) The Biological Productivity of the Ocean. *Nat. Educ. Knowl.* **3**, 21.

Signal developers (2021) *signal: Signal processing*.,

Stehfest K., Toepel J. and Wilhelm C. (2005) The application of micro-FTIR spectroscopy to analyze nutrient stress-related changes in biomass composition of phytoplankton algae. *Plant Physiol. Biochem.* **43**, 717–726.

Stone A. G., Traina S. J. and Hoitink H. A. J. (2001) Particulate organic matter composition and pythium damping-off of cucumber. *Soil Sci. Soc. Am. J.* **65**, 761–770.

Sun C.-H., Zhu M.-X., Ma W.-W., Sun Z.-L., Zhang X.-R., Ding K.-Y. and Liu S.-H. (2020) Examining bulk and iron-associated organic carbon through depth in margin sea sediments (China) under contrasting depositional settings: Chemical and NEXAFS spectral characterization. *J. Mar. Syst.* **207**, 103344.

Tejedor-Tejedor M. I., Yost E. C. and Anderson M. A. (1992) Characterization of benzoic and phenolic complexes at the goethite/aqueous solution interface using cylindrical internal reflection Fourier transform infrared spectroscopy. 2. Bonding structures. *Langmuir* **8**, 525–533.

Tejedor-Tejedor M. I., Yost E. C. and Anderson M. A. (1990) Characterization of benzoic and phenolic complexes at the goethite/aqueous solution interface using cylindrical internal reflection Fourier transform infrared spectroscopy. Part 1. Methodology. *Langmuir* **6**, 979–987.

Thimsen C. A. and Keil R. G. (1998) Potential interactions between sedimentary dissolved organic matter and mineral surfaces. *Mar. Chem.* **62**, 65–76.

Thomas J. E. and Kelley M. J. (2008) Interaction of mineral surfaces with simple organic molecules by

diffuse reflectance IR spectroscopy (DRIFT). *J. Colloid Interface Sci.* **322**, 516–526.

Vannote R. L., Minshall G. W., Cummins K. W., Sedell J. R. and Cushing C. E. (1980) The river continuum concept. *Can. J. Fish. Aquat. Sci.* **37**, 130–137.

Wang Y., Zhang Z., Han L., Sun K., Jin J., Yang Yu, Yang Yan, Hao Z., Liu J. and Xing B. (2019) Preferential molecular fractionation of dissolved organic matter by iron minerals with different oxidation states. *Chem. Geol.*

Wittekindt C. and Marx D. (2012) Water confined between sheets of mackinawite FeS minerals. *J. Chem. Phys.* **137**, 054710.

Wold S., Sjöström M. and Eriksson L. (2001) PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **58**, 109–130.

Wolthers M., Charlet L., van Der Linde P. R., Rickard D. and van Der Weijden C. H. (2005) Surface chemistry of disordered mackinawite (FeS). *Geochim. Cosmochim. Acta* **69**, 3469–3481.

Yang Y., Duan J. and Jing C. (2013) Molecular-scale study of salicylate adsorption and competition with catechol at goethite/aqueous solution interface. *J. Phys. Chem. C* **117**, 10597–10606.

Yost E. C., Tejedor-Tejedor M. I. and Anderson M. A. (1990) In situ CIR-FTIR characterization of salicylate complexes at the goethite/aqueous solution interface. *Environ. Sci. Technol.* **24**, 822–828.

Zhao B., Yao P., Bianchi T. S., Shields M. R., Cui X. Q., Zhang X. W., Huang X. Y., Schröder C., Zhao J. and Yu Z. G. (2018) The role of reactive iron in the preservation of terrestrial organic carbon in estuarine sediments. *J. Geophys. Res. Biogeosciences* **123**, 3556–3569.

Appendix

Table A1. Elemental analysis results for NOM endmembers and sediment extracted NOM.

Size Fraction	NOM Source	C/N	%OC
HMW	Soil	14.1 (0.4)	32.0 (1.3)
	Plankton	7.3 (0.03)	29.1 (0.8)
	Corn	21.8 (5.9)	29.8 (1.2)
LMW	Soil	0.2 (0.08)	1.4 (0.3)
	Plankton	8.7 (0.3)	20.8 (1.6)
	Corn	14.6 (2.9)	4.3 (0.2)
Unfractionated	Sediment SAG-05	17.3 (0.9)	0.5 (0.004)
Unfractionated	Sediment STN 19	11.3 (1.2)	0.3 (0.009)

C/N: Average carbon to nitrogen ratio; %OC: average percent organic carbon of dried NOM by mass, as determined by mass spectrometry. Standard deviations in parentheses.

Hotelling's T^2 measures the distance of each data point's scores to the centre of the hyperplane defined by the PLSR model's LVs. At 2 LVs, the hyperplane forms an ellipse circumscribed by the 95% confidence limit as determined by the *t*-distribution. With normalized scores, as in the case of PLSR, the ellipse becomes circular. When plotting the scores of LV 2 vs. LV 1 in Figure A.1, we see that two samples consisting of HMW corn endmembers may act as leverage points, particularly in the first LV. However their T^2 values fall well below the 95% confidence level cutoff of 7.85.

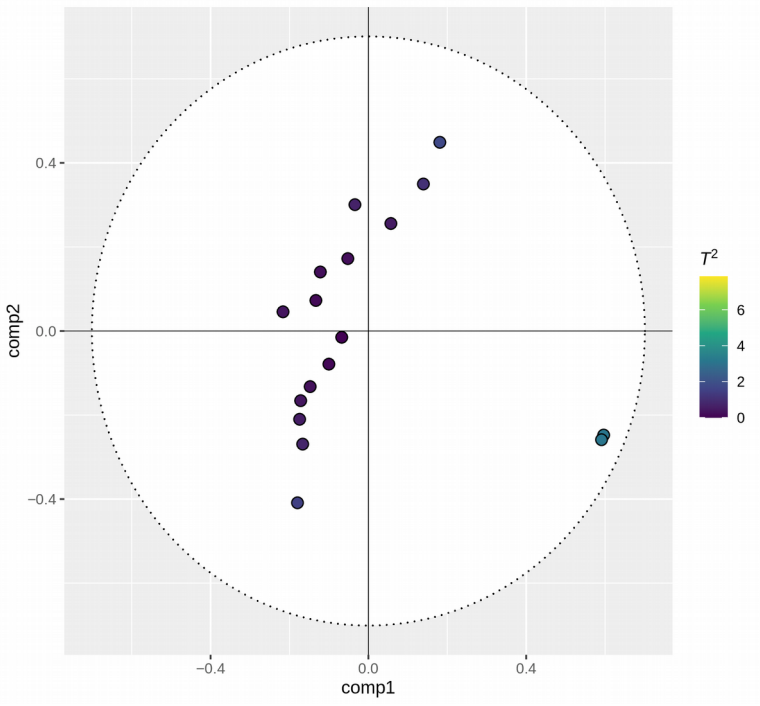

Figure A.1. The dotted perimeter represents the 95% confidence limit for each of LV/component 1 and 2. Shading represents the Hotelling's T^2 value for each individual observation, with cutoff value calculated at 7.85 for the 95% confidence level.