
Georgina Anne McKee
Old Dominion University

Recommended Citation
https://digitalcommons.odu.edu/chemistry_etds/39

This Dissertation is brought to you for free and open access by the Chemistry & Biochemistry at ODU Digital Commons. It has been accepted for inclusion in Chemistry & Biochemistry Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
THE NATURE, ORIGIN AND PRESERVATION OF
AMIDE ORGANIC NITROGEN IN ORGANIC MATTER

by

Georgina Anne McKee
M. Chem. July 2006, University of Sheffield, United Kingdom

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

CHEMISTRY

OLD DOMINION UNIVERSITY
August 2011

Approved by:

Patrick G. Hatcher (Director)

Andrew S. Gordon (Member)

Thomas Isenhour (Member)

Jingdong Mao (Member)
ABSTRACT

THE NATURE, ORIGIN AND PRESERVATION OF AMIDE ORGANIC NITROGEN IN ORGANIC MATTER

Georgina Anne McKee
Old Dominion University, 2011
Director: Dr. Patrick G. Hatcher

Past studies have found much of nonliving sedimentary/aqueous nitrogen-containing organic matter (NCOM) is composed of amides, assumed to be peptides/proteins. Their lability calls into question their survival, and several hypotheses have been suggested to explain this. Using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) and nuclear magnetic resonance (NMR) spectroscopy I investigated the molecular amide composition in sedimentary/aqueous systems while reassessing their preservation and formation routes.

Development of a suitable methodology is essential for successful NCOM study due to electrospray ionisation source requirements: sediment samples need to be rendered into solution. Based on NMR and FT-ICR-MS analysis, I determined that pyridine extraction was optimal. Additionally, I developed a graphical method to visualise the important components of NCOM finding that N/H versus O/C was most suitable.

Extracting with pyridine, I investigated molecular forms of NCOM in a lacustrine sediment (Mangrove Lake, Bermuda, MLB). With 2D NMR and FT-ICR-MS, I identified alkyl (di)amides, and confirmed their structure by gas chromatography MS. I suggest a formation pathway through amidation of naturally occurring glyceride esters for these newly discovered compounds. FT-ICR-MS analysis of the Lake and pore water NCOM show alkyl (di)amide prevalence throughout the Lake, and the increasing ammonia contents with depth correlates with increasing alkyl (di)amide contents.
downcore supporting an amidation process.

Peptide preservation hypotheses were additionally examined by synthesising designer peptides, and monitoring their fate in two natural waters, MLB and Elizabeth River, VA. Results indicate peptide length, not chemical structure, is important for determining removal rates. Additionally, removal rates are significantly affected by location, probably because of micro-organism community differences and/or adduction to naturally present Michael receptors (e.g., quinones). The adduction of a quinone to designer peptides was found to have no effect on peptide removal rates. New compounds are formed from peptide-adducts seem to incorporate parts of the peptide as a stable product demonstrating a new route for partial peptide preservation in natural waters.

Overall, advanced analytical techniques have led to new insights into amide N composition, providing further evidence of likely formation and preservation pathways in natural systems.
This dissertation is dedicated to my family.
ACKNOWLEDGMENTS

The completion of this dissertation was only possible by the guidance and support of many people. I would firstly like to thank my advisor, Dr. Hatcher, who suggested I embark upon the journey of completing a PhD, and for being there for me every step of the way. His patient teaching of all the skills I have learnt as being a graduate student in his group will immensely assist my future research. I am especially grateful for the opportunity to learn and become adept at a wide variety of techniques not available to many graduate students. In addition, his careful and thoughtful critic of my writing has much improved my style. The high level of funding I received throughout my degree, enabled me to attend many conferences, teaching me about effective presentation skills and allowing me to network with my peers. He has provided me with an amazing start to my career, and I will always be grateful. I would also like to thank my committee members, Dr. Isenhour, Dr. Gordon and Dr. Mao for their continued advice throughout my degree, and for dedicating many hours to the improvement of this document.

Completion of my degree would not have been possible without the support of my research group. I would like to thank each of them for their advice through the years that have enabled me to pursue my degree. I would like to thank in particular, Dr. Rachel Sleighter who has spent many hours answering my questions and assisting in all aspects of my research. Her friendly face when I first arrived at ODU made me feel welcome and at home immediately in my new state of Virginia. I would also like to Dr. Zhanfei Liu who worked with me to learn the intricacies of peptide synthesis from scratch and advised me on many topics, both related to my research and on my personal life. Another individual who has been an incredible help for my research is Megan Kobiela, whose
outstanding organisation and planning skills made much of the syntheses and analyses of
the compounds used in this dissertation possible. The amazing dynamic of the Hatcher
group has made me feel so grateful for coming to ODU and consequently, I am able to
call so many my friends. I will never forget the annual ‘purple elephant’ parties, the
Larchmont gatherings, intramural football and the annual writing retreats. My time at
ODU truly will be remembered with fond and happy memories.

There are many additional people I would like to thank who have made my
research possible with their advice and assistance including Dr. Elodie Salmon, and Dr.
Hussain Abdulla. I particularly thank Susan Hatcher of the College of Sciences major
instrument cluster (COSMIC) whose patient answering of my many questions and
assistance with all aspects of FT-ICR-MS analysis enabled me to be successful in my
research. I would also like to thank Mahasilu Amunugama for his assistance with FT-
ICR-MS and Dr. Junyan Zhong for teaching me how to analyse my samples and
answering my questions related to nuclear magnetic resonance (NMR) spectroscopy.
Particular thanks go to Dr. Aron Stubbins and Joshua Marsh for accompanying me on the
field trip to Bermuda, and their help in collecting my samples. I would also like to thank
Dr. Mopper and Dr. Mullholland for the occasional use of the instruments in their labs.
In addition, I would like to thank many individuals at the Bermuda Institute of Ocean
Sciences (BIOS), who greatly assisted in all aspects of collecting my samples in Bermuda
in particular Jane Burrows, Steven Bell, and Bruce Williams. The use of one of their
labs, and many pieces of their equipment allowed me to collect and process my samples
within my allotted time efficiently. I would also like to acknowledge the R/V HSBC
Atlantic Explorer for use of their freezer to transport some of my samples back to ODU.
Many people of USDA assisted in obtaining the appropriate permits that allowed me to
transport my samples back to ODU. In addition, many companies that I worked with in
organising the trip to Bermuda provided the necessary equipment I needed and I am
grateful for their services. I would also like to thank the Dean of the College of Sciences
for providing much of the funding for the Bermuda trip. Funding for my research was
also provided by the Frank Batten Endowment Fund and the National Science Foundation
Grant number OCE-0726759.

Lastly I would like to thank my family for providing the support I needed to
complete my degree. Their encouragement during the many difficult times I
countered, gave me the strength I needed to continue. My sister, Sarah, in particular
has always been my inspiration to be a success in all aspects of my life with her amazing
examples of courage when times are challenging. I have no doubt that my husband,
Steven, has made it possible for me to become the person I am today by his continuous
support. His ability to still appear enthusiastic about my research on a daily basis even if
he has heard it thousands of times before makes me appreciate how important he is for
achieving my life goals. He is an unwavering supporter of all my work even if it means
that he has to make sacrifices for me. I am truly lucky to have such a fantastic support
network that I can rely on in all circumstances.
## NOMENCLATURE

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$% C_e$</td>
<td>Percent of C in the extract</td>
</tr>
<tr>
<td>$% C_r$</td>
<td>Percent of C in the residue</td>
</tr>
<tr>
<td>$% C_s$</td>
<td>Percent of C in the original sediment</td>
</tr>
<tr>
<td>$% N_e$</td>
<td>Percent of N in the extract</td>
</tr>
<tr>
<td>$% N_r$</td>
<td>Percent of N in the residue</td>
</tr>
<tr>
<td>$% N_s$</td>
<td>Percent of N in the original sediment</td>
</tr>
<tr>
<td>AAFRV</td>
<td>Alanine-alanine-phenylalanine-arginine-valine</td>
</tr>
<tr>
<td>AAVF</td>
<td>Alanine-alanine-valine-phenylalanine</td>
</tr>
<tr>
<td>AFRV</td>
<td>Alanine-phenylalanine-arginine-valine</td>
</tr>
<tr>
<td>AFVA</td>
<td>Alanine-phenylalanine-valine-alanine</td>
</tr>
<tr>
<td>AVFA</td>
<td>Alanine-valine-phenylalanine-alanine</td>
</tr>
<tr>
<td>BG</td>
<td>Bayou Grande, Pensacola, Florida</td>
</tr>
<tr>
<td>BIOS</td>
<td>Bermuda Institute of Ocean Sciences</td>
</tr>
<tr>
<td>C</td>
<td>Carbon</td>
</tr>
<tr>
<td>COSMIC</td>
<td>College of Sciences major instrument cluster</td>
</tr>
<tr>
<td>CPMAS</td>
<td>Cross polarisation magic angle spinning</td>
</tr>
<tr>
<td>DAPI</td>
<td>4, 6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DBE</td>
<td>Double bond equivalent</td>
</tr>
<tr>
<td>DCAA</td>
<td>Dissolved combined amino acids</td>
</tr>
<tr>
<td>DEGA</td>
<td>Aspartic acid-glutamic acid-glycine-alanine</td>
</tr>
<tr>
<td>DFAA</td>
<td>Dissolved free amino acids</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethyl formamide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>DMSO-$d_6$</td>
<td>Dimethyl sulfoxide-$d_6$</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DOC</td>
<td>Dissolved organic carbon</td>
</tr>
<tr>
<td>DOM</td>
<td>Dissolved organic matter</td>
</tr>
<tr>
<td>DON</td>
<td>Dissolved organic nitrogen</td>
</tr>
<tr>
<td>DN</td>
<td>Dissolved nitrogen</td>
</tr>
<tr>
<td>EA</td>
<td>Elemental analysis</td>
</tr>
<tr>
<td>ER</td>
<td>Elizabeth River, VA</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionisation</td>
</tr>
<tr>
<td>ESI-FT-ICR-MS</td>
<td>Electrospray ionisation coupled to Fourier transform ion cyclotron resonance mass spectrometry</td>
</tr>
<tr>
<td>ETFE</td>
<td>Ethylene tetrafluoroethylene</td>
</tr>
<tr>
<td>FA</td>
<td>Phenylalanine-alanine</td>
</tr>
<tr>
<td>FEP</td>
<td>Fluorinated ethylene propylene</td>
</tr>
<tr>
<td>FID</td>
<td>Free induction decay</td>
</tr>
<tr>
<td>FMOC</td>
<td>Fluorenylmethyloxycarbonyl chloride</td>
</tr>
<tr>
<td>FT-ICR-MS</td>
<td>Fourier transform ion cyclotron resonance mass spectrometry</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Gas chromatography mass spectrometry</td>
</tr>
<tr>
<td>HBTU</td>
<td>2-(1H-benzotriazol-1-yl)-1, 1, 3, 3-tetra-methyluronium hexafluorophosphate</td>
</tr>
<tr>
<td>HMBC</td>
<td>Heteronuclear multiple bond correlation</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>IC</td>
<td>Inorganic carbon</td>
</tr>
<tr>
<td>IUPAC</td>
<td>International Union of Pure and Applied Chemistry</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LW</td>
<td>Lake water</td>
</tr>
<tr>
<td>MD</td>
<td>Mississippi Delta</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MLB</td>
<td>Mangrove Lake, Bermuda</td>
</tr>
<tr>
<td>MLF</td>
<td>Mud Lake, Florida</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass/charge</td>
</tr>
<tr>
<td>N</td>
<td>Nitrogen</td>
</tr>
<tr>
<td>Nap</td>
<td>1, 4-naphthoquinone</td>
</tr>
<tr>
<td>Nap-A</td>
<td>1, 4-naphthoquinon-2-yl-alanine</td>
</tr>
<tr>
<td>Nap-AFRV</td>
<td>1, 4-naphthoquinon-2-yl-alanine-phenylalanine-arginine-valine</td>
</tr>
<tr>
<td>Nap-AREG</td>
<td>1, 4-naphthoquinon-2-yl-alanine-arginine-glutamic acid-glycine</td>
</tr>
<tr>
<td>Nap-AVFA</td>
<td>1, 4-naphthoquinon-2-yl-alanine-valine-phenylalanine-alanine</td>
</tr>
<tr>
<td>NCOM</td>
<td>Nitrogen-containing organic matter</td>
</tr>
<tr>
<td>NHMFL</td>
<td>National High Magnetic Field Laboratory</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>O</td>
<td>Oxygen</td>
</tr>
<tr>
<td>ODU</td>
<td>Old Dominion University</td>
</tr>
<tr>
<td>OM</td>
<td>Organic matter</td>
</tr>
<tr>
<td>P</td>
<td>Phosphorus</td>
</tr>
<tr>
<td>PAH</td>
<td>Polycyclic aromatic hydrocarbon</td>
</tr>
<tr>
<td>PDA</td>
<td>Photodiode array</td>
</tr>
<tr>
<td>PON</td>
<td>Particulate organic nitrogen</td>
</tr>
<tr>
<td>PTFE</td>
<td>Poly(tetrafluoroethane)</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene Fluoride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PW</td>
<td>Pore water</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RuBisCO</td>
<td>Ribulose-1, 5-phosphate carboxylase/oxygenase</td>
</tr>
<tr>
<td>S</td>
<td>Sulfur</td>
</tr>
<tr>
<td>SWGA</td>
<td>Serine-tryptophan-glycine-alanine</td>
</tr>
<tr>
<td>SOM</td>
<td>Sediment(ary) organic matter</td>
</tr>
<tr>
<td>SON</td>
<td>Sediment(ary) organic nitrogen</td>
</tr>
<tr>
<td>SPE</td>
<td>Solid phase extraction</td>
</tr>
<tr>
<td>SPPS</td>
<td>Solid phase peptide synthesis</td>
</tr>
<tr>
<td>S/N</td>
<td>Signal/noise</td>
</tr>
<tr>
<td>TC</td>
<td>Total carbon</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TOC</td>
<td>Total organic carbon</td>
</tr>
<tr>
<td>TN</td>
<td>Total nitrogen</td>
</tr>
<tr>
<td>VFA</td>
<td>Valine-phenylalanine-alanine</td>
</tr>
<tr>
<td>VPL</td>
<td>Valine-proline-leucine</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>xiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>xv</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 WHY STUDY N ORGANIC MATTER?</td>
<td>1</td>
</tr>
<tr>
<td>1.2 DISSERTATION OBJECTIVES</td>
<td>5</td>
</tr>
<tr>
<td>1.3 BACKGROUND INFORMATION RELEVANT TO THE FOCUS OF THIS DISSERTATION</td>
<td>7</td>
</tr>
<tr>
<td>2. A NEW APPROACH FOR MOLECULAR-LEVEL CHARACTERISATION OF N-CONTAINING COMPOUNDS IN SEDIMENTS BY FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY: EXTRACTION OPTIMISATION</td>
<td>21</td>
</tr>
<tr>
<td>2.1 INTRODUCTION</td>
<td>21</td>
</tr>
<tr>
<td>2.2 MATERIALS AND METHODS</td>
<td>24</td>
</tr>
<tr>
<td>2.3 RESULTS AND DISCUSSION</td>
<td>28</td>
</tr>
<tr>
<td>2.4 CONCLUSIONS</td>
<td>53</td>
</tr>
<tr>
<td>3. A METHOD TO INVESTIGATE THE N-CONTAINING MOLECULES IN SEDIMENTARY ORGANIC MATTER AS DETERMINED BY FOURIER TRANSFORM-ION CYCLOTRON RESONANCE-MASS SPECTROMETRY</td>
<td>55</td>
</tr>
<tr>
<td>3.1 INTRODUCTION</td>
<td>55</td>
</tr>
<tr>
<td>3.2 MATERIALS AND METHODS</td>
<td>58</td>
</tr>
<tr>
<td>3.3 RESULTS AND DISCUSSION</td>
<td>59</td>
</tr>
<tr>
<td>3.4 CONCLUSIONS</td>
<td>76</td>
</tr>
<tr>
<td>4. ALKYL AMIDES IN TWO ORGANIC-RICH ANOXIC SEDIMENTS: A POSSIBLE NEW ABIOTIC ROUTE FOR N SEQUESTRATION</td>
<td>78</td>
</tr>
<tr>
<td>4.1 INTRODUCTION</td>
<td>78</td>
</tr>
<tr>
<td>4.2 MATERIALS AND METHODS</td>
<td>81</td>
</tr>
<tr>
<td>4.3 RESULTS AND DISCUSSION</td>
<td>87</td>
</tr>
<tr>
<td>4.4 CONCLUSIONS</td>
<td>109</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>26</td>
</tr>
<tr>
<td>1.</td>
<td>26</td>
</tr>
<tr>
<td>2.</td>
<td>30</td>
</tr>
<tr>
<td>3.</td>
<td>36</td>
</tr>
<tr>
<td>4.</td>
<td>63</td>
</tr>
<tr>
<td>5.</td>
<td>86</td>
</tr>
<tr>
<td>6.</td>
<td>102</td>
</tr>
<tr>
<td>7.</td>
<td>129</td>
</tr>
<tr>
<td>8.</td>
<td>154</td>
</tr>
<tr>
<td>9.</td>
<td>160</td>
</tr>
<tr>
<td>10.</td>
<td>186</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure | Page
---|---
1. The N cycle in an aqueous environment showing the importance of OM throughout the cycle. | 3
2. 1D $^1$H liquid state NMR of intensity scaled solvent extracted MLB sediment from 2.8 – 2.85 m | 33
3. Negative ion analysed FT-ICR-MS spectrum of MLF 0 m sediment solvent extracts | 38
4. Negative ion analysed FT-ICR-MS expanded spectra (352 – 353 m/z) for MLF 0 m sediment | 41
5. The total number of formulae assigned in three sediments from MLF and MLB extracted in five different solvents and analysed using FT-ICR-MS negative ionisation mode | 44
6. The total number of formulae types assigned in three sediments from MLF and MLB in methanol and pyridine analysed using negative ion FT-ICR-MS | 46
7. FT-ICR-MS spectra of BG (0.21 m) ((a), expanded in (b)) and MD (0 m) ((c), expanded in (d)) pyridine extracted sediments analysed by negative ionisation mode | 49
8. Four part Venn diagram comparing assigned formulae for BG, MLB, MLF and MD pyridine extracted FT-ICR-MS analysed sediments | 51
9. Charts displaying the abundances for the assigned formulae containing N in FT-ICR-MS negatively analysed pyridine extracts of 4 sediments | 61
10. van Krevelen diagram of select CHNO-containing formulae (as listed in the key) of the pyridine extract of MLB sediment analysed by FT-ICR-MS in negative ion mode | 65
11. Plot of Kendrick mass defect vs. nominal Kendrick mass for MLB pyridine extracted sediment analysed by FT-ICR-MS negative ion mode | 67
12. Diagram plotting the ratio of DBE/C against O/C of FT-ICR-MS negative ion mode analysed pyridine extract of MLB sediment | 69
13. O/C vs. N/H diagrams of negatively analysed MS data of MLB extracted sediment | 72
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.</td>
<td>O/C vs. N/H diagrams of 4 sediment extracts analysed by negative ion mode FT-ICR-MS</td>
</tr>
<tr>
<td>15.</td>
<td>NMR $^{13}$C CPMAS solid-state spectra of whole sediment samples from (a) MLB of depth 0.52 m, (b) MLB of depth 2.8 m, (c) MLF of 0 m depth, and (d) MLF of depth 3.0 m</td>
</tr>
<tr>
<td>16.</td>
<td>NMR $^{15}$N CPMAS solid-state spectra of whole sediment samples from (a) MLB of depth 0.52 m, (b) MLB of depth 2.8 m, (c) MLF of 0 m</td>
</tr>
<tr>
<td>17.</td>
<td>HMBC NMR spectrum of liquid-state analysis of (a) MLB sediment pyridine extract of 2.8 m and (b) MLF sediment pyridine extract of 3.0 m</td>
</tr>
<tr>
<td>18.</td>
<td>FT-ICR-MS spectra of MLB sediment from 2.8 m depth</td>
</tr>
<tr>
<td>19.</td>
<td>FT-ICR-MS of the pyridine extract from MLF sediment, depth 3 m</td>
</tr>
<tr>
<td>20.</td>
<td>FT-ICR-MS analysis of Behenamide dissolved in pyridine (a) in negative ionisation mode, (b) in positive ionisation mode</td>
</tr>
<tr>
<td>21.</td>
<td>GC-MS selected ion chromatogram (m/z 59) of MLB extract from sediment of depth 2.8 m</td>
</tr>
<tr>
<td>22.</td>
<td>The percentage of formula types from MLB at (a) 0.52 m and (b) 2.8 m and MLF (c) 0 cm and (d) 3.0 m analysed using negative ionisation mode FT-ICR-MS</td>
</tr>
<tr>
<td>23.</td>
<td>Magnitude weighted percent of alkyl amides in FT-ICR-MS negative ionisation data of depths 0.22 m to 3.0 m depth of pyridine sediment extracts from MLB</td>
</tr>
<tr>
<td>24.</td>
<td>FT-ICR-MS spectrum of fatty acid and alkyl amide mixed standard analysed using negative ionisation mode</td>
</tr>
<tr>
<td>25.</td>
<td>FT-ICR-MS spectra analysed using negative ionisation mode of sunflower oil (a) is before reaction, and (b) after reaction with ammonium carbonate</td>
</tr>
<tr>
<td>26.</td>
<td>A map showing the location where samples were collected (shown with an X) in MLB in November 2009</td>
</tr>
<tr>
<td>27.</td>
<td>Photos from left clockwise, full core barrel on the raft, raft floating on MLB, full core barrel showing handle, stainless steel coupling and stopper, full core barrel while still in water with piston inside, amber vials used for sample storage and Glove Bag containing 4 sample bags ready for PW sampling</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>28.</td>
<td>Depth profiles of TN concentration of the PW (mM), N % of dry sediment weight and NH$_4^+$ concentration in the PW (mM) in the samples collected from MLB in November 2009. 127</td>
</tr>
<tr>
<td>29.</td>
<td>Pyridine extract mass spectra of 4 sediment depths collected from MLB in 2009, analysed by FT-ICR-MS negative ion mode. 130</td>
</tr>
<tr>
<td>30.</td>
<td>PW DOM extract mass spectra (336 – 337 Da) of a depth profile from MLB sediment collected in November 2009 and analysed by FT-ICR-MS negative ion mode. 133</td>
</tr>
<tr>
<td>31.</td>
<td>LW DOM PPL extracted mass spectra from MLB analysed by negative ion FT-ICR-MS. 135</td>
</tr>
<tr>
<td>32.</td>
<td>Formula type classifications as determined by N, O, S and P contents, calculated as a percent of total number of formulae assigned for PW and LW DOM and sediment extracts analysed by FT-ICR-MS negative ion mode. 137</td>
</tr>
<tr>
<td>33.</td>
<td>Plots of N/H versus O/C for two sediment samples, two PW samples and one LW sample from MLB. 140</td>
</tr>
<tr>
<td>34.</td>
<td>Three plots displaying characteristics of PW, LW and sediment collected from MLB. 143</td>
</tr>
<tr>
<td>35.</td>
<td>Maps showing the locations of water collected in MLB and ER and used for peptide incubations. 152</td>
</tr>
<tr>
<td>36.</td>
<td>Plot of concentration against time for the incubation of AFRV and AAFRV in ER water. 162</td>
</tr>
<tr>
<td>37.</td>
<td>Concentration plots of peptide removal for single and quad incubations in water collected from ER and single vs dual incubations in water from MLB. 165</td>
</tr>
<tr>
<td>38.</td>
<td>Mass spectra of DOM from ER and MLB water PPL extracted and analysed by negative ionisation mode in FT-ICR-MS. 170</td>
</tr>
<tr>
<td>39.</td>
<td>van Krevelen diagrams plotting H/C against O/C of DOM from ER and MLB waters analysed by negative ionisation FT-ICR-MS. 172</td>
</tr>
<tr>
<td>40.</td>
<td>HPLC chromatograms of AFRV and Nap-AFRV incubations in collected ER water. 187</td>
</tr>
<tr>
<td>41.</td>
<td>Concentration profiles of (a) AFRV + Nap-AFRV, (b) AVFA + Nap-AVFA and (c) AREG + Nap-AREG incubations in ER water measured by HPLC. 190</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>42.</td>
<td>Tandem MS of peak 3 during the ER water incubation of Nap-AVFA</td>
</tr>
<tr>
<td>43.</td>
<td>Each of the suggested fragments displayed in Fig. 42 is indicated on the peak 3 structures</td>
</tr>
<tr>
<td>44.</td>
<td>Structure of Nap-AVFA (top) and the proposed structure of peak 3 (bottom), formed during the degradation of Nap-AVFA in ER water collected by HPLC and analysed by positive ion mode FT-ICR-MS</td>
</tr>
<tr>
<td>45.</td>
<td>Structure of Nap-AFRV and proposed structure of peak 1, formed during the incubation of Nap-AFRV in collected ER water</td>
</tr>
<tr>
<td>46.</td>
<td>Structure of Nap-AREG and proposed structure of peak 5 formed during the incubation of Nap-AREG in ER water</td>
</tr>
<tr>
<td>47.</td>
<td>The proposed structure of peak 2, 4 and 6 formed during the incubation of Nap-AFRV, Nap-AVFA and Nap-AREG respectively in ER water</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1. Why study N organic matter?

The importance of non-living organic matter in global nutrient cycling is immense due to its extremely high amount compared to living organic matter (~ $3300 \times 10^{15}$ g $C_{\text{non-living}}$ versus ~ $600 \times 10^{15}$ g $C_{\text{living}}$) (Hedges et al., 2000; Hedges and Keil, 1995). Non-living organic matter (herein referred to as organic matter (OM)), is key for balancing nutrient cycles (e.g., C, N, and O) globally, since they rely heavily on the remineralisation of OM to enable survival of all organisms on the planet. The highly efficient remineralisation process ensures that only a small amount (~ 0.5 % of OM) escapes to become preserved in dissolved organic matter (DOM) of marine and fresh water systems, as well as soils and sediments (Hedges et al., 2000). The reasons why these particular compounds escape remineralisation and are preserved for time periods ranging from a few days to thousands of years are still largely unknown despite decades of research (Hartnett et al., 1998; Hedges et al., 2000; Hedges and Keil, 1995; Mopper et al., 2007). One of the main difficulties in uncovering the answer is that OM matrices are very complex and vary substantially according to environmental conditions (Baldock et al., 2004).

The study of nitrogen-containing organic matter (NCOM) is one area of research

This dissertation has been formatted according to the journal Geochimica et Cosmochimica Acta.
that has received considerable attention for decades (Aluwihare et al., 2005; Berman and Bronk, 2003; Herbert, 1999; McCarthy et al., 1997; Schulten and Schnitzer, 1998; Worsfold et al., 2008; Zehr and Ward, 2002). N exists naturally in many different forms, both inorganic and organic, including nitrate, nitrite, ammonia, amino acids, DNA and urea (Fig. 1). All of these forms are essential for a balanced N cycle that in turn affects the cycles of other key nutrients, e.g., C and O (Arrigo, 2005; Hulth et al., 2005). My research focuses on some key aspects of the N cycle that impact the manner in which peptides (both biologically-produced and abiotically-produced) are preserved or degraded in natural waters and sediments.

In recent years, many additional anthropogenically produced N compounds have been released into the environment that threatens to destroy the carefully regulated natural N cycle (Galloway et al., 2008; Petit et al., 1999; Vitousek et al., 1997). Polycyclic aromatic hydrocarbons (PAHs), more specifically their oxidised forms (quinones), are one of the anthropogenically produced classes of compounds that can have dramatic effects on the N cycle through covalent bonding to existing NCOM including peptides (Hsu and Hatcher, 2005; Means et al., 1982; Parris, 1980). In Chapter 6 as one focus of my dissertation, I investigate the effect of covalent coupling of peptides to a quinone giving new insight into how these compounds can impact the N cycle.
Essential to understanding the organic N contribution to the N cycle in aqueous systems, another focus of this dissertation, is a clear knowledge of its structural composition in DOM and sediments. Many studies have sought to characterise the organic N in DOM and sediments worldwide (e.g., Aluwihare et al., 2005; Burdige and Zheng, 1998; McCarthy et al., 1997; Worsfold et al., 2008). However, despite many advances, there still remain many questions regarding the molecular identity of NCOM
and how this affects its preservation in DOM and sediment. Past molecular studies of immature sediments have suggested that the entire N in these systems is composed of amides, assumed to be peptides/proteins (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; McCarthy et al., 1997). This is due, in large part, to the fact that measurements of peptide and amino acid concentrations in aqueous environments reveal that they can often be a major component of NCOM (Nguyen and Harvey, 1998; Nguyen and Harvey, 2001). The conundrum is, however, that peptides/proteins are considered to be very labile in natural systems, and consequently, they should not be well preserved in sediment organic matter (SOM) or DOM. The unexpected preservation of amide N requires explanation and many mechanisms, both chemical and physical, have been suggested as to how amide N is preserved including: selective preservation of recalcitrant material (Hatcher et al., 1983), encapsulation into more recalcitrant OM complexes (Knicker and Hatcher, 1997; Zang et al., 2000), sorption to mineral matter (Kirchman et al., 1989; Kleber et al., 2007), inherent chemical recalcitrance (Nguyen and Harvey, 2001), oxygen exposure (Gray et al., 2002; Hartnett et al., 1998), and chemical reactions with existing OM (Burdige, 2007; Hsu and Hatcher, 2005; Knicker, 2004; Maillard, 1916; Rillig et al., 2007; Roth and Harvey, 2006). However, much of this previous work relies on a series of assumptions regarding the exact molecular character of the amide N and how it interacts with existing matter already present in the system. In this dissertation I seek to re-evaluate some of these assumptions with molecular characterisation of the structural composition of the amide N and testing of some of the above proposed hypotheses in natural systems. To accomplish this molecular characterisation, I employ some new approaches involving the use of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) and advanced nuclear magnetic
resonance (NMR) spectroscopy (Chapters 1 and 2).

1.2 Dissertation objectives

The overarching hypothesis for my dissertation is that environmental recalcitrance of amide organic N is determined by its structure, its production and its reactivity upon release. To study this broadly encompassing hypothesis, I have sought to address it, in part, through several sub-hypotheses that I will answer in this dissertation:

1. Amide N is not comprised solely of peptides/proteins in immature sediments, but may be composed of additional compounds not previously identified, and that these compounds are a major fraction of recalcitrant amide N in sediments while the more microbially accessible peptides are degraded. The use of advanced MS and NMR methods gives me the unique opportunity to examine this hypothesis in a manner that has never been accomplished previously (Chapter 4). I use a 2D NMR technique, heteronuclear multiple bond correlation (HMBC) to specifically examine the amide N present in the sample. Using the extraction protocol developed in Chapter 1, I analyse two organic-rich sediments, Mangrove Lake, Bermuda (MLB) and Mud Lake, Florida (MLF), on the molecular level by employing FT-ICR-MS and gas chromatography mass spectrometry (GC-MS) analysis, specifically identifying those formulae that likely contain amide functionality. Organic-rich sediments such as those from MLF and MLB are useful study locations to learn about processes involving sedimentary organic N (SON) that are occurring in organic-poor sites, due to the enhanced ability to detect more subtle changes in the OM composition than in sediments having lower amounts of OM.
2. The conclusions reached in addressing the above hypothesis (Chapter 4) have uncovered the existence of abiotically-produced alkyl amides. *I hypothesise that the abiotic incorporation of inorganic N is through an amidation reaction and is a key pathway for production and recalcitrance of amide N compounds in lacustrine environments.* In Chapter 4 I demonstrate that amidation occurs with model compounds (triglycerides from vegetable oils). In Chapter 5 I examine the evolution of alkyl amides in a natural lacustrine system (MLB) as evidence that they are produced in the environment by reaction with naturally produced ammonia by molecularly characterising the sediment, pore water and lake water NCOM and comparing their amide N to the production of inorganic N in the sediment system.

3. Addressing another portion of the N cycle, that deals with biodegradation of peptides and factors that control this biodegradation, *I examine the hypothesis that the residue length and natural environment of peptide release are the major determinants for peptide preservation in aqueous environments* (Chapter 6).

The inherent chemical diversity of properties of peptides as determined by their amino acid composition seems a likely reason why some peptides are preserved and others are remineralised rapidly. Chemical structure is therefore, investigated in Chapter 6 by comparing rates of removal of differently sequenced peptides in natural waters. Residue number also seems a likely reason why some peptides are preserved, since larger peptides should be less accessible for microbial attack than those that are smaller due to stronger interactions. Consequently, peptides of different length are studied and their removal rates compared in Chapter 6. The environments into which peptides are released are each different in terms of their OM composition, bulk properties and microbial population and therefore, it seems probable that peptide preservation is controlled by the
inherent conditions of the water environment. Therefore, two different water samples are used in Chapter 6 so that a comparison is possible between a lacustrine environment using MLB lake water and a riverine environment using water from the Elizabeth River, VA. The results of Chapter 6, therefore, provide new details regarding some of the most important factors governing the preservation of peptides in natural waters.

4. Previous studies have hypothesised that the covalent binding of peptides to PAH type structures through Michael reaction chemistry can dramatically influence their degradation in natural systems and is likely a major route of PAH bioremediation (Hsu and Hatcher, 2005). In Chapter 7, I test this concept by synthesising naphthoquinone adducts to peptides and subjecting these to biodegradation in natural waters. Due to the increase in anthropogenically produced compounds, e.g., PAHs, these structures could interact with peptides to enhance their stability in natural systems and aid in their preservation (Hsu and Hatcher, 2005; Hsu and Hatcher, 2006; Means et al., 1982; Novoszad et al., 2005; Parris, 1980). I examine this possibility by covalently binding 1, 4-naphthoquinone (Nap), an example oxidised PAH, to a series of synthesised peptides and monitor their rate of removal in natural water compared to the un-adducted peptides. In doing so, I can determine whether Michael adduction is a preservation route for peptides in natural waters.

1.3 Background information relevant to the focus of this dissertation

1.3.1 Molecular studies of amide N in OM

Many studies have demonstrated that immature sediments are primarily composed of amide N (Hatcher et al., 1983; Knicker and Hatcher, 1997; Nguyen and Harvey, 1998;
Vairavamurthy and Wang, 2002) with contributions from aminosugars in some locations (Benner and Kaiser, 2003; Cowie and Hedges, 1994; Dauwe and Middelburg, 1998; Niggemann and Schubert, 2006). As amide N is degraded with time, both biologically and thermally mediated, it is chemically transformed, and consequently, several studies have found that more mature sediments are composed primarily of pyrolic and, pyridinic type structures (Boudou et al., 2008; Kelemen et al., 2006; Knicker et al., 2002; Knicker et al., 1996; Schimmelmann and Lis, 2010). Interestingly, pyrolic and pyridinic structures have also been identified in a surface upwelling region, probably because the thermally mature OM has been displaced upwards (Patience et al., 1992). These N-containing aromatic structures are likely precursors for formation of recalcitrant structures that constitute a major component of kerogen, coals, petroleum and lignites (Kelemen et al., 2002; Kelemen et al., 1999; Li et al., 1997; Mitra-Kirtley et al., 1993a; Mitra-Kirtley et al., 1993b).

However, despite the high lability of many amide structures, they appear to be preserved for long periods of time unexpectedly (Burdige, 2007; del Rio et al., 2004; Derenne et al., 1993; Hedges and Keil, 1995). Some of the hypotheses that have been suggested to facilitate the preservation of amides in sediments were briefly discussed above; I will discuss them in more detail here. Physical protection of amide N in sediments is one of the methods that have been proposed in past studies, this includes protection by sorption to mineral matter (Arnarson and Keil, 2001; Gordon and Millero, 1985; Kirchman et al., 1989; Kleber et al., 2007; Lambert, 2008; Mayer, 1994a). This is a probable pathway since inorganic matter is often the major component in sediment systems. Encapsulation of amide N in more recalcitrant OM such as algaenan, by sandwiching between cell walls, has also been proposed as a physical protection pathway
for amide preservation (Knicker and Hatcher, 1997; Zang et al., 2000); however, to date this hypothesis has not been proven in natural systems. It is possible that some amide containing compounds in sediments possess an inherent recalcitrance to remineralisation through their chemical and physical properties; thereby they are selectively preserved in the sediment (Hatcher et al., 1983; Nguyen and Harvey, 2001). Chemical interactions with existing OM, by formation of adducts such as Michael adducts, or through reactions with more recalcitrant material such as lignin or tannin could also serve to preserve amide N (Hsu and Hatcher, 2005; Hsu and Hatcher, 2006; Rillig et al., 2007). The earliest chemical reaction suggested to protect amide N, is through a Maillard reaction, during which carbohydrates and proteins react together, forming the more stable product of melanoidins (Hedges and Keil, 1995; Maillard, 1916). An additional method that would be observed as preservation of amides in sediment even though new material is being produced is through incorporation of inorganic N into existing OM and has been observed to occur in a few studies (Clinton et al., 1995; Dong and Yuan, 2009; Knicker et al., 1997; Thorn and Mikita, 1992). All or some of these proposed methods could be occurring in immature sediments serving to preserve the abundant amide N for many years, making it an important component of the global N cycle.

It is commonly thought that the composition of immature sediment amide is primarily due to preservation of proteins and peptides (Knicker and Hatcher, 2001; Nguyen and Harvey, 1998; Rillig et al., 2007; Tappin et al., 2007; Vairavamurthy and Wang, 2002). This conclusion is based upon the fact that peptides and proteins are often a major component of living organisms and upon their death, peptides and proteins are released and distributed throughout aquatic systems. Consequently, past molecular studies of dissolved organic nitrogen (DON) have identified proteins, and peptides in a
variety of different locations, including a freshwater ecosystem (Chen et al., 2010; Maie et al., 2006; Maie et al., 2007), lacustrine environments (Hatcher, 1978; Rosenstock and Simon, 2001; Simon, 1998), in ocean and seawater (Aluwihare et al., 2005; McCarthy et al., 1996; McCarthy et al., 1997; Tanoue, 1995; Tanoue et al., 1996), and in coastal areas (Harvey and Mannino, 2001; Herbert, 1999; Kuznetsova and Lee, 2001; Kuznetsova et al., 2004; Sommerville and Preston, 2001): The global contribution of peptides and proteins to all DON is estimated to be about 5 – 20 % (Sharp, 1983), although it is often difficult to quantify them probably because of the methods that are preserving them (e.g., encapsulation, absorption). The amount of peptide/protein in certain systems can be much higher than this, especially under anoxic conditions (Hartnett et al., 1998; Nguyen and Harvey, 1997; Orem et al., 1986). Therefore, the role peptides/proteins play in the N cycle in terms of their preservation and degradation is a key component that requires extensive study.

1.3.2 Peptide removal from natural water

Peptides and proteins are one of the compound classes that constitute dissolved combined amino acids (DCAA), since they are composed of covalently bound amino acids. In order to make the DCAA pool more bioavailable, microorganisms in aqueous systems degrade these complexes in a variety of ways, primarily through hydrolysis, so that they can be utilised in their systems. These methods include extracellular hydrolysis (Arnosti, 2011; Kuznetsova and Lee, 2001; Obayashi and Suzuki, 2005; Pantoja and Lee, 1999) which is required to decompose compounds that are larger than ~ 600 Da to pass through the cell membrane (Alexander, 1973; Nikaido and Vaara, 1985; Weiss et al., 1991) by releasing enzymes into the aquatic environment. Upon extracellular hydrolysis,
the smaller peptide fragments are then taken into the cell of the micro-organism (Mulholland and Lee, 2009; Pantoja et al., 1997). If the peptide is small enough it can be directly taken into the cell without prior extracellular hydrolysis. The rate at which uptake and consequently, loss of measurable peptide in the water occurs has been measured in several studies, with values ranging from a few hours to several days (Liu et al., 2010; Nunn et al., 2003; Pantoja and Lee, 1999; Pantoja et al., 1997; Pantoja et al., 2009; Roth and Harvey, 2006). The hydrolysis of larger proteins and peptides has been demonstrated to occur in a stepwise manner, and specific (in terms of cleavage point) pathway indicating that there are some fundamental rules regarding peptide hydrolysis (Pantoja et al., 1997). Despite this highly efficient process of peptide removal from natural waters, amino acids can be detected as dissolved free amino acids (DFAA) indicating that some escape and can be preserved (Dauwe and Middelburg, 1998; Kuznetsova et al., 2004; Pantoja et al., 2009; Rosenstock and Simon, 2001) in addition, to some intact peptides and proteins. The processes that govern peptide/protein and amino acid degradation are therefore, very complicated and although many advances have been made in this field more extensive study is required before all factors controlling hydrolysis are understood.

1.3.3 Study of lacustrine OM

Similar to other environments, lacustrine OM sources can range from terrestrial to algal derived, depending on the conditions surrounding the lake of interest (Meyers and Ishiwatari, 1993; Meyers and Teranes, 2001; Wetzel, 2001). Consequently, OM in these systems is composed of a variety of different compound classes including carbohydrates, lipids, lignin, tannin, and proteins. N cycling in lakes is of particular interest since the
use of inorganic N in many industrial processes, has significant run off into rivers and streams that often sources lacustrine environments (Foley et al., 2005; Vitousek et al., 1997). As the inorganic N concentrations increase in lakes as a result of the new anthropogenic sources, eutrophication occurs, dramatically altering the N cycle within the lake (Cabana and Rasmussen, 1996; Carpenter et al., 1998; Downing and McCauley, 1992). It is of high importance, therefore, to have an in-depth knowledge of the N cycle within these environments so that one can make meaningful alterations to the N cycle in these systems in order to combat this issue.

The composition of OM within a lacustrine environment plays an important role in regulating the inorganic N concentration, as part of the N cycle (Fig. 1). Many studies have been directed to unravelling the constituents of lacustrine OM (e.g., Emerson, 1976; Meyers and Ishiwatari, 1993; Meyers and Teranes, 2001). In this dissertation, much of my research uses samples from organic-rich lakes (MLF and MLB), which are sensitive indicators of organic N composition and chemistry for locations with significantly lower N contributions. In particular, MLB, has been extensively studied in the past regarding its NCOM composition (Boudreau et al., 1992; Hatcher, 1978; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Knicker et al., 2002; Zang and Hatcher, 2002; Zang et al., 2000). The organic-rich nature of this lake is due mainly to the anaerobic conditions that the OM is subject to upon deposition (Hatcher, 1978). As discussed above, much of the NCOM is composed of amides (Section 1.3.1), which is generally considered to be due to peptides/proteins prevalence (Knicker and Hatcher, 2001). In addition, the algal nature of the OM source in MLB means that aliphatic functionality is common (Hatcher, 1978). The SOM in MLB is dominated at shallow depths by carbohydrates, after which, substantially lower concentrations are observed in sediment from deeper horizons (> 3
m), suggesting their selective diagenesis. Amide N is the only N functionality observed throughout the first ~10 m of sediment, after which aromatic N becomes more important. This trend was discussed above in Section 1.3.1. From these past studies, MLB presents a system that is representative of many other locations in terms of NCOM trends. Therefore, as described in Section 1.2, samples from MLB will be utilised to investigate in more detail the molecular composition and preservation pathways of NCOM.

1.3.4 Advanced analytical techniques for the study of NCOM

It is necessary for my study of NCOM to use advanced analytical techniques that are capable of examining in detail the processes/characterisation under study (Hatcher et al., 2001; Mopper et al., 2007). Many techniques have been applied to assist in characterising NOM, including infrared spectroscopy (Chang and Laird, 2002; McLellan et al., 1991; Niemeyer et al., 1992; Rossel et al., 2006), NMR spectroscopy (Derenne et al., 1993; Knicker, 2000; Kögel-Knabner, 1997; Simpson et al., 2011), GC-MS (Knicker et al., 2001; Schulten and Gleixner, 1999; Schulten et al., 1997; Sorge et al., 1993), HPLC (Cowie and Hedges, 1992; Liu et al., 2010; Poirier et al., 2005), and FT-ICR-MS (Fu et al., 2006; Purcell et al., 2007; Wu et al., 2003). In this dissertation I primarily use two techniques that have been successfully employed to study DON and SON, NMR and FT-ICR-MS. A variety of NMR techniques have been used in a number of studies, and consequently, DON has been suggested to be composed of a number of structures. These include: aminosugars in deep marine water (Aluwihare et al., 2005), proteins/peptides in ultrafiltered riverine water (Kaiser et al., 2003) and in freshwater (Lu et al., 2003), and pyridinic in estuarine and river water (Maie et al., 2006). NMR studies of sediments have revealed that it is
primarily composed of protein/peptide in immature sediment (Burdige, 2007; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Nguyen and Harvey, 1998) but upon maturation, pyridinic N is dominant (Kelemen et al., 2006; Knicker et al., 2002). These past studies demonstrate that NMR is an effective technique for examining N-containing compounds in DOM and SOM from a variety of environments. However, NMR is only capable of examining N molecular character of OM for the entire sample which although it is very useful for assessing the general nature of a sample; it is not capable of identifying individual molecular species which is essential for a complete and detailed examination of the DON and SON.

FT-ICR-MS is a relatively new technique that has successfully examined a variety of different OM sample types. These include: DOM from fresh/marine water (Hertkorn et al., 2006; Kujawinski et al., 2009; Reemtsma et al., 2008; Sleighter and Hatcher, 2008), rainwater DOM (Altieri et al., 2009), pore water DOM (D'Andrilli et al., 2010; Koch et al., 2005; Schmidt et al., 2009; Tremblay et al., 2007), water soluble aerosol OM (Kroll et al., 2011; Mazzoleni et al., 2010; Schmitt-Kopplin et al., 2010b; Wozniak et al., 2008), meteorite OM (Schmitt-Kopplin et al., 2010a) and soil OM (Kujawinski et al., 2002; Ohno et al., 2010). FT-ICR-MS identifies individual molecular species, a character that is immensely useful in making new advances in understanding N chemistry (Fu et al., 2006; Laskin et al., 2009; Purcell et al., 2007; Reemtsma et al., 2008). Together, these two techniques can provide high levels of information regarding NCOM, in particular of lacustrine OM, enabling new insights into the molecular chemistry of this OM.

One of the requirements of FT-ICR-MS is that the sample needs to be in liquid form for injection into the electrospray (ESI) source. This makes the study of SOM more
difficult since a suitable liquid medium needs to be utilised in order to enable its examination by FT-ICR-MS. This has been achieved in past studies of soil samples by extracting humic/fulvic acids and dissolving in basic water prior to analysis (Kujawinski et al., 2002; Ohno et al., 2010). A study by Schmidt-Kopplin et al. (2010) investigated extraction of meteorite OM by a variety of solvents, concluding that polar protic solvents were most suitable. In addition, pyridine was used to extract a wood, kerogen and immature sediment sample with high levels of success (Zhong et al., 2011). Essential to successful extraction of SOM is removal of a representative fraction of the original SOM sample so that any inferences regarding the N character obtained by FT-ICR-MS are meaningful. To this end, this dissertation begins with a thorough examination of a variety of extraction solvents to enable FT-ICR-MS analysis of a representative fraction of SOM (Chapter 2). The extraction efficiency of OM removal is assessed by both bulk and NMR examination effectively quantifying the ability of each solvent tested to remove the most dominant structural entities from the sediment. In order to make this study more relevant to a variety of studies, four different sediments are investigated, including two organic-rich and two organic-poor sediments. The source of the two organic-rich sediments are MLB and MLF, both of which have been extensively examined in past studies (Bates et al., 1995; Boudreau et al., 1992; Bradley and Beard, 1969; Canfield et al., 1998; Filley et al., 2001; Filley et al., 2002; Hatcher, 1978; Hatcher et al., 1982; Knicker and Hatcher, 2001; Orem et al., 1986; Spiker and Hatcher, 1984; Zang and Hatcher, 2002). The two organic-poor sediment samples examined are from Bayou Grande, Pensacola, Florida (BG) and the South Pass of the Mississippi Delta (MD), whose past studies have revealed much about their sediment characteristics (Bennett et al., 1991; Bennett et al., 1977; Bennett and Faris, 1979; Lewis et al., 2001; Simpson et
The results of this Chapter are used for successful examination of the N characteristics of SOM in future Chapters (3 – 5).

One of the advantages and disadvantages of FT-ICR-MS analysis of OM samples is the ability to identify hundreds to thousands of molecular species. This both increases the amount of information obtained but also the complexity of the data analysis. In most past studies utilising FT-ICR-MS analysis of OM samples, graphical methods have been employed to enable easier analysis of the resulting spectra. These methods include the van Krevelen diagram (Kim et al., 2003a), Kendrick mass defect analysis (Hughey et al., 2001), carbon oxidation state (Kroll et al., 2011), double bond equivalent (DBE) versus oxygen (Kujawinski et al., 2009), DBE/C versus C (Hockaday et al., 2006) and N/C versus H/C (Wu et al., 2004). All of these diagrams have been successful for examining various components of the OM samples analysed. In Chapter 3 I investigate the usefulness of some of these above mentioned diagrams and two new plots with an emphasis on examining the N-species in the most time-effective and informative way. The goal of the study therefore, is to determine the most suitable diagram that examines only those formulae that contain N without pre-sorting of data, and also to separate data into likely structural sources. The same four samples that are analysed in Chapter 2 are examined in this chapter namely MLB, MLF, MD and BG sediments demonstrating the applicability of the chosen diagram to a range of sediment sample types.

1.3.5 Development of a synthetic pathway for peptides and adducts production

Paramount for the investigation into peptide preservation in natural water is the synthesis of high purity peptides and Michael adducts. During my PhD I have devoted much time to developing protocols to facilitate this. The primary synthetic route that I
employ is solid phase peptide synthesis (SPPS); that has been used in past studies to reproducibility produce highly pure peptides (Chan and White, 2000; Fields et al., 1991; Merrifield, 1963; Stewart, 1976; Wang, 1973). This synthetic route relies on using a solid resin support to attach amino acids together in a stepwise, automated process. Although simple in principle, the development of SPPS for the research in this dissertation required much trial and error to determine the optimum reaction times, and the most reliable amino acids to use in terms of attachment probability, side reaction minimisation and ease of workup after synthesis. The resin used for SPPS in this dissertation is Wang, which attaches to the C-terminus of the first amino acid in the chain (Wang, 1973). The advantage of the Wang resin is that it is acid labile so that its removal is facilitated through a simple cleavage reaction after addition of all amino acids. The attachment of the first amino acid is more difficult than later amino acids due to the possibility of side reactions; therefore, for all syntheses in this study, the first amino acid was bought pre-attached to the Wang resin. Resins are key for SPPS since they provide a support for the reaction to occur upon and they are able to swell in the presence of certain solvents, allowing easier access to the amino acid involved in the reaction.

The following reactions all occur using the Peptide Synthesiser (PS3, Protein Technologies). The first step in attachment of an amino acid to the preloaded amino acid-resin complex is deprotection of the Fluorenylmethyloxycarbonyl chloride (FMOC) group on the amino acid amine group. This involves the addition of piperidine (C₅H₁₁N) in a solution of dimethyl formamide (DMF, Sigma Aldrich) that selectively cleaves the bond between the amine of the amino acid and the FMOC group. Having prepared the resin-amino acid complex for reaction the new amino acid needs to be prepared also. This is facilitated by activating the carboxyl group of the amino acid with an activating
agent, 2-(1H-benzotriazol-1-yl)-1, 3, 3-tetra-methyluronium hexafluorophosphate (HBTU), in the presence of DMF (Fields et al., 1991). The activated amino acid is then added to the resin-amino acid complex, in the presence of DMF and allowed to react for 10 minutes to an hour depending on the amino acids. This process is repeated until all of the amino acids have been added to the peptide sequentially. During this entire process side chains of the amino acids are protected with O-\text{t}Bu, so that side reactions are minimised. The resin attached peptide is then removed from the Peptide Synthesiser for further workup.

The workup involves the removal of the side chain protecting groups and the resin, so that only the peptide remains in solution, following the general guidelines provided by the Novabiochem online protocol (www.novabiochem.com). Before this step, the resin and peptide are washed thoroughly with DMF, ethanol (Fisher) and methylene chloride (Fisher) under vacuum and placed in a desiccator for one hour under vacuum. The reaction, known as the “cleavage reaction” involves adding the solid peptide to the cleavage cocktail of 9 mL trifluoroacetic acid (TFA, Fisher), 0.5 mL anisole (Sigma Aldrich) (or 0.5 mL thioanisole (Sigma Aldrich) and 0.3 mL 1, 2-ethanediol (Fisher), depending on the sequence), and 0.5 mL MilliQ water. The cleavage reaction is allowed to react for 30 minutes to 5 hours with stirring. In the cleavage reaction the TFA cleaves the bond between the resin and the peptide and the anisole/thioanisole act as scavengers that immediately react with the cleaved products to prevent unwanted side reactions (King et al., 1990). The slurry is then filtered under vacuum, and the TFA is removed completely by rototary evaporation. The residue is then dissolved in 5 mL 10 % acetic acid (Fluka), and extracted into the aqueous phase while leaving the protecting groups in the 10 mL chloroform (Fisher) organic phase.
In order to purify each synthesised peptide, I used HPLC to systematically collect the desired peptide in high concentration. This was possible after determining the optimum conditions for each peptide, in terms of flow rate, mobile phase and gradients. The retention time for each peptide was confirmed by collecting all relevant peaks and determining their mass and structure using FT-ICR-MS and NMR analysis, thereby confirming that the correct peptide had been produced. The HPLC uses a preparative C₁₈ column (Grace, Apollo 5 µm, 150 mm x 10 mm), with detection using a photodiode array detector (210 nm for all peptides but alanine-arginine-glutamic acid-glycine (AREG) at 200 nm), and collection of peptides using the attached fraction collector (Shimadzu). The mobile phases were generally 0.05 M Na₂HPO₄ and methanol. The peptides chosen for this study are all components of the sequence of Ribulose 1, 5-bisphosphate carboxylase oxygenase, a ubiquitous protein in aqueous systems. The following peptides were synthesised as part of this study, although not all were used in this dissertation as indicated due to synthesis difficulties and prior use in another study by Liu et al. (2010): aspartic acid-glutamic acid-glycine- alanine (DEGA, not used here), valine-proline-leucine (VPL, not used here), serine-tryptophan-glycine-alanine (SWGA, used in previous study), alanine-valine-phenylalanine-alanine (AVFA), alanine-alanine-valine-phenylalanine (AAVF, used in previous study), alanine-phenylalanine-valine-alanine (AFVA, used in previous study) phenylalanine-alanine (FA, used in previous study), valine-phenylalanine-alanine (VFA, used in previous study), alanine-phenylalanine-arginine-valine (AFRV), alanine-alanine-phenylalanine-arginine-valine (AAFRV), and AREG. The synthesised peptides used in this study were then stored at -14 °C until use in the experiments described in Chapters 6 and 7.

In Chapter 7, I investigate the effect of Michael adduction to peptide removal
rates in natural waters. Therefore, I needed to synthesise highly pure Michael adducts of the same peptides used in Chapter 6 in enough quantity (~ 200 – 500 mg) for these experiments. Synthesis of quinone adducted peptides to produce 1, 4-naphthoquinon-2-alanine-yl (Nap-A), involved adapting a method that was previously used by Bittner et al. (2002) to covalently bind glycine to the quinone. Unfortunately, due to the larger size of alanine in this study, this route had very low yield and therefore, approximately 5 – 10 batches of the reaction was required for each peptide adduct to synthesise enough adduct to attach to the rest of the peptide via SPPS. The method employed to synthesise the Nap-A is as follows: approximately 20 mmol of unprotected A (Acros) was dissolved in water (MilliQ) and added to a hot 40 mmol ethanol solution of Nap (Acros). This solution was stirred at room temperature for 24 – 72 hours, a time that was extended from that of Bittner et al. (2002) to increase yield, after which the solution was removed by filtration, and the solid brought to dryness by rotoary evaporation. Purification of Nap-A used a silica gel column (grade 62, 60 – 200 mesh, 150 Å, Sigma-Aldrich) eluting the product with a 9:1 v/v mixture of methylene chloride/methanol (Fisher) with average yields of 4 %. SPPS for the adducted peptides was adapted following the optimisation of the technique by not using a deprotection step in the addition of Nap-A to the rest of the peptide chain. Much of the workup was similar to the peptides, except the low solubility of the adducts presented difficulty in extraction into an aqueous phase. This was bypassed by conducting the purification steps by HPLC using the organic phase. The production of the peptides and adducts in this study was therefore, developed and optimised specifically for my dissertation, enabling the successful monitoring of their concentration in natural waters in Chapters 6 and 7.
CHAPTER 2

A NEW APPROACH FOR MOLECULAR-LEVEL CHARACTERISATION OF N-CONTAINING COMPOUNDS IN SEDIMENTS BY FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY: EXTRACTION OPTIMISATION

2.1 Introduction

Despite the fact that only ~ 1% of organic matter present in the oceanic water column is preserved in oceanic sediments, this sedimentary pool of organic matter is almost equal to that of living organic matter today (1000 x 10^{15} g C sediment versus 603 x 10^{15} g C for land and marine organisms) (Hedges et al., 2000). Understanding sediment molecular-level chemistry comprehensively is an important challenge if I am to ever realise the true impact this pool of organic matter has on current global cycles. However, this task has eluded scientists for decades. Approximately only about half of the sedimentary organic matter can be characterised employing a mixture of wet chemical, bulk physical and chemical techniques (Hedges et al., 2000).

This text is being prepared for submission to Organic Geochemistry.

Most analytical instruments that are typically available for molecular level studies examine only a small fraction of whole sediment samples, mainly because they require solubility and/or volatility. If extraction is not an analytical avenue, then most approaches have been to chemically or thermally degrade the organic matter to yield soluble and/or volatile products (Challinor, 2001; Clifford et al., 1995; del Rio et al., 1998; Goñi and Hedges, 1995; Hedges and Oades, 1997; Niemeyer et al., 1992). Use of non-invasive chemical analysis has always been a preferred approach to characterisation. Some studies have been of bulk composition using techniques such as Nuclear Magnetic Resonance (NMR) spectroscopy and Infrared Spectroscopy (Knicker, 2004; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Kögel-Knabner, 1997; Mopper et al., 2007; Niemeyer et al., 1992; Stevenson and Goh, 1971; Tremblay and Gagne, 2002). These bulk characterisation approaches can be used to describe the organic matter only in an average manner as in Tremblay and Gagne (2002), and to reveal the bulk changes in composition along a degradative pathway as in Knicker and Hatcher (2001). Other studies of sedimentary organic matter normally consist of a treatment of the sediment to remove a particular fraction from the whole and analyse it to learn about the source, degradation and composition. Sediment fractions traditionally extracted include those containing sulphur compounds (Filley et al., 2002), lipids (Canuel and Martens, 1996; Marsh and Weinstein, 1966; Sauer et al., 2001), proteins (Cowie and Hedges, 1992; Dauwe and Middelburg, 1998; Nunn and Keil, 2006), carbohydrates (Burdige et al., 2000; Gerchakov and Hatcher, 1972; Pakulski and Benner, 1992) and lignin (Dittmar and Lara, 2001; Goñi and Montgomery, 2000; Gordon and Goñi, 2004; Wysocki et al., 2008). Although often informative regarding the changes in sediments for these compound classes, these selective extraction techniques cannot provide information regarding
molecular changes for the bulk of sedimentary organic matter. With the introduction of some advanced analytical techniques capable of non-invasive molecular characterisation, this is now more readily possible.

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) has been used successfully to characterise organic matter in many different systems. These include dissolved organic matter (Hertkorn et al., 2006; Kujawinski et al., 2009; Mopper et al., 2007; Sleighter and Hatcher, 2007; Sleighter et al., 2009), pore waters (Schmidt et al., 2009; Tremblay et al., 2007), water soluble atmospheric molecules (Wozniak et al., 2008), meteorite organic matter (Schmitt-Kopplin et al., 2010a), rainwater (Altieri et al., 2009) and extracts of soil organic matter (Hockaday et al., 2006; Kujawinski et al., 2002). In order to apply this technique, solubility of the organic matter is required. Thus, choice of an appropriate solvent that allows for the maximum extraction of the organic matter being targeted is essential. This is especially a requisite for examination of sedimentary organic matter where the majority exists in a sparingly soluble form. To comprehensively assess the nature of sedimentary organic matter as a whole, it must be ascertained that the organic matter extracted is representative of the whole (Zhong et al., 2011). The fraction extracted must not only be sufficiently large but should have the same analytical characteristics as the whole sediment (Salmon et al., 2011). To adequately use the information obtained from FT-ICR-MS for whole sediment characterisation purposes it is imperative that I establish that the solvent used meets the above requirements. Such a comparison of possible extraction solvents has been investigated by Schmidt-Kopplin et al. (2009), when examining a meteorite sample by FT-ICR-MS, concluding that polar protic solvents such as methanol are most suitable.

The study described herein investigates five different possible extraction solvents
of varying chemical character including methanol, in order to determine analytically the
most appropriate solvent for molecular level analysis of some modern sedimentary
organic matter. Of important consideration for choice of an optimal solvent is removal of
a large fraction of organic matter, the degree to which the solvent extracted organic
matter is representative of the whole sediment organic matter, and diversity of
information obtained by FT-ICR-MS analysis. I chose for my study two organic-rich
lacustrine sediments that have been under detailed investigation by my group, one a
freshwater lake (Mud Lake, Florida, MLF) and the other a saline lake (Mangrove Lake,
Bermuda, MLB). While the high organic contents make these sediments unique among
typical lacustrine deposits, I feel that my strategy and results can be applied to the more
common, low-organic content sediments in both lacustrine and marine systems, because
the solvents used are specifically removing organic matter. To demonstrate the wider
applicability of my methodology, I apply my extraction protocol involving the suitably
determined solvent to two additional sediment samples of low organic matter content,
one from the Mississippi Delta (MD) and the other from Bayou Grande (BG) in
Pensacola, Florida. It should be noted that only thermally immature sediments will be
considered in this study, solvent extraction has been investigated in more mature
sediments (kerogen) elsewhere (Salmon et al., 2011).

2.2 Materials and Methods

2.2.1 Samples and their preparation

The two initially examined organic-rich sediments in this study, MLF and MLB,
are algal derived sediments and have been investigated in several past studies (Bates et
al., 1995; Boudreau et al., 1992; Filley et al., 2001; Filley et al., 2002; Hatcher et al., 1982; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; McKee and Hatcher, 2010; Orem et al., 1986; Spiker and Hatcher, 1984). Collection of the sediments analysed by MS from MLB was performed in 1982; the details are included in Orem et al. (1986). The MLB sediment (2.8 – 2.85 m depth) analysed by 1D $^1$H NMR was collected in November 2009, in a similar way to the 1982 sediments. The MLF sediment was collected in 1993, details of which are described in Filley et al. (2001). Two depths from the 1982 MLB (1.4 and 3.0 m) and one from MLF (0 m) sediment were used in this study for MS analysis; these were chosen based on the quantity of sediment available.

The two organic-poor sediments are from BG and the MD. BG is a tidal bayou, of approximately 4.3 km$^2$, with organic matter inputs from both fresh (terrestrial and marine) and anthropogenic sources (Lewis et al., 2001; Simpson et al., 2005). The collection of BG sediment (0.21 – 0.31 m, bottom) is described by Lewis et al. (2001). The area of the MD collection point is in the Southwest Pass of MD (Bennett et al., 1991), previously studied for its microstructure (Bennett et al., 1991; Bennett et al., 1977; Bennett and Faris, 1979; Bohlke and Bennett, 1980).

For FT-ICR-MS analysis all sediment sample extractions were treated the same and analysed using the same experimental conditions in order to fulfil a closer comparison between solvents. Approximately 50 mg of each sediment (see Table 1 for exact quantities) was used in each extraction; 1 mL of each solvent was added to each sediment and then extracted for 3 days (determined optimum time for maximum extraction, data not shown) at room temperature on an orbital shaker (VWR model 57018-754) at 280 rpm. The extract mixture was then filtered through a non-contaminating (Kauppila et al., 2006) 0.2 μm poly(tetrafluoroethane) (PTFE) filter to
produce a filtrate that was stored in a refrigerator (4 °C) until analysis. The following solvents were used: aqueous base (1M ammonium hydroxide, Fisher Scientific), chloroform (Fisher Scientific), methanol (Fisher Scientific), pyridine (Acros) and MilliQ water. All glassware was precombusted (to 450 °C) before use.

Table 1

The masses (mg) of the MLF, MLB, MD and BG sediments used for solvent extraction in FT-ICR-MS negative mode analysis.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>MLF 0 m</th>
<th>MLB 1.4 m</th>
<th>MLB 3.0 m</th>
<th>MD 0 m</th>
<th>BG 0.15 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aqueous base</td>
<td>45.76</td>
<td>NC</td>
<td>58.84</td>
<td>NE</td>
<td>NE</td>
</tr>
<tr>
<td>Chloroform</td>
<td>57.69</td>
<td>60.58</td>
<td>48.16</td>
<td>NE</td>
<td>NE</td>
</tr>
<tr>
<td>Methanol</td>
<td>61.53</td>
<td>49.73</td>
<td>63.92</td>
<td>NE</td>
<td>NE</td>
</tr>
<tr>
<td>Pyridine</td>
<td>52.19</td>
<td>49.21</td>
<td>62.60</td>
<td>NE</td>
<td>NE</td>
</tr>
<tr>
<td>Water</td>
<td>56.60</td>
<td>46.02</td>
<td>53.45</td>
<td>NE</td>
<td>NE</td>
</tr>
</tbody>
</table>

NC - not collected, lost upon processing sample.

NE - not extracted

For the analysis by NMR, each of the sediment samples of approximately 800 mg was added to 16 mL of each solvent. The sediments were extracted at room temperature on an orbital shaker at 250 rpm for 3 days, followed by removal of the extraction solvents from the sediment via filtration through a 0.2 μm PTFE filter. The solvents were evaporated under nitrogen and the solid residue was refrigerated at 4 °C until NMR analysis. This residue was redissolved in 600 μL of dimethyl sulfoxide-d6 (DMSO-d6, Acros, 99.9 % purity) prior to analysis. The sediment residues remaining after extraction were additionally utilised in the calculations of efficiency. Solvents were removed completely by blowing down with nitrogen gas followed by lyophilisation, after which sediments were weighed and used for elemental analysis.
2.2.2 Total organic carbon (TOC) and total nitrogen (TN) determination

TOC and TN was determined for the NMR residues of MLB 2.8 – 2.85 m sediments, prepared as described above and for MD and BG whole sediments. Sediments were analysed in triplicate; a mean value was determined and standard deviation was calculated. TOC and TN percentages were calculated by use of a 5 point calibration curve of nicotinamide (Sigma-Aldrich). This was accomplished by using a Thermo Finnigan Flash EA 1112 Series elemental analyser (CE Instruments, UK) that employs a combustion approach.

2.2.3 FT-ICR-MS analysis

All of the liquid extracts were diluted by 100 (except for high concentration water extracts, diluted by 10) with methanol (Sigma-Aldrich, LC-MS grade) and adjusted to slightly basic pH with ammonium hydroxide prior to analysis for negative ionisation MS. All FT-ICR-MS analyses were performed on a Bruker Daltonics 12 Tesla Apex Qe 175 FT-ICR-MS at the College of Sciences major instrumentation cluster (COSMIC) facility. The diluted extracts were infused into the Apollo II electrospray (ESI) ion source at a rate of 120 μL/hr using a syringe pump. Before addition of samples the instrument was externally calibrated using polyethylene glycol which was accumulated for 1.0 sec in the hexapole before being transferred to the ICR cell and 20 transients were co-added. Sample extracts were analysed using a 1.0 sec accumulation time in the hexapole with 300 co-added transients. ESI voltages were altered for each sample accordingly. A 4 MWord time domain was used for the free induction decay (FID). Zero-filling was completed once on the summated FID signal and then Sine-Bell apodised before Fourier transformation followed by magnitude calculation using the Bruker Daltonics Data
Internal mass calibration was achieved with sample-contained fatty acids using a previously described method (Sleighter et al., 2008). Molecular formulae were assigned using the Molecular Formula Calculator (Molecular Formula Cal. v. 1.0, (NHMFL, 1998)) with a maximum error allowed as 1 ppm. Unrealistic formulae were removed using a conservative set of rules described in Wozniak et al. (2008) developed for Matlab (v. 7.4.0, The Mathworks Inc., Natick, MA).

2.2.4 NMR analysis

For 1D $^1$H liquid state analysis of the MLB 2009 sediments, each of the DMSO-$d_6$ solvated extracts was analysed using solvent presaturation, in the case of the base, methanol and water extracts, the large residual water signal was suppressed, and for the chloroform and pyridine extracts DMSO-$d_6$ was suppressed. For each of the spectra, 128 scans were accumulated with 31248 data points. An exponential apodisation of 1 Hz was applied to each spectrum.

2.3. Results and Discussion

2.3.1 Quantification of extraction efficiency

An important issue when comparing different extraction solvents is the efficiency for removing the organic matter from the sediment. Determining the mass of the extract is one way to quantify the amount of sediment that is removed, after removal of the solvent. Table 2 lists the percentage of sediment removed by mass for each extraction solvent for MLB 2.8 – 2.85 m 2009 sediment. It is clear from Table 2 that chloroform...
and base remove less sediment by mass than the other three solvents. Unfortunately, mass determination is an unreliable method to use to calculate the efficiency of a solvent to remove organic matter from sediment due to three main reasons. First, because solvents such as base and water are particularly efficient at removing salts and inorganic carbon from the sediment which contribute to the mass of the extract but not the amount of organic matter removed. Second, there are some unquantifiable losses due to filtration and sample handling. Third, some solvents are difficult to remove completely from the organic extract. The amount of organic matter in each of the extracts is therefore, very difficult to quantify by mass alone. As an alternative, the elemental composition can be used to calculate the percent of C and N removed by the extracting solvent. The following equations demonstrate how this was performed for this study:

\[
\text{Percent C or N in extract} = \left[\left(\% \text{ C or N}_s \times m_s\right) - \left(\% \text{ C or N}_r \times m_r\right)\right]/m_s \times 100 \quad (1)
\]

\[
\text{Percent C or N removed} = \left[\left(\% \text{ C or N}_c \times m_c\right) / \left(\% \text{ C or N}_s \times m_s\right)\right] \times 100 \quad (2)
\]
Table 2

Information regarding bulk extraction calculations for different solvent extracts of sediments from MLB 2.8 – 2.85 m (2009) and whole sediments from MD 0 m and BG 0.21 – 0.31 m depths. The first column lists the percentage of sediment removed by mass compared to the amount of sediment originally used for extraction.

<table>
<thead>
<tr>
<th>Solvent extract</th>
<th>Percent removed by mass</th>
<th>Percent of C removed by extraction</th>
<th>Percent of N removed by extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLB - Aqueous base</td>
<td>33.9</td>
<td>18.6</td>
<td>20.6</td>
</tr>
<tr>
<td>MLB - Chloroform</td>
<td>10.7</td>
<td>27.5</td>
<td>27.7</td>
</tr>
<tr>
<td>MLB - Methanol</td>
<td>33.3</td>
<td>19.2</td>
<td>20.3</td>
</tr>
<tr>
<td>MLB - Pyridine</td>
<td>12.5</td>
<td>26.7</td>
<td>26.9</td>
</tr>
<tr>
<td>MLB - Water</td>
<td>35.0</td>
<td>18.1</td>
<td>20.2</td>
</tr>
<tr>
<td>MD - Pyridine</td>
<td>N/A</td>
<td>3.84*</td>
<td>1.88*</td>
</tr>
<tr>
<td>BG - Pyridine</td>
<td>N/A</td>
<td>12.3*</td>
<td>2.56*</td>
</tr>
</tbody>
</table>

N/A - not analysed

* - these values are absolute C and N amounts in the whole sediment because elemental compositions of extracts were not available.

Where % Cₜ and % Nₜ are the percents of C and N in the original sediment, mₛ is the mass of the sediment used for extraction, % Cᵣ and % Nᵣ are the percents of C and N in the residue left after extraction, mᵣ is the mass of the residue left after extraction, % Cₑ and % Nₑ are the percents of C and N in the extract and mₑ is the mass of the extract. The residue is defined as the solid left after removal of the extract and the extract is the sediment removed by the solvent. The calculated percentages are listed in Table 2. It should be noted that mₑ is determined by difference of the original sediment mass and the residue mass. Pyridine and chloroform extract the largest quantity of organic C and N, over 26 % C and N present in the original sediment, compared with extract yields of less than 19 % for the other solvents. Although a smaller mass of sediment is removed by pyridine and chloroform, the extract is richer in organic C and N than the other three solvent extracts, indicating that on the bulk level, these two solvents are more efficient.
organic matter extractors. The most probable reason for these differences is the higher extraction of inorganic C and N from the sediment by water, base and methanol than chloroform and pyridine, possibly saturating these solvents so that their ability to remove organic matter is reduced. The ultimate goal of this study is examination of sediments by FT-ICR-MS therefore, the most important determination of solvent extraction efficiency, in addition to the relative removal efficiency, should be related to the molecular functionality and diversity of structures contained in those extracts. Consequently, lesser importance regarding extraction efficiency determination should be placed on elemental composition of bulk sediments since it provides no information regarding the molecular complexity of the organic matter extracted by each solvent.

2.3.2 1D NMR analysis

A much better approach to evaluating the efficacy of various solvents for extracting representative fractions of sedimentary organic matter is use of an advanced analytical technique that is capable of assessing the various functionalities of compounds removed in each solvent extract. Proton liquid-state NMR analysis is ideally suited for this task and a spectrum of each solvent extract of MLB sediment collected in November 2009 is shown in Fig. 1. The goal is to determine which solvent removes the most representative fraction of organic matter from the sediment in terms of the dominant functionality, but also which one removes the larger number of compounds. Unfortunately this technique is not able to examine nonprotonated C containing functional groups; consequently, the following discussion considers only those functional groups with protons incorporated with the acknowledgement of the under-representation of unprotonated compounds. It is immediately clear when examining these spectra that
they are immensely complicated and require a more detailed study that is beyond the scope of this study in order to assign the functionality of each peak. Nonetheless, one can evaluate the relative amounts of the various functionalised structures by focusing on the absolute intensities of the various peaks.

The spectra are divided into different chemical shift regions to assist in their discussion (denoted by dotted lines). Region I, labelled as aromatic and phenolic, displays few peaks. The three solvent peaks of pyridine and one for ammonium hydroxide in the base extract do not prove to substantially overlap organic matter peaks, since they are located in the sparsely populated region I. It appears that the water and base extracts are proportionally extracting some aromatic compounds that the three organic solvents are unable to extract from the MLB sediment. Region II is assigned to unsaturated aliphatics and alcohol groups and also has few peaks falling within this region. This has been observed in solid state $^{13}$C NMR spectra and is an expected result considering the age of the MLB sediment in this study (Knicker and Hatcher, 1997; McKee and Hatcher, 2010). Region III displays several peaks from many functional groups including but not limited to esters, amides, carboxyl, and amine groups. Peaks deriving from carbohydrates, amino acids, fatty acids, and amines all have chemical shifts that fall in this region. The complexity and broadness of peaks in this region suggests that there are multitudes of compounds in these extracts, which are highly functionalised.
Fig. 2. 1D \(^1\)H liquid state NMR of intensity scaled solvent extracted MLB sediment from 2.8 – 2.85 m. 2.6 – 3.3 ppm is expanded and is inset to each spectrum. Peaks in inset spectra are labelled a – f and * indicates a solvent peak. Whole spectra are divided into functional group classifications regions I – IV.
A small section of this region is expanded (2.59 – 3.3 ppm) and is inset in each panel of Fig. 2. There are 6 different peaks in this region, associated with protons in environments of α-amine, α-amide and α-carboxyl/ester functional groups. The pyridine extract contains 5 of the 6 peaks, labelled a – f, indicating for this region at least it is a more efficient extracting solvent for these types of functional groups. It should be noted that peak d is a spinning sideband of the DMSO-\textit{d}_6 and therefore, is present in large amounts in those samples that have not selectively suppressed this solvent peak (water, base). The relative intensities of the 6 peaks vary considerably between extracts, for example, peak a is extremely dominant in the pyridine extract, but can barely be differentiated in the chloroform and methanol extracts, whereas in water it cannot be observed at all. The expanded region is generally characteristic of the rest of the spectrum in terms of the number of peaks resolved in each region for the different extracts. The reasons why certain peaks are larger in some extracts than others are difficult to ascertain due to the considerable complexity inherent in examining an organic matter sample. What can be gathered from examining this small expanded region is that pyridine and methanol seem to be extracting a larger fraction of the organic matter that can be analysed by NMR. The last region in these NMR spectra is region IV; that associated with protons in environments associated with aliphatic functionalities. Due to the fact that most of the organic matter in MLB is derived from algae, where long alkyl chains are a major component, this region is expected to dominate the spectra (McKee and Hatcher, 2010). The water and in particular, the base extract seem to be particularly inefficient at extracting these types of compounds from the sediment.

The predominance of solvent peaks in each spectrum makes the observation of some of the underlying sample peaks difficult, and, although every effort was made to
remove each of the solvents from the extract residue, small amounts remained to complicate the spectra. Ideally, the chosen solvent should possess minimal solvent peak interference with the organic matter sample. In this respect, methanol can be regarded as least valuable, since it contains two peaks at 3.16 and 4.01 ppm that overlap quite substantially with many organic matter peaks. Water and base extracts have similar overlap issues, although with only one peak that is an issue. It appears, from the standpoint of solvent peak overlap, that pyridine and chloroform are the most suitable.

A way to overcome the inclusion of solvent peaks in the spectra is to selectively integrate the peaks in each spectrum that are clearly assigned to organic matter in the extract. Table 3 shows the regionally-defined integrals for each of the five solvent extract spectra. The 0 – 1 ppm chemical shift range of selected sample integrals are similar for most of the samples except that of the water extract and the chloroform extract that are 100 times smaller and 10 times larger than the others, respectively. The aqueous base and pyridine extracts have similar values for this chemical shift range. As mentioned above, this region is assigned to the methyl protons of purely aliphatic compounds present in each of the samples. These types of compounds are expected to be a dominant compound class due to the algal nature of the sediment. The fact that water is displaying a substantially smaller quantity of this peak type leads me to believe that this extract is quantitatively less suitable as an extraction solvent than the other four solvents. The most intense region in the spectra is between 1 – 4 ppm; most of the peaks for this sample are located in this region because of the wide variety of compound classes whose chemical shifts fall here. Pyridine yields approximately 10 – 1000 times the peak intensity in this region as the other solvents, indicating that it is the best solvent to remove the highly structurally diverse level of compounds that are located in this
chemical shift range. Water is again displaying less intensity in this region than the other solvents and this is generally true for all regions of the spectrum. The ranges of 4 – 6, 6 – 8 and 8 – 10 ppm are generally lower absolute integrals than the upfield regions previously discussed, which is expected of the organic matter contents of this sediment.

Table 3

Absolute integral values of chemical shift regions of 1D $^1$H NMR spectra of solvent extracted sediment from MLB 2.8 – 2.85 m (2009). Sample peaks are selected and integrals for all peaks in each region are summed together and included in this Table.

<table>
<thead>
<tr>
<th>Solvent extract</th>
<th>NMR chemical shift region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 – 1 ppm</td>
</tr>
<tr>
<td>MLB - Aqueous base</td>
<td>6.43x10$^8$</td>
</tr>
<tr>
<td>MLB - Chloroform</td>
<td>1.35x10$^9$</td>
</tr>
<tr>
<td>MLB - Methanol</td>
<td>1.44x10$^8$</td>
</tr>
<tr>
<td>MLB - Pyridine</td>
<td>6.17x10$^8$</td>
</tr>
<tr>
<td>MLB - Water</td>
<td>4.36x10$^6$</td>
</tr>
</tbody>
</table>

The aqueous base, water and methanol all represent a lower intensity throughout the spectra and this is apparent in Fig. 2. Chloroform and pyridine both universally extract larger amounts of structurally diverse compounds from the sediments as is obvious from visual and quantitative inspection of the NMR spectra. Due to the higher number and intensities of peaks in the important region of 1 – 4 ppm, pyridine seems to be the best choice. This region is especially important when noting that this study is intended to identify the most suitable solvent for MS analysis, where ionisation of O and S aliphatic type containing compounds is most probable. This selective integration data is similar to what was observed on the bulk level in Table 2, chloroform and pyridine both extracted larger quantities of organic matter from the sediment. On the molecular level, as observed with NMR, pyridine seems to be the better choice, due to the larger
diversity of compounds observed (Fig. 2) and quantity of those compounds as determined by integration (Table 2 and 3).

2.3.3 FT-ICR-MS comparison for MLF and MLB sediment

Fig. 3 shows the FT-ICR-MS of MLF’s surficial sediment (depth of 0 m) by negative ion mode analysis using the five different solvents. All of these spectra were scaled to the same peak magnitude to highlight the differences in signal/noise (S/N) between each of these extraction solvents. As can be observed in each solvent extract, several hundred to thousand peaks are successfully ionised and consequently observed in this Figure. In the aqueous base extract of MLF 0 m, the magnitude of the peaks in this extract is less when compared to that in Fig. 3, those corresponding to organic extracts. The extraction with water produces a similar result corroborating with the results from the NMR data.
Fig. 3. Negative ion analysed FT-ICR-MS spectrum of MLF 0 m sediment solvent extracts. All peak magnitudes are expanded to the same scale.
One property that occurs in all of these extracts is that peaks normally assigned to fatty acids are of high peak magnitude (especially in the mass/charge (m/z) region of 250 – 300). Fatty acids are a ubiquitous group of compounds observed in many natural organic matter samples analysed using FT-ICR-MS (Goñi and Hedges, 1995; Sleighter et al., 2008), and are considered to be very easily ionised in the ESI source. It is for this reason the peak magnitude was expanded and set off-scale, so that less easily ionised peaks in the mass spectra could be observed. It appears that a greater similarity is observed in the larger magnitude peaks in the methanol and pyridine extracts, when compared with the other solvent extracts, a similar conclusion as that for the NMR data. Closer examination is required to investigate which solvent is ultimately the most appropriate for FT-ICR-MS analysis.

Fig. 4 displays the selected mass range of 352 – 353 m/z for each of the five extraction solvents. This mass range was chosen arbitrarily, designed to display a wide range of assigned formulae for each of these solvents. Peaks labelled “salt” are commonly observed in natural organic matter samples when analysed by FT-ICR-MS, particularly when samples originate from a high salinity environment (Hedges and Oades, 1997). The presence of salts is commonly viewed as detrimental to analysis by ESI because they out-compete organic matter for charge, thereby suppressing the signals for the organic compounds. (Clifford et al., 1995; King et al., 2000; Stenson et al., 2002), however, in this sample it does not appear to be as problematic except perhaps for the water and base extracts. There have been many attempts to assign elemental formulae to peaks labelled as “salt”, generally located at high mass defect, by including common elements that form salts in molecular formula calculations, e.g., chlorine, however, it has not been possible to assign formulae and is a topic for future work. “Salt” peaks only
appear to be an issue for the water and base extracts, and in sediments with higher salinity, water and base may become completely unsuitable due to this issue.

Examining the assigned formulae in Fig. 4 provides some more useful evidence to assist in deciding on an appropriate extraction solvent. Eleven different peaks are identified in this mass region between the five solvents. For the base and water extracts only peaks 2, 5 and 10 are removed from the sediment and visualised by MS. The three organic solvents, chloroform, methanol, and pyridine (Fig. 4) remove additional compounds that are not observed in the two aqueous extracts. The additional peaks include a $^{13}$C isotopologue, peak 9, ($^{12}$C$_{21}$H$_{42}$NO$_2$), CHO compounds, peaks 4 and 8, (C$_{22}$H$_{41}$O$_3$ and C$_{24}$H$_{49}$O$^-$), a CHOS compound, peak 6, (C$_{20}$H$_{33}$O$_3$S) and CHNO containing compounds, peaks 1 and 3, (C$_{22}$H$_{46}$NO$^-$ and C$_{21}$H$_{38}$NO$_3$). The organic solvents, particularly pyridine, possess a higher ability to remove more diverse formula types from the sediment based on their higher polarity and similar electrochemical properties.
Fig. 4. Negative ion analysed FT-ICR-MS expanded spectra (352 – 353 m/z) for MLF 0 m sediment. Peaks are labelled and formulae are listed in table insets for each spectrum.
The extracts from MLB at two different depths (0 m and 3 m) were also analysed by FT-ICR-MS. Their spectra are very similar when examined in a general way as was completed for MLF sediment in Fig. 3 and 4 and it will not be repeated for the MLB sediments. The same general characteristics and conclusions regarding solvent suitability are true for these two depths, including many of the same peaks observed in the expanded mass region in Fig. 4.

It should be noted that each extract was treated in the same way for analysis by FT-ICR-MS; this includes approximate mass of the sediment added to each solvent (and consequently, the same volume of solvent for each), the preparation of each extract prior to analysis (filtration and dilution with methanol) and analytical procedure with the FT-ICR-MS instrument. One of the possible reasons for the lower number of peaks observed in the aqueous extracts is that the dilution of the extracts prior to analysis reduces the concentration of the aqueous extracts so that the ionisation efficiency of the unobserved compounds is too low to be detected in the ICR cell. To assess whether this situation is occurring I analysed the water extract of MLB 3.0 m sediment at a higher concentration (factor of 10 increase). The higher concentration water extract ionises even fewer compounds (1106) than the low concentration extract (1302), thereby proving that the low number of peaks in the water extract is inherently low and not concentration dependent. The possibility exists that the salts present in this extract is suppressing the ionisation of the organic matter. It is interesting that the methanol extract also contains salts but they apparently do not have a strong influence on spectral quality. Perhaps, this is indicative of the fact that water is simply not extracting enough ionisable organic matter, and this, compounded by the salt problem, limits the usefulness of aqueous solvents for these sediments. The bulk C and N measurements in Table 2 confirm this.
From my examination of the mass spectra (Fig. 3 and 4) and NMR data (Fig. 2) it appears that the aqueous extracts are the least suitable solvents for extracting organic matter from organic-rich sediments. However, it is not clear which one of the three organic solvents is most suitable. Another piece of evidence for solvent suitability that can be used is examining the number of formulae that are identified in each spectrum, noting that the aqueous extracts contained many high defect “salt” peaks, probably reducing the number of peaks and consequently, assigned formulae. The total number of assigned formulae is displayed in Fig. 5. Fig. 5 clearly shows that water is universally removing the smallest number of molecular formulae for all three sediments. The large variability of the number of identified formulae of the base and chloroform extracts for these three sediments calls into question the reliability of these two solvents to provide the highest quality data in all samples. Pyridine and methanol both provide large amounts of information by FT-ICR-MS analysis identifying over 1250 different formulae for the three sediments.
When I consider all of the information obtained thus far for bulk composition, functional group diversity, representativeness of the original sediment and quality of MS data I now have enough information to disregard some of the solvents as sediment extractors. Water and base remove a less representative fraction of organic matter from the sediment, consequently displaying less structural diversity in both NMR and MS data and their MS spectra are of inferior quality, plagued with issues of “salt” interference and concentration. Chloroform seems to remove a similar amount of organic matter from the sediment as pyridine (Table 2), but the lower functional group diversity and unreliability of the MS data to provide consistently high levels of information, points to the unsuitability of chloroform as an extraction solvent. My remaining two solvents,
methanol and pyridine both seem equally well suited for solvent extraction on the molecular level as determined by the points listed above, although pyridine has a higher organic matter concentration determined by bulk and NMR quantification. Further MS analysis of the two solvent extracts is therefore warranted, in order to decide on the ultimate solvent for sediment extraction.

My more in-depth analysis of the pyridine and methanol extracts in the three sediments involved dividing the assigned elemental formulae into general classes. Fig. 6 displays this information as the sum of peaks in each class compared to the total number of peaks, calculated as a percentage. There are eleven classes of formulae identified in these sediments containing the commonly occurring elements, C, H, O, N, S and P. Although other elemental compositions could be possible, bulk elemental analysis (data not shown) leads me to believe that inclusion of additional elements will not yield many differences in my elemental assignments. It is immediately evident that the dominant formula type is CHO. As was discussed earlier, due to the fact that this analysis was completed in negative ion mode and that fatty acids are known to ionise very easily in the ESI source, CHO formulae should dominate negative spectra as they do so here (Cech and Enke, 2001). The next largest percentage of formulae identified for all of the sediments is for those containing CHNO (about 25 % assigned). Other N-containing formulae of significance are the ones containing CHNOS and CHNOP. Another significant formula is that of CHOS. Pyridine removes a higher number of CHNO-containing compounds than methanol for MLB sediments, while methanol seems to be more efficient for the MLF sediment. The ultimate goal of this study is to chose the solvent that provides the most comprehensive and detailed amount of molecular information regarding the sediment as possible, therefore, the one providing the most
diverse molecular diversity would be the more ideal solvent. Using all of the information obtained thus far, pyridine is the superior solvent to remove a representative fraction of the sediment, as demonstrated by NMR, more functional group diversity also observed in the NMR spectrum, reliable removal of high numbers of compounds, correlating to individual molecular formulae comprised of a large diversity of formulae types.

![Graph showing the total number of formulae types assigned in three sediments from MLF and MLB in methanol and pyridine analysed using negative ion FT-ICR-MS.](image)

Fig. 6. The total number of formulae types assigned in three sediments from MLF and MLB in methanol and pyridine analysed using negative ion FT-ICR-MS.

2.3.4 Application of chosen solvent to organic-poor sediments

Although pyridine is shown to be an effective extraction solvent for MLB and MLF sediments, it is important to test whether pyridine extraction can be used for other types of sediments. Therefore, the same extraction procedure and analysis by FT-ICR-MS is conducted on an additional two Gulf of Mexico sediments; BG and MD. These
two sites have different chemistries and lower organic matter contents to that of the lacustrine sediments previously studied, and therefore, should present a useful test to assess the wider applicability of pyridine extraction for sediments. Past studies of BG have investigated bulk nutrient levels such as total nitrogen, nitrite, nitrate and ammonia of the surface water and elemental characteristics of the sediment. The TOC and TN percentages (12.3 and 2.56, respectively, Table 2) are three to four times lower than that of MLF and MLB sediments (Filley et al., 2001; Hatcher et al., 1982), making molecular investigations more challenging. Solid state $^{13}$C NMR of the BG sediment revealed that alkyl functionalities dominate the organic matter, similar to that of MLB and MLF sediments of shallower depth (McKee and Hatcher, 2010; Simpson et al., 2005).

The other sediment analysed is from the MD. Deltas comprise a major fraction of sediments globally (44 % total C (Killops and Killops, 2005)). The MD is one such example, with a drainage area covering 40 % of the continental US and 70 – 90 % of its water contributes to the freshwater in the Gulf of Mexico (Coleman et al., 1998; Goñi et al., 1998; Gordon et al., 2001; Turner and Rabalais, 1994). The TOC has been measured throughout the Delta, with values of $< 2$ % (dry weight), while the TN values are generally $< 0.2$ % (dry weight) (Gordon and Goñi, 2004; Gordon et al., 2001; Turner and Rabalais, 1994). For the MD samples analysed here, TOC and TN values are slightly larger than those observed in previous studies, however the difference is not significant (Table 2). Several studies have been directed at identifying the source of the sediment organic matter entering the Delta, and while it is agreed that the Mississippi River is the main source, there is some difference of opinion regarding whether it is terrestrial or marine in nature (Goñi et al., 1998; Keil et al., 1997; Kennicutt II et al., 1995; Turner and Rabalais, 1994; Waterson and Canuel, 2008). The identity as to the exact molecular
composition of BG and MD sediment organic matter has not been further characterised.

Fig. 7 shows the resulting mass spectra of the pyridine extraction of BG and MD sediments. Fig. 7a displays the BG pyridine extracted sediment entire broadband mass spectrum, and Fig. 7c shows the complementary MD spectrum. These two spectra show that pyridine is able to extract a significant amount of organic matter, despite the fact that these two sediments have less organic matter within them. Comparing these two spectra to MLF extracts, Fig. 3, it is clear that these sediments have very different distributions of ionisable peaks as expected. Expanded mass spectra of the same mass region in Fig. 4, is displayed in Fig. 7b and 7d for BG and MD sediments respectively.

One important matter that needs to be verified is the reproducibility of the pyridine extraction, consequently BG and MD sediments were extracted and analysed in duplicate. The broadband spectra are very similar (data not shown), indicating that the sediment is homogeneous enough that pyridine extracted organic matter is similarly ionised in the ESI source. Comparing formula assignments between each duplicate spectrum using the Duplicate Remover™ add-in (ableBits.com) in Microsoft Excel indicates that 70 % of formulae for BG and 76 % for MD spectra are the same. When I increase the S/N threshold of the BG samples to a higher level (10) the percentage of identical formulae between the replicates increases to 81 %. I increased the S/N threshold in order to be certain that peaks falling just below this threshold are real sample peaks and not due to noise.
Fig. 7. FT-ICR-MS spectra of BG (0.21 m) ((a), expanded in (b)) and MD (0 m) ((c), expanded in (d)) pyridine extracted sediments analysed by negative ionisation mode. The assigned formulae to the peaks in (b) and (d) are labelled and listed in table insets.
The remaining 19 % of the BG assigned formulae that are not duplicated in both BG samples is attributable to a variety of problems including: peaks in one sample falling just below the S/N level (9.4 %), split peaks at high mass (5.3 %), assignment differences due to formulae that are close to the allowed assignment error (1 ppm) (3.5 %) and only 0.5 % is due to peaks that are not ionised in one of the duplicates at all. In only one of the samples of BG sediment extracted can split peaks be observed, and is due to a high concentration of ions in the ICR cell. When this occurs, a single peak appears as a double peak with an incorrect m/z output leading to an incorrect formula assignment. In comparing with the other sediment extracts, I use the BG replicate that does not contain split peaks. Therefore, there is an extremely small difference in peaks observed and assigned a formula in the two replicate spectra, indicating that pyridine extraction is very reproducible for a sediment sample analysed by FT-ICR-MS. Given this information, the peak variability that is observed between the four sediment samples are due to real differences in the organic matter of each sample that pyridine is fully capable of removing.
Fig. 8. Four part Venn diagram comparing assigned formulae for BG, MLB, MLF and MD pyridine extracted FT-ICR-MS analysed sediments. All percentages are calculated by comparison to the total number of assigned formulae for the four spectra combined.

A useful diagram to compare the differences in the assigned formulae is a Venn diagram, calculating the percent of unique or common formulae between each of the sediments. Fig. 2.7 displays the resulting 4 part Venn diagram, for MLF, MLB (3.0 m depth), BG and MD sediments' formula assignments. There are 3483 different assigned formulae when combining all of the mass spectra for the four sediments, each comparison in the Venn diagram is calculated as a percentage of the total number of assigned formulae for all spectra. Most of the formula percents are located in the unique formulae for each sediment, with the largest being that of BG at 23.3%. BG unique formulae are
approximately 50% CHOS containing, this large amount of S compared to the other three sediments is consistent with the fact that BG sediment is close to many anthropogenic sources (Bates et al., 2002; Brüchert and Pratt, 1999; Simpson et al., 2005). There is no dominant formula type that is unique to MD sediments. It is difficult to ascertain the exact reason but it could be associated with the fact that there is a mixed organic matter source contributing to the sediments of MD (Goñi et al., 1998; Gordon and Goñi, 2004; Gordon et al., 2001). The unique formulae for MLF are very heavily dominated (~90%) by CHO containing formulae. This is somewhat unexpected since past studies of MLF have not indicated a predominance of organic O; upon closer inspection of these formulae it is apparent that many are highly oxygenated and saturated. This could be explained by the fact that MLF sediment has not undergone substantial degradation especially since this sediment is one of the shallowest of all of the four sediments and predominately anaerobic. For MLB, the unique formulae that are assigned to this sediment are predominantly CHNO (~35%) and to a lesser extent CHOS (~20%) containing. In a previous study of MLB sediment, it was discovered that a large fraction of the CHNO containing formulae can be assigned a structure of alkyl amide (McKee and Hatcher, 2010). Many of the unique formulae for MLB are those that can be distinguished as alkyl amide.

The similarity of all four sediments is the largest percentage (4.9%) when comparing any of the sediments with each other. The common formulae for all four sediments are predominantly CHO (64%) containing. When examining the exact formulae many of the homologous series match those of fatty acids, generally considered to be ubiquitous for organic matter (Sleighter et al., 2008). 12% of the common formulae for all four sediments are CHNO containing. Several of these formulae can be
categorised as alkyl amide. Although MLB sediments contain the majority of alkyl amides, the match of formulae previously identified as alkyl amide by McKee and Hatcher (2010), in the two sediments from the Gulf of Mexico indicates that alkyl amides could be present in a variety of different sediment organic matter types. For the other sediment comparisons (2 – 3 sediments compared) most of the other common peaks are CHO containing; this is probably because of the ubiquitous nature of CHO containing compounds in lignin, waxes and oils commonly found in sediments (Hedges et al., 2000; Hedges and Keil, 1995). The fact that there are so many unique formulae for each sediment demonstrates the fact that pyridine extraction does not remove the same fraction of organic matter from any given sediment. Fig. 8 shows that it is possible to obtain meaningful molecular level information from a variety of different sediments.

2.4 Conclusions

This study was undertaken to determine an effective method to analyse the complex nature of sediment organic matter by FT-ICR-MS. Due to the need for samples to be in liquid form to be injected into the ESI source and analysed by FT-ICR-MS I investigated the most appropriate extraction solvent to provide the most comprehensive amount of information possible. Often time is an important factor to consider when analysing sediment organic matter using FT-ICR-MS; while it is ideal to use a multitude of solvents to extract the sediment organic matter, this is not always possible and so my study is designed to determine the most suitable solvent. Five different solvents were used to extract sediment initially from two different organic-rich lakes, MLF and MLB. Initial examination of these extracts by 1D $^1$H NMR revealed that the organic solvents
(particularly pyridine and methanol) were removing more compounds from the sediment. After further analysis by FT-ICR-MS, I ultimately determined that pyridine provided the most information regarding the organic compounds present in these two lakes. While I understand that this study utilises only negative ion mode analysis and that some compounds can additionally be observed in positive ionisation mode, many of the compounds I observed are O- and S-containing and can be identified in either ionisation mode. Under the assumption that the organic-rich sediments are sensitive to small differences in N composition, additional solvent testing was not undertaken in examination of further samples. To prove that pyridine is an effective solvent for other sediment types two additional organic-poor sediments were analysed by FT-ICR-MS, both from the Gulf of Mexico. I was able to successfully obtain molecular level information regarding these two sediments and in addition a comparison with the two lacustrine sediments revealed significant differences in their organic character. With the proof of wider applicability of pyridine as an extraction solvent for both organic-rich and poor sediments, the methodology described here can readily be used to study the complex nature of sediment organic matter in many locations.
CHAPTER 3

A METHOD TO INVESTIGATE THE N-CONTAINING MOLECULES IN SEDIMENTARY ORGANIC MATTER AS DETERMINED BY FOURIER TRANSFORM-ION CYCLOTRON RESONANCE-MASS SPECTROMETRY

3.1 Introduction

Sedimentary organic matter (SOM) is an important component of global elemental cycles. Nitrogen (N) is one such element, whose molecular signature has been under investigation for many years because of its complex role in biogeochemical systems (Goñi et al., 1998; Hedges et al., 2000; Hedges and Keil, 1995; Knicker and Hatcher, 2001; McKee and Hatcher, 2010; Schimmelmann and Lis, 2010). Numerous analytical techniques have been applied to probe the complex molecular and macromolecular chemistry of N within the challenging matrix of sediments, with varying levels of success (Challinor, 2001; Hatcher et al., 2001; Kögel-Knabner, 1997; Niemeyer et al., 1992; Sleighter and Hatcher, 2007). Studies of the organic N in modern sediments have concluded that amide functional groups predominate; these are mostly derived from peptides/proteins but some most recent studies indicate that alkyl amides are also present. 

This text is being prepared for submission to Journal of Mass Spectrometry.

important (Knicker and Hatcher, 1997; McKee and Hatcher, 2010; Nguyen and Harvey, 2001; Rillig et al., 2007). The ability to molecularly differentiate the chemical form of N-containing compounds, other than what is readily extractable in organic solvents and analyzed by gas chromatography mass spectrometry has emerged from the implementation of advanced multidimensional nuclear magnetic resonance (NMR) spectroscopy and electrospray ionisation coupled to Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS).

FT-ICR-MS has been successfully employed to study several types of natural organic matter, including dissolved organic matter (DOM) (Hertkorn et al., 2006; Koch et al., 2005; Kujawinski et al., 2002; Reemtsma et al., 2008; Sleighter and Hatcher, 2007), atmospheric water-soluble OM (Kroll et al., 2011; Mazzoleni et al., 2010; Wozniak et al., 2008), soil OM (Kramer et al., 2004; Kujawinski et al., 2002; Ohno et al., 2010), pore water DOM (D'Andrilli et al., 2010; Schmidt et al., 2009; Tremblay et al., 2007) and SOM (McKee and Hatcher, 2010). One of the biggest obstacles to studying sediments using ESI-FT-ICR-MS is the requirement of samples to be soluble in an appropriate solvent, usually methanol/water. From my recent study (Chapter 2) I determined that extraction into a suitable solvent capable of removing a representative fraction of the whole can provide a wealth of information regarding the OM composition (McKee and Hatcher, 2011; Zhong et al., 2011). Pyridine was the suggested solvent for extraction prior to ESI-FT-ICR-MS analysis as it was shown to provide the greatest extraction yield and NMR studies indicated that it was most representative of the whole organic fraction of the sediment.

The ultrahigh resolution of FT-ICR-MS makes it possible to assign individual elemental formulae to each ionisable peak. However, teasing out the most relevant and
informative formulae defining the N character of the OM from the assigned formulae is difficult to achieve without using elemental compositional diagrams to display various components of the FT-ICR-MS data. Several diagrams have been used in past studies, including the van Krevelen diagram (Kim et al., 2003a), Kendrick mass defect analysis (Hughey et al., 2001), carbon oxidation state (Kroll et al., 2011), double bond equivalent (DBE) versus oxygen (Kujawinski et al., 2009), DBE/C versus C (Hockaday et al., 2006) and N/C versus H/C (Wu et al., 2004). Each of these elemental diagrams have proved useful for study of various molecular-level aspects of organic matter, however, few have specifically selected N-containing formulae to gather information of the N-containing molecules present.

In this study, I introduce new types of plots for FT-ICR-MS data to specifically examine differences in the N-containing formulae among a few sediment samples. Traditionally employed diagrams will be assessed for their suitability and compared with new ones being introduced. A suitable diagram should be able to sufficiently compartmentalise N-containing formulae by level of unsaturation, N, O and C contents so that information regarding the nature and source of the N can be evaluated. Four sediments are used in this study from the following locations: Mangrove Lake, Bermuda (MLB), Mississippi Delta (MD), Mud Lake, Florida (MLF) and Bayou Grande (BG), Pensacola, Florida. Both organic-rich (MLB and MLF) and organic-poor (MD and BG) sediments are examined here in order to demonstrate the applicability of my approach to a wide range of sediment types.
3.2 Materials and Methods

3.2.1 Samples and their preparation

The two organic-rich sediments, MLB and MLF are lacustrine, algal derived and have been investigated in several past studies (Bates et al., 1995; Boudreau et al., 1992; Filley et al., 2001; Filley et al., 2002; Hatcher et al., 1982; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; McKee and Hatcher, 2010; Orem et al., 1986; Spiker and Hatcher, 1984). The two organic-poor sediments, MD and BG are both from the Gulf of Mexico region, BG from a tidal bayou in Pensacola, Florida harbour area, of approximately 4.3 km², with OM inputs from both fresh (terrestrial and marine) and anthropogenic sources (Lewis et al., 2001; Simpson et al., 2005) while MD sediment was collected in the Southwest Pass (Bennett et al., 1991), with OM inputs mainly from the Mississippi River (Bennett et al., 1991; Bennett et al., 1977; Bennett and Faris, 1979; Bohlke and Bennett, 1980)

Approximately 50 mg of each sediment was extracted into 1 mL of pyridine over a period of 3 days (determined a sufficient time to remove maximum quantity of sediment, data not shown) at room temperature on an orbital shaker (VWR model 57018-754) at 280 rpm. After this time, the extract was filtered through a 0.2 μm poly(tetrafluoroethane) filter and stored in a refrigerator (4 °C) until analysis completed on the same day.

3.2.2 FT-ICR-MS analysis

Each of the extracts was diluted with methanol (Sigma-Aldrich, LC-MS grade) by 100 fold and adjusted to slightly basic pH with ammonium hydroxide (Fisher)
immediately prior to analysis for negative ionisation MS. All ESI-FT-ICR-MS analyses were performed on a Bruker Daltonics 12 Tesla Apex Qe 175 FT-ICR-MS at the College of Sciences Major Instrument Cluster facility at Old Dominion University. The diluted extracts were continually infused into the Apollo II ESI source at a rate of 120 μL/hr using a syringe pump. Before analysis of samples the instrument was externally calibrated using polyethylene glycol which was accumulated for 1.0 sec in the hexapole before being transferred to the ICR cell and 20 transients were co-added. Sample extracts were analysed using a 1.0 sec accumulation time in the hexapole with 300 co-added transients. ESI voltages were altered for each sample accordingly. A 4 MWord time domain was used for the free induction decay (FID). Zero-filling was completed once on the summed FID signal and then Sine-Bell apodised before Fourier transformation followed by magnitude calculation using the Bruker Daltonics Data Analysis software v. 1.3™.

Internal calibration was achieved using fatty acids present in the samples for negative ion mode using the method described in Sleighter et al. (2008). Molecular formulae were assigned using the Molecular Formula Calculator (Molecular Formula Cal. v.1.0, (NHMFL, 1998)) with a maximum error allowed as 1 ppm. Unrealistic formulae were removed using a conservative set of rules described in Wozniak et al. (2008) developed for Matlab (v. 7.4.0, The Mathworks Inc., Natick, MA).

3.3 Results and Discussion

My investigation began with an examination of the types of N-containing formulae observed in each sediment extract. Assigned elemental formulae can be divided
into different formula classes depending on their heteroatom contents, and is done so in Fig. 9, considering only those formulae containing N. Percentages are calculated based on the total number of assigned formulae for each sediment extract, those containing CHO being the most abundant in all cases. It is clear from this Figure that CHNO containing formulae are the dominant N-containing formula type for all four sediments. The difference between organic-rich and organic-poor sediments is obvious from this Figure since the percent of N-containing formulae on the whole is lower for MD and BG sediment extracts compared with the organic-rich samples. Such a trend is expected because the organic-rich sediments are deposited in anaerobic environments where N-containing sources, e.g., proteinaceous materials, are expected to be preserved to a greater extent than in organic-poor sediments from coastal areas which typically tend to be more aerobic. CHNOS and CHNOP are the next largest N-containing formula types for the four sediments, although the exact abundances vary, probably due to ionisation efficiency differences and OM compositions in each sediment. This Figure gives some detail about the N-containing formulae assigned in each sediment extract but does not give any information regarding the exact nature of the formulae and any possible links regarding their structure or source.
Fig. 9. Charts displaying the abundances for the assigned formulae containing N in FT-ICR-MS negatively analysed pyridine extracts of 4 sediments. The relative abundances of formulae in each sediment are calculated as a percentage of the total number of assigned formulae for each spectrum, including those of the most abundant CHO formula types (not shown).

In order to focus on the most important information that can be derived from the FT-ICR-MS data, I will only consider the dominating N-containing formulae which are CHNO containing, but my approach is applicable to other formula types that contain N. Only one sediment is studied in detail (MLB) since it contains the greatest abundance of N-containing formulae of the 4 sediment extracts. However, one can apply the same methodology to other samples. Table 4 shows the general CHNO formula types that are observed in MLB. This Table lists every formula type in the CHNO class and sorts them according to series in which a common feature exists, namely the number of N and O
atoms in the formula. Only formulae with more than two peaks within each series are listed, since three formulae are required to form a realistic series. There is a large variety of formulae observed in this sediment as shown in Table 4; demonstrating not only the ability of pyridine to extract these compounds from the sediment but also the value of FT-ICR-MS as an analytical tool to observe them. The ability to observe such a diverse number of N-containing formulae is unique for solvent extraction, and is due to the high extraction efficiency of pyridine. The number of formulae in each series varies considerably ranging from 3 to 58. In cases where a large number of formulae are identified within a general formula series, I observe several types of homology that include varying levels of unsaturation as defined by the relative abundance of C and H atoms in the series. This suggests that several of the formulae have similar structures probably because they are sourced from the same type of compound. Although a trend in the number of assigned formulae for a particular formula type may be difficult to recognise at this stage, in general, it appears that there are fewer assigned formulae in which higher numbers of N and O are present. This result is reasonable when considering that the source of the organic matter in MLB is primarily algae and is not diverse. Perhaps this singular source type confers a chemistry composed mainly of compounds containing lower numbers of heteroatoms.
Table 4

The general formulae of CHNO containing compounds (series with > 2 formulae identified) and the number of formulae observed in that series by FT-ICR-MS negative mode analysis of the pyridine extract of MLB sediment.

<table>
<thead>
<tr>
<th>General formula</th>
<th>Number in series</th>
<th>General formula</th>
<th>Number in series</th>
</tr>
</thead>
<tbody>
<tr>
<td>C₅H₆NO³</td>
<td>24</td>
<td>C₅H₆N₃O</td>
<td>6</td>
</tr>
<tr>
<td>C₅H₆NO₂ᵃ</td>
<td>58</td>
<td>C₅H₆N₃O₃</td>
<td>5</td>
</tr>
<tr>
<td>C₅H₆NO₃ᵃ</td>
<td>53</td>
<td>C₅H₆N₃O₄</td>
<td>11</td>
</tr>
<tr>
<td>C₅H₆NO₄</td>
<td>35</td>
<td>C₅H₆N₄O</td>
<td>15</td>
</tr>
<tr>
<td>C₅H₆NO₅ᵃ</td>
<td>17</td>
<td>C₅H₆N₄O₂</td>
<td>3</td>
</tr>
<tr>
<td>C₅H₆NO₆</td>
<td>10</td>
<td>C₅H₆N₆O²ᵃ</td>
<td>30</td>
</tr>
<tr>
<td>C₅H₆N₂O₂</td>
<td>5</td>
<td>C₅H₆N₆O₂</td>
<td>26</td>
</tr>
<tr>
<td>C₅H₆N₂O₄</td>
<td>8</td>
<td>C₅H₆N₇O</td>
<td>9</td>
</tr>
<tr>
<td>C₅H₆N₂O₅</td>
<td>3</td>
<td>C₅H₆N₇O₂</td>
<td>10</td>
</tr>
<tr>
<td>C₅H₆N₂O₆</td>
<td>4</td>
<td>C₅H₆N₇O₃</td>
<td>11</td>
</tr>
<tr>
<td>C₅H₆N₃O</td>
<td>6</td>
<td>C₅H₆N₇O₄</td>
<td>3</td>
</tr>
</tbody>
</table>

ᵃ Series chosen for subsequent comparison.

Unfortunately, due to the complex nature of the data that are produced with FT-ICR-MS it is sometimes necessary to focus on just a few compounds in order to derive meaningful differences between samples. Accordingly, I decided to focus on five series of CHNO compounds (a subset of those indicated in Table 4) with varying numbers of assigned formulae within each series. Diagrams commonly used to study FT-ICR-MS data generally focus on examining trends in CHO containing formulae. Some of these diagrams will be utilised in this study but will be examined in regard to the information available for N-containing formulae in particular, the five series chosen above. The diagram that most readily presents useful information regarding the characteristics of the five series will then be used to compare all 4 sediment samples in this study, and used to make inferences regarding their source and degradative pathway. The key components that will be assessed in each diagram investigated include: ability to observe differences
of formulae in N contents, O contents, saturation level and separation according to likely sources. Additionally, the most suitable diagram would be one that examines the N characteristics of the sample without pre-sorting of N-containing assigned formulae. The ability to examine only the molecular N characteristics of this sediment is possible due to the ability of pyridine to extract a representative fraction of the sediment as demonstrated by NMR (McKee et al. 2011).

The van Krevelen diagram is commonly used in OM studies in which FT-ICR-MS data are employed (Kim et al., 2003a; Schmidt et al., 2009; Sleighter and Hatcher, 2007; Sleighter and Hatcher, 2008; Wozniak et al., 2008). A van Krevelen plot of the five series of CHNO formulae is shown in Fig. 10. The van Krevelen diagram provides a simple and convenient way of evaluating the compound source represented by the formula. Thus, I can determine if a peak and its assigned formula belong to a class of compounds whose broad structural makeup is lipid, lignin, carbohydrate, protein, or condensed aromatic. It is apparent, that the formulae displayed in Fig. 10 primarily group in the lipid region of the van Krevelen diagram and that they lack aromaticity as would be inferred by a low H/C ratio. When considering that the nature of the organic matter source is mostly algal and the fact that these sediments are thermally immature, the location of these compounds on the van Krevelen diagram is as expected. Within each series, I generally observe that the points lie along a trend line that intersects an ordinate value of 2.0 but also some formulae lie in a vertical orientation to each other indicating that they differ primarily in their H/C molar ratio. This implies that many of the formulae within the series differ by their CH₂ content or, in the case of vertically aligned formulae, by their degree of unsaturation. For example, the formula type CHNO₃ is located to the right of the rest of the points, with a higher amount of O per C atom. In
addition, the points are distributed in a line indicating homology among the series.

![van Krevelen diagram](image)

Fig. 10. van Krevelen diagram of select CHNO-containing formulae (as listed in the key) of the pyridine extract of MLB sediment analysed by FT-ICR-MS in negative ion mode. The plot is divided into general regions to indicate the likely structure of formulae displayed.

However, none of this information refers to the N content of the formula, and without prior selection from other non-N-containing formulae it would not be clear that these points contain N in their formula. In addition, if N-containing formulae are of the same type (i.e., lipid/protein) then the points in several different series overlap making it difficult to make further inferences about their character. The main characteristics that can be gleaned from this diagram for the N-containing formulae plotted are that they are protein/lipid in nature and that there appears to be series of homology. However, I wondered if additional structural information could be contained within the dataset, prompting me to seek other possible plots that would provide additional information for these series of CHNO compounds.
Another common diagram plotted to identify trends in FT-ICR-MS data involves Kendrick Mass Defect analysis (Hughey et al., 2001; Kendrick, 1963). Kendrick mass rescales the IUPAC mass of a particular formula by a repeating unit (e.g., CH$_2$ and O$_2$). Using the Kendrick mass, the Kendrick mass defect can be calculated easily, identifying homologous series of formulae because they have the exact same Kendrick mass defect. This type of diagram can be useful for exploring N-containing formulae, and is investigated for the MLB sediment homologous series identified in Table 4; Fig. 11 is the result.

Fig. 11 clearly shows the homologous series for each of the 5 formula types, as would be expected of a Kendrick mass defect analysis. It appears that generally heteroatom poor N-containing formulae (e.g., CHNO) have higher mass defect and lower nominal Kendrick mass, therefore, falling in the top left corner of Fig. 11. The van Krevelen diagram hints at homology but the Kendrick mass defect analysis clearly shows each individual series along a horizontal line. Where the data fall on this line indicates the level of saturation and the mass range of that series. The main advantage of Kendrick mass defect analysis is the ability to identify homology of a series: however, it does not specifically target N-containing formulae unless prior selection is done as is shown above. Without individually identifying each data point it is not possible to ascertain whether it contains N and what its source character could be. It is possible prior to constructing a plot of the data to manually select those formulae that contain N thereby ensuring that the N characteristics are under investigation. However, this type of pre-selection is not conducive to a rapid method to analyse the MS data; a more useful diagram would be one which specifically targets those formulae that contain N as part of its calculations.
Fig. 11. Plot of Kendrick mass defect vs. nominal Kendrick mass for MLB pyridine extracted sediment analysed by FT-ICR-MS negative ion mode. 5 CHNO containing formulae are included and separated according to the key.

DBE calculation is a useful measure of the level of saturation of an assigned formula, $C_{c}H_{h}N_{n}O_{s}S_{p}$ and is calculated by the following equation:

$$DBE = \frac{[(2 \times c) - h + n + p + 2]}{2}$$

DBE normalised to C (DBE/C), was first used to analyse condensed aromatic ring structures in soils exposed to fire (Hockaday et al., 2006) and may prove useful in this study due to its dual inclusion of saturation level and N content (albeit N contents have less influence on this value). Therefore, a plot of (DBE)/C against elemental O/C provides an additional manner to compare the various formulae within each series for the
selected CHNO compounds displayed in Fig. 12. In this Figure the five selected CHNO series are distributed across the axes depending on their level of saturation (DBE), number of N (included in DBE calculation), and number of C and O. In addition, since this is a ratio plot, homologous series are offset diagonally making them easier to identify. All of the homologous series trend lines intercept 0, meaning their slope is:

\[
\text{Slope of homologous series trend line} = \frac{1}{\text{DBE}} \times \text{amount of O} \tag{4}
\]

The result of equation 4 means that the amount of O in any particular formula has a large effect on the slope of that particular homologous series. The fact that the origin of all slopes is at 0, makes this diagram unique compared to the above considered plots. Consequently, the fully saturated CHNO series has a slope of 1 and can be used as a central line (S = 1) in the plot to further divide into general regions as shown on Fig. 12 (labelled S = 2.5, 0.5 and 0.25, with the number denoting the slope). Data points located to the left of S = 2.5, are those that have high amounts of O compared to N and are completely saturated. Data points between the slopes of 2.5 and 1 have more O than N but these have 1 or 2 unsaturated bonds. Data points between S = 1 and 0.5, are saturated series but with more N than O. Below the slope of 0.25 represents those series that have one or two unsaturations and have substantially more N than O in the formula. Another feature of this diagram is that it allows me to clearly distinguish homology within a series as a linear trend among the formulae that are related. Such a trend is also observed in the traditional van Krevelen diagram and Kendrick mass defect analysis diagram discussed above, particularly Kendrick mass defect analysis. Therefore, this diagram does not possess additional information regarding series homology. For example, the CHNO₃
formula type has three clear homologous series with one, two and three unsaturation points. I can speculate that the formulae within each series are derived from the same structural entity differing only in CH$_2$ and that the three series are made up of different functional groups. For example, formulae with one unsaturation could contain an amide and three hydroxyl groups. Although all formulae in this series are homologous in formula, the positions of the three functional groups might vary along the chains of the series and there maybe branching of the aliphatic chain. Since Fig. 10 indicated that all of these series are lipid and protein like, due to the low saturation level of the CHNO$_3$ series discussed above, they are probably lipid like formulae, with functional groups located at the ends of the aliphatic chains.

Fig. 12. Diagram plotting the ratio of DBE/C against O/C of FT-ICR-MS negative ion mode analysed pyridine extract of MLB sediment. Five series of CHNO containing formulae are included on this diagram as displayed in the key. Annotations are provided on the diagram indicating homology of series; these lines are labelled according to their slope (E.g., S = 2.5 indicates a slope of 2.5).
Although the DBE/C vs. O/C plot provides information regarding the nitrogen and oxygen contents, the level of unsaturation and homology of an assigned formula, pre-selection of N-containing formulae is necessary in order to investigate only their characteristics. As mentioned above pre-selection of N-containing formulae is necessary for the van Krevelen and Kendrick mass defect analysis diagrams, therefore, Fig. 12 does not present the ideal diagram for investigating only N-containing formulae. A plot that involves N as a key part of the calculation for one of the plot axes will select those formulae that contain N. A diagram plotting O/C against N/H is one example. Any formula that does not contain N or O will be located on either the x or y axis, meaning that pre-sorting of formulae is not necessary for this plot. The co-inclusion of N and O in a formula is expected for MS data analysed by negative ion mode, due to the ionisation requirements of the ESI source; Fig. 9 illustrates the dominance of this type of formula for the four sediment extracts in this study. Fig. 13a shows the resulting O/C vs. N/H diagram for the five series indicated in Table 4. As can be observed from this Figure, the five series are distributed in different regions of the plot primarily according to their N and O contents, similar to Fig. 10 and 12. In addition, the level of saturation in a formula is a factor in determining the position of a data point in the diagram (since H and C contents both determine the slope), therefore, series homology, with formulae differing by CH$_2$ can also be differentiated. One homologous series is labelled (CHNO$_5$ series labelled homology) on the Figure to demonstrate that the slope of this line originates at 0, similar to Fig. 12 and that all formulae in this series fall on the same line. The trend line at which the number of N equals the number of O in a series is labelled N = O on the Figure and has a slope of 0.5, to be used as a central point to determine the general characteristics of the formulae assigned in a sample. Points below this line are more O
rich than N, and above this line, they contain more N than O. This Figure therefore, has all of the requirements listed above in order to plot quickly and informatively, the N characteristics of lipid like formulae of FT-ICR-MS analysed data.

The five series selected in Fig. 13a demonstrate the information that can be obtained for lipid like formulae. Although in these five selected series, lipids are the likely predominant source, peptides and aminosugars could also contribute. Peptides have been considered in some samples to be the dominant type of N species and are therefore, important to consider when plotting N-containing formulae (Knicker and Hatcher, 2001). In addition, some aminosugars, namely muramic acid and N-acetyl glucosamine, are known structural components in bacteria and consequently, these and other aminosugars have been identified in natural OM samples (Aluwihare et al., 2005; Benner and Kaiser, 2003; Niggemann and Schubert, 2006). These two classes of compounds were plotted on an O/C vs. N/H diagram in addition to three series previously plotted in Fig. 13a for comparison. Five peptides are chosen to plot on this Figure each containing 5 identical amino acids that represent the most aromatic, aliphatic, N, O and S rich in nature, illustrating the maximum area that peptides could be expected to be located. Example aminosugars that have been identified in natural organic samples (Benner and Kaiser, 2003; Niggemann and Schubert, 2006), and bacteria (Salton, 1965) that are likely contributing to OM are plotted on this Figure, assuming that formulae sourced from these types of structure maintain approximately the same amount of N, O, C and H contents as the original source. Each class of compounds is located in distinct regions in Fig. 13b, similar to the areas of the traditional van Krevelen diagram.
Fig. 13. O/C vs. N/H diagrams of negatively analysed MS data of MLB extracted sediment. (a) displays five selected formulae as shown in the legend, and (b) shows three of the same formula series in addition to formulae examples of peptides and aminosugars. Annotations are provided on both diagrams showing on (a) their characteristics in terms of their N and O contents and on (b) the positions of the formulae from each source. The peptides used as examples are labelled on the Figure, and the aminosugar examples are numbered and identified in the legend on the right.
Although it appears that the peptide region is extensive it should be kept in mind that the 5 example peptides are extreme points and that it is likely that the majority of data points for a real sample will be located in the centre of this region. The lipid region is divided into two sections (N rich and poor) although it is probable that data points will be located between these two regions. Fig. 13a and 13b illustrate that an O/C vs. N/H diagram is ideal for studying the N characteristics of any organic matter sample analysed by FT-ICR-MS. It selects only those formulae that contain N, is able to separate formulae based on N and O contents, displays homology of series based on their level saturation and clear regions are defined based on these characteristics for the most likely compound types present in samples. Therefore, this diagram will be used to plot all the data from the analysis of the 4 sediment extracts and I will specifically examine their N characteristics.

Fig. 14 displays the resulting O/C vs. N/H plots of all assigned formulae in the four sediment extracts without pre-selection of N-containing formulae. The plots on the left show the entire O/C vs. N/H diagram of each of the sediment extract MS data while the plots on the right show only the N-poor lipid region. As shown in Fig. 13b, the entire O/C vs. N/H diagram can be divided into sections according to formula types that can be linked to certain types of structures. Each of the four plots of the entire diagrams show that although they have similar N amounts according to their organic-rich (MLB and MLF) and -poor (MD and BG) contents (Fig. 9), the structural types differ significantly. N-poor lipids are the dominant formula type in all of the four extracts, although the amount and number of homologous series vary. None of the samples exhibit significant aminosugar content which is expected since sugars are generally not well ionised by the ESI source.
Fig. 14. O/C vs. N/H diagrams of 4 sediment extracts analysed by negative ion mode FT-ICR-MS. Left side plots are the entire plot, with regions regarding likely sources labelled. Right side plots are expanded N-poor lipid regions with CHNO series slopes annotated. Vertical formula saturation homology is indicated by a box (on MLB plot).
MD contains the most peptide-like formulae of the four samples, probably due to its extensively variable sources from the Mississippi River. Although not as widely distributed as the MD sample, BG displays several peptide-like formulae. MLB and MLF plots look similar, which is not surprising since they have similar OM sources (algae) and are both organic-rich sediments. N-rich lipids are present in similar amounts for the four extracts, perhaps suggesting their ubiquity in sediments. Each O/C vs. N/H diagram shows series homology, a notable feature of this diagram.

Since N-poor lipid like formulae are universally the most observed formula type in the four samples, this region is expanded and is displayed in the plots on the right of the Figure. Several trend lines are drawn on these Figures to show the positions of each of the series that seem to dominate (CHNO, CHNO₂, CHNO₃, CHNO₄ and CHNO₅). The plot of MLB seems to display the most homology for these 5 series and for the rest of the diagram in general. The lack of series homology in the N-poor lipid region for BG, and MLF samples suggests that the source and reactions of the organic matter are very diverse producing a wide variety of formula types. Many of the CHNO formulae identified in these four samples are probably alkyl amides, previously identified in MLB sediment (McKee and Hatcher, 2010). Possible structures of the formulae CHNO₂, are likely a nitro group or an amide, with an additional hydroxyl or carbonyl group, attached to an aliphatic chain. Structures of formulae containing CHNO₃, CHNO₄ and CHNO₅ likely contain combinations of these functional groups in addition to a long aliphatic chain.

Noticeable in several locations of the N-poor lipid region O/C vs. N/H diagrams are vertical lines of data points, especially in the BG, MD and MLB samples. One of these sets of data points is highlighted in the MLB plot and labelled saturation series.
The formulae correlating to these data points are C$_{22}$H$_{37}$NO$_3$, C$_{22}$H$_{39}$NO$_3$, C$_{22}$H$_{41}$NO$_3$, C$_{22}$H$_{43}$NO$_3$ and C$_{22}$H$_{45}$NO$_3$. Therefore, vertically aligned data points correlate to the same formula but with varying amounts of unsaturation, similar to the traditional van Krevelen diagram. This pattern of unsaturation of the same base formula in three of the four extracts studied here implies that either there is a common OM source that produces similar structures or that there is a stepwise removal of H to produce a variety of formulae that differ only by their level of saturation. If the stepwise removal of H via a dehydrogenation reaction is occurring in this sediment and in BG and MD sediments, this could be molecular evidence of diagenesis.

3.4 Conclusions

Four sediment extracts were analysed by FT-ICR-MS negative ion mode to study their N-characteristics on the molecular level. It is important that the high amount of information produced by FT-ICR-MS analysis of these samples and others is utilised in the most effective and efficient way. It is very difficult to visualise and obtain meaningful data for the large dataset obtained, especially with regard to the N-containing compounds. The traditional van Krevelen diagram is very effective for studying possible sources of CHO formulae and their series homology, however, this diagram is not specific for N-containing formulae without pre-selection. Pre-selection of N-containing formulae is also necessary for the plotting of data in a Kendrick mass defect analysis and a DBE/C vs. O/C diagram. Both of these plots display homology of CH$_2$ series effectively, by plotting data points in a series either horizontally (Kendrick) or at a slope originating at 0 (DBE/C vs. O/C). While the DBE/C vs. O/C plot was able to separate
likely sources of assigned formulae, the Kendrick mass defect analysis plot is less able.
A fourth diagram is proposed with the idea to incorporate all of the strengths of the
previous 3 diagrams. An O/C vs. N/H specifically plots N-containing formulae through
the initial calculation of the ratios plotted. In addition, homologous series for these
formulae can be readily observed with slopes originating at 0 so that patterns regarding
the nature of compounds in a particular sample can be determined. Similar to the
traditional van Krevelen diagram, the modified version for N-containing formulae plots
compound types in different regions that give details regarding the type of organic matter
present. CHNO formulae are the dominant N-containing formulae, and of these, N-poor
lipids are the predominant type in all four sediment extract samples. Although I am
unable to ascertain the exact reasons for the presence of the formulae identified it appears
that there are several patterns regarding degradative pathways that are readily observed
with the O/C vs. N/H diagram. This methodology for the analysis of sediment samples
by negative ionisation FT-ICR-MS with the use of my proposed O/C vs. N/H diagram
can be used for other natural organic matter samples to quickly and efficiently evaluate
the nitrogenous species in large numbers of samples. By employing this diagram in
future studies, new information can be obtained regarding specifically the N-containing
formulae assigned from FT-ICR-MS analysis, and the relative importance of N-
containing structural entities such as peptides, N-containing lipids and aminosugars can
be determined.
CHAPTER 4

ALKYL AMIDES IN TWO ORGANIC-RICH ANOXIC SEDIMENTS: A POSSIBLE NEW ABIOTIC ROUTE FOR N SEQUESTRATION

4.1. Introduction

Fixed nitrogen (N) plays an integral role in global cycling of biologically important elements. N is an unusual nutrient since it is transformed through a variety of oxidation states by many different microbes, forming several N species ($\text{NO}_2^-$, $\text{NO}_3^-$, $\text{NH}_2^-$, $\text{NH}_4^+$, $\text{N}_2\text{O}$, organic N and $\text{N}_2$) in what is known as the N cycle (Arrigo, 2005; Hulth et al., 2005; Knicker, 2004; Zehr and Ward, 2002). At the heart of the N cycle is ammonia, used by primary producers to fix organic N. Upon death of primary producers and other organisms that incorporate N through various trophic levels, the organic N is available for remineralisation which occurs within a matter of days to produce energy and form inorganic N, mostly ammonia/ammonium. The small fraction ($\sim 1\%$) that escapes remineralisation to become recalcitrant organic N can play a huge role in N.

The following text was published in Geochimica et Cosmochimica Acta in 2010. See below for the full citation. The copyright permission is included in Appendix A.

bioavailability within the entire biosphere since it can remain sequestered for thousands to millions of years, mostly in sediments (Hedges and Keil, 1995).

Despite extensive study, a substantial fraction of the total organic nitrogen sequestered in sediments remains unidentified. There have been many tentative suggestions made as to the functional form of recalcitrant N in sediments. These include heteroaromatic N functionalities which have been observed in coals (Knicker, 2004), chitinous residues (Baas et al., 1995), N-acetyl amino polysaccharides in high molecular weight dissolved organic matter (Aluwihare et al., 2005) and proteins/peptides (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Zang and Hatcher, 2002). Past studies using $^{15}$N Nuclear Magnetic Resonance (NMR) (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001) proposed that nearly all of the N in some thermally immature organic-rich sediments, such as the sediments in this study and in many other sediments and soils as well, exists as amides (Knicker, 2004). The identity of this amide functionality belongs to proteins and/or peptides but other amides such as amino sugars are also possible. In contrast, N-compounds in thermally mature sediments are generally considered to be of the pyridinic and aromatic form (Boudreau et al., 1992; Knicker et al., 2002; Schimmelmann and Lis, 2010).

Proteins/peptides are normally one of the more labile components in sediments, as are amino sugars, and so preservation of these labile amide components for hundreds to thousands of years seems remarkable. To explain this long-term preservation of proteins/peptides, several suggestions have been made, falling into two generic classifications: (a) physical processes and (b) chemical reactions. From the standpoint of physical processes, all rely on the fact that microbes and extracellular enzymes have restricted access to physically protected proteins/peptides for hydrolysis (Nguyen and
Harvey, 2001), including the trapping of the proteins/peptides within mineral matter (Baldock et al., 2004; Hedges and Keil, 1995) and encapsulation into refractory natural organic matter (Knicker and Hatcher, 1997). There are several alternative suggestions for the preservation involving chemical coupling with various compounds present within the sediment (Hedges and Keil, 1995; Rillig et al., 2007). These include Maillard reactions between carbohydrates and proteins where the stable products melanoidins are preserved (Hedges et al., 2000; Maillard, 1916), the degradation and reconstitution of polymers by condensation reactions (Knicker, 2004), and the production of Michael adducts and Schiff bases that can occur between aromatic structures and proteins/peptides (Hsu and Hatcher, 2005).

Armed with some new analytical methods for characterising natural organic matter preservation, I examine two organic-rich sediments, where organic N is enriched (relative to marine sediments) and has been shown to be sequestered (Filley et al., 2001; Filley et al., 2002; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001). I employ several advanced NMR and mass spectrometry (MS) techniques to examine the average organic composition of the sediments. High resolution solution-state NMR is used to obtain two-dimensional data for solid sediment sample extracts, followed by a new MS technique capable of molecularly assessing the presence of the various organic N forms in these extracts - Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). My goal is to evaluate the N-containing organic matter with the most modern tools at my disposal so that I can perhaps discern the nature of the sequestered organic N and, consequently, the mechanism for sequestration.
4.2 Materials and Methods

4.2.1 Sample preparation

4.2.1.1 Sediment sample preparation. The organic-rich sediments are from two lakes; the first is Mangrove Lake, Bermuda (MLB), formed from algal-derived sedimentary organic matter deposited anaerobically (Hatcher et al., 1982; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Orem et al., 1986; Zang and Hatcher, 2002). The second is Mud Lake, Florida (MLF), a freshwater lake of 180 ha size accumulating high levels of algal-derived organic matter. Details of the collection of the sediments used for this study from MLB and MLF are described in previous studies (Bates et al., 1995; Filley et al., 2001; Hatcher et al., 1983). Prior to analysis with FT-ICR-MS, each sediment sample was extracted with pyridine (Acros) (50 mg/1 mL) for 3 days at room temperature using an orbital shaker (VWR model 57018-754). I chose as an extract solvent, pyridine, mainly because previous studies have shown it to be an extremely good solvent for coal (Wu et al., 2003). Admittedly, sediments examined here are far from being coal, but I recognise that pyridine’s solvating power would assure me that a significant fraction of the sedimentary organic matter would be rendered soluble.

The extract was then filtered by a 0.2 μm Millipore™ filter effectively stopping the extraction procedure. For gas chromatography mass spectrometry (GC-MS) analysis, chloroform (Fluka) was used as an extracting solvent, which seemed to remove a higher proportion of alkyl amides from the sediment than pyridine. Chloroform extraction involved the same procedure as that for the pyridine extractions as described above. For analysis by liquid-state NMR, the sediments were extracted in the same manner as for MS using larger volumes of sample and pyridine due to the inherent insensitivity of the
NMR technique. After filtration, the pyridine was removed from the extracts by evaporation using nitrogen gas at room temperature. The remaining solid extracts were redissolved in 600 µL of dimethyl sulfoxide-$d_6$ (DMSO-$d_6$) (Cambridge Isotope Laboratories Inc.). For analysis by solid-state NMR approximately 80 mg of each whole sediment sample (dry weight) was packed into a 4 mm rotor (Kel-F™, Bruker Biospin).

4.2.1.2 Sunflower oil treatment. Approximately 1 g of sunflower oil (Richfood) was added to 2 mL of 3M aqueous solution of ammonium carbonate (Sigma-Aldrich) and the mixture was placed in an oven maintained at 30 °C. The pH of the two phase system was approximately 8 and stable throughout the four weeks of this reaction.

4.2.2 Solid-state Cross Polarisation Magic Angle Spinning (CPMAS) $^{13}$C NMR

A Bruker Biospin 400 MHz NMR spectrometer equipped with a gradient capable coil and solid-state multinuclear probe ($^1$H, $^{13}$C, $^{15}$N) was used for all NMR analyses in this study. The spectrometer is located at the College of Sciences major instrument cluster (COSMIC) at Old Dominion University. A ramp cross polarization pulse programme was used for analysis, employing a magic angle of 54.7° and spinning at 14 kHz. A 4 µs 90° pulse was applied at the proton frequency and the magnetisation was transferred to $^{13}$C during a ramped contact time of 1 ms for MLB and 2 ms for MLF (the contact times used were determined to be optimum for each sediment and the ramping involved changing the proton power level from 50 to 100 %). The recycle delay time used for all samples was 1 s followed by a data acquisition time domain of 512 data points. Data processing used the Bruker Topspin v. 1.3 software™ to zero-fill to 4096 data points and exponentially filter (1 Hz line broadening) the Free Induction Decay (FID).
4.2.3 Liquid-state Heteronuclear Multiple Bond Correlation (HMBC) $^1H$-$^{13}C$ NMR

The pulse programme used heteronuclear zero and double quantum coherence with a gradient pulse time of 1000 μs and no decoupling. Long range couplings were selected based on a delay of 1 s. The dwell time of 93.8 μs was used between the acquisition of two successive data points. A total of 4096 time domain data points were acquired for the $^{13}C$ channel and 128 for the $^1H$ channel using digital quadrature detection. Using Bruker Topspin software zero-filling both domains to 1024 data points was applied to the acquired data and exponential line broadenings of 0.3 Hz for $^{13}C$ and 1 Hz for $^1H$.

4.2.4 FT-ICR-MS

4.2.4.1 Sediment analysis. The pyridine extracts of MLB sediments were diluted by 1000-fold in methanol (addition of 1 μL to 1000 μL methanol, Sigma-Aldrich), and extracts from MLF were diluted by only 100-fold prior to injection into the electrospray (ESI) source to aid ionisation. The pH was adjusted for each sample to either pH 8 for negative ionisation mode by addition of small quantities of ammonium hydroxide (Sigma-Aldrich) or to pH 6 for positive ionisation mode by addition of formic acid (Fluka). Polyethylene glycol was used to externally calibrate the instrument to an accuracy of 1 ppm prior to analysis of the samples. All analyses were completed on a Bruker Daltonics 12 Tesla Apex Qe FT-ICR-MS at the COSMIC facility. Samples were infused using a syringe pump at a rate of 120 μL/h into the Apollo II ESI ion source. Voltages were adjusted for each sample and for the ionisation mode being utilised (positive/negative). Ions were accumulated in the hexapole for 1 sec before being transferred to the ICR cell. Exactly 300 co-added transients were acquired in broadband
mode. A 4 MWord time domain was used for the FID. The Bruker Daltonics data analysis software v. 1.3™ was used to zero-fill once and Sine-Bell apodise the FID prior to magnitude-calculated Fourier transformation. Internal calibration was performed on each of the samples using naturally present fatty acids (Sleighter et al., 2008) for negative ionisation mode and homologous series within all the spectra were used to calibrate the positive ionisation mode. Elemental formulae were calculated using the Molecular Formula Calculator (v. 1.0, National High Magnetic Field Laboratory, Tallahassee, FL, 1998) with an accuracy of less than 1.0 ppm, and final assignment choices were aided using a series of conservative rules inserted into a programme developed using Matlab (v. 7.4.0, The Mathworks Inc. Natick, MA ) (Wozniak et al., 2008).

4.2.4.2 Sunflower oil samples. For analysis of the sunflower oil, in negative ionisation mode before and after treatment, samples were diluted to a concentration ratio of 1:50000 (v/v) with methanol, and the pH was adjusted as described above for the pyridine extracts. An ion accumulation time of 1 s was used and 20 co-added transient FID’s were acquired and analysed as described above.

4.2.4.3 Alkyl amide standard. For analysis of behenamide (Sigma-Aldrich) by FT-ICR-MS, samples were prepared separately for each ionisation mode. For negative ionisation mode, a small quantity of standard (approx. 5 mg) was diluted in pyridine (1 mL), immediately prior to analysis, the standard was diluted further by methanol (1:1000 pyridine:methanol v/v) to achieve an approximate final concentration of 0.005 mg/mL. For positive ionisation mode, 0.015 mg/mL stock solution in pyridine was prepared, immediately prior to analysis, this was further diluted 1:10 using methanol (v/v), with a final concentration of 0.0015 mg/mL. As in previous FT-ICR-MS analyses the pH was adjusted using ammonium hydroxide for negative and formic acid for positive ionisation.
analysis. For analysis in both ionisation modes an ion accumulation time of 1 sec was used and 20 co-added transient FID’s were acquired and analysed as described above.

4.2.4.4 Fatty acid and alkyl amide mixed standard. The fatty acid and alkyl amide standards were prepared separately and then mixed immediately prior to analysis with negative ionisation mode FT-ICR-MS. For the alkyl amide stock solution the following alkyl amides were used: tetradecanamide (Alfa Aesar), hexadecanamide (VWR), octadecanamide (Sigma-Aldrich), oleamide (Chemservice), behenamide, and 13-docosanamide (Cayman Chemicals). The approximate concentration of each alkyl amide in the stock solution was 0.015 mg/mL in pyridine. For the fatty acid stock solution the fatty acids used included pentadecanoic acid (Sigma-Aldrich), hexadecanoic acid (Sigma-Aldrich), nonadecanoic acid (Sigma-Aldrich), eicosanoic acid (Sigma-Aldrich), docosanoic acid (Sigma-Aldrich) and tetracosanoic acid (Sigma-Aldrich), prepared in a concentration similar to that of alkyl amides (~ 0.015 mg/mL). Table 5 displays the exact concentrations for each individual compound in these two stock solutions. The stock solutions were stored at 4 °C until analysis. For the mixed standard, the fatty acid and alkyl amide stock solutions were further diluted in a proportion of 1:100:1000 (v/v), respectively with methanol. Ammonium hydroxide was used to adjust the pH to approximately 8. An ion accumulation time of 1 sec and 20 co-added transient FID’s were acquired and analysed as described above.
Table 5

Exact concentrations of the two stock solutions: fatty acid and alkyl amide used for FT-ICR-MS analysis.

<table>
<thead>
<tr>
<th>Alkyl amide</th>
<th>Concentration (mg/mL)</th>
<th>Fatty acid</th>
<th>Concentration (mg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetradecanamide</td>
<td>0.0155</td>
<td>Pentadecanoic acid</td>
<td>0.0202</td>
</tr>
<tr>
<td>Hexadecanamide</td>
<td>0.0180</td>
<td>Hexadecanoic acid</td>
<td>0.0151</td>
</tr>
<tr>
<td>Oleamide</td>
<td>0.0170</td>
<td>Nonadecanoic acid</td>
<td>0.0160</td>
</tr>
<tr>
<td>Octadecanamide</td>
<td>0.0165</td>
<td>Eicosanoic acid</td>
<td>0.0151</td>
</tr>
<tr>
<td>13-Docosenamide</td>
<td>0.0158</td>
<td>Docosanoic acid</td>
<td>0.0196</td>
</tr>
<tr>
<td>Behenamide</td>
<td>0.0163</td>
<td>Tetracosanoic acid</td>
<td>0.0172</td>
</tr>
</tbody>
</table>

4.2.5 GC-MS analysis

The residue of the chloroform extracts of MLB sediments was obtained by removing the chloroform by evaporation with N₂ and redisolution in 2 mL of dichloromethane:methanol (3:2 v/v) which was used successfully to observe alkyl amides produced under hydrothermal conditions (Rushdi and Simoneit, 2004). A solvent blank was analysed prior to and after each sample to ensure that no contaminants were present in the GC column. Standard alkyl amides of tetradecanamide, hexadecanamide, octadecanamide, oleamide, behenamide, and 13-docosenamide were used as a basis with which to identify alkyl amides present within the sample. The modified methodology used for all samples followed that of Rushdi and Simoneit (2004), on a capillary GC-MS instrument, a Hewlett Packard 6890 series GC system which is connected to a Pegasus II time-of-flight mass spectrometer (LECO Corp.). The GC column used was a 15 m x 0.25 mm (i.d.) fused silica capillary column with a film thickness of 0.25 m (5 % phenylpolysiloxane and 95 % methylpolysiloxane, Restek Rtx 5). 1 μL samples were injected using a split/splitless injector which was selected to operate in splitless mode at 280 °C for these samples.

The column, with helium carrier gas running at a constant flow of 1.0 mL/min,
was temperature programmed using an initial hold time of 4 minutes at 40 °C followed by ramping to 300 °C at a rate of 20 °C/min, and with an isothermal time of 5 minutes at final temperature. For the mass spectrometer, the ionisation mode was electron ionisation at 70 eV with an ion source temperature of 200 °C and a transfer line temperature of 280 °C. Collection of the mass spectra was at a rate of 20 spectra per second after a solvent delay time of 120 sec. Acquisition of masses between m/z 35 and 500 was used. All analyses used the LECO Pegasus II software (v. 3.22).

4.3 Results and Discussion

4.3.1 Solid-state \textsuperscript{13}C NMR

Fig. 15 shows the solid-state CPMAS \textsuperscript{13}C NMR spectra obtained from analysing whole sediment samples from MLB and MLF. In the spectra from shallow depths (Fig. 15a and 15c) alkoxyl (66 - 92 ppm and 92 - 115 ppm) from carbohydrates is the dominant functional group, which is a logical observation when considering the source of the organic matter is primarily algae in both lakes. Resonances at 10 - 48 ppm are assigned to aliphatic functionalities which are the second most dominant signals, presumably deriving from long alkyl chains present. The peak between 167 and 188 ppm observed in both Figures cannot be unambiguously assigned to one functionality since both carboxyl and amide are possible. It is perhaps surprising that these peaks are so intense, when compared to the \textsuperscript{13}C NMR spectrum of the peat of MLB (Knicker and Hatcher, 2001). Aromatic functionality (124 - 142 ppm) is the smallest peak in these spectra, indicating that the sediments do not have significant contributions from terrestrial organic matter that is usually laden with aromatic carbon. Comparing the \textsuperscript{13}C NMR
spectrum of MLB sediment collected at depth (2.8 m) (Fig. 15b) and MLF sediment from 3 m (Fig. 15d), it is clear that aliphatics proportionally increase with depth, presumably because they are selectively enhanced as more labile alkoxy groups of carbohydrates are removed from the organic matter (Hatcher et al., 1983). These observations are consistent with the concept of selective preservation, with more labile groups being degraded initially leaving only the recalcitrant groups to be preserved at lower depths. The aromatic peaks in Fig. 15a, 15b, 15c and 15d do not alter significantly, seemingly because the sediments have only undergone early diagenesis and no additional inputs of terrestrial material are observed. Another dominant functional group observed in Fig. 15a, 15b, 15c and 15d is the amide/carboxyl resonance (~175 ppm). Since both of these functionalities are associated with compounds thought to be fairly labile (proteins, peptides, fatty acids, etc.), this peak would be expected to decrease, mirroring alkoxy behaviour, but their persistence indicates some preservation of these groups. The spectral overlap of amide and carboxyl groups in solid-state $^{13}$C NMR spectra at 165 – 190 ppm makes it difficult to determine which functional group is most preserved or degraded. It should be noted that this discussion pertains only to those environments that contain nonprotonated C atoms.
Fig. 15. NMR $^{13}$C CPMAS solid-state spectra of whole sediment samples from (a) MLB of depth 0.52 m, (b) MLB of depth 2.8 m, (c) MLF of 0 m depth, and (d) MLF of depth 3.0 m.

The solid-state $^{15}$N NMR spectra of these sediments (Fig. 16) reveal that amides are the only major functionality observed at all depths. This is the same conclusion made by past studies of these sediments (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Knicker et al., 2002; Zang and Hatcher, 2002).
Fig. 16. NMR $^{15}$N CPMAS solid-state spectra of whole sediment samples from (a) MLB of depth 0.52 m, (b) MLB of depth 2.8 m, (c) MLF of depth 0 m. In each spectrum the dominant N functionality is associated with amide functional groups.

4.3.2 HMBC solution-state NMR of MLB and MLF sediment extracts

Fig. 17 shows the 2-dimensional HMBC solution-state NMR spectrum of the pyridine extract of MLB sediment at depth 2.8 m (Fig. 17a) and that of MLF sediment of depth 3.0 m (Fig. 17b) where higher spectral resolution is realised compared to solid-state NMR data. Analyses of samples from other depths of MLB and MLF reveal that their spectra are very similar to Fig. 17a and 17b. There are three main spectral regions, aliphatic, aromatic and amide, however, since the focus of this paper is on N functionality, I examined in detail only the amide region, even though the high resolution of the spectrum in other chemical shift regions suggests that a wealth of information can
be gleaned. The only amide peaks which would be expected in this region are from proteins/peptides, and, indeed, I observe peaks at 4.01/171, 2.7/171 and 2.58/171 ppm that are readily assigned to proteins/peptides. However, I observe additional cross peaks at 2.2/172 and 2.1/175 ppm that cannot be assigned to proteins/peptides. My search for assignments uncovers these cross peaks as alkyl amides, correlating to an interaction 3 bond lengths distance, structural entities not recognised in previous studies. These assignments are made based on predictions and analyses of HMBC spectra of model alkyl amides, and peptides (Using 2D NMR predictor, ACD labs v. 9.0), and my knowledge of these samples. The HMBC spectrum (not shown) of a mixture of 4 alkyl amides confirms the above assignment. I also verified that the subject cross peaks cannot be assigned to other amide functionalities such as amino sugars by using the 2D NMR predictor of some example sugars. For example, N-acetylglucosamine shows a signal at 1.93 ppm, 177 ppm, for protons and carbons α to the carbonyl carbon, respectively, a region that is remote to the one assigned to the alkyl amides. Alkyl amides are therefore, observed in HMBC spectra for both lakes suggesting that they could have widespread occurrence in thermally immature organic-rich anoxic sediments.
Fig. 17. HMBC NMR spectrum of liquid-state analysis of (a) MLB sediment pyridine extract of 2.8 m and (b) MLF sediment pyridine extract of 3.0 m. The amide region is expanded and the peaks are identified.

4.3.3 FT-ICR-MS analysis

4.3.3.1 MLB and MLF sediment extracts. The assignment of a new class of compounds not previously considered in thermally immature sediments, warrants further investigation with a technique capable of molecular characterisation. FT-ICR-MS has been successfully used to identify individual organic compounds at the molecular level for several different types of natural organic matter, - including dissolved organic matter (Kujawinski et al., 2004; Sleighter and Hatcher, 2007; Stenson et al., 2003), humic extracts of soil organic matter (Hockaday et al., 2006; Hughey et al., 2008; Kujawinski et al., 2002), and pore water dissolved organic matter (Koch et al., 2005; Tremblay et al., 2007). Recently, there have been some dissolved organic matter and pore water studies focused on N-containing compounds by FT-ICR-MS (Reemtsma et al., 2008; Schmidt et al., 2009). Reemtsma et al. (2008) examined hydrophobic and hydrophilic fractions of marine waters, identifying many N-containing compounds. The study by Schmidt et al.
(2009), examines N compounds in pore water obtained from a variety of sources.

Fig. 18. FT-ICR-MS spectra of MLB sediment from 2.8 m depth. Pyridine extract analysed by (a) negative ionisation mode and (b) positive ionisation mode. Insets in both spectra show expanded spectra of peaks with formulae similar to that of two CHNO compounds chosen arbitrarily. CHNO and CHN₂O₂ as well as the high magnitude peaks are identified. In the table inset, various formula types are assigned symbols, (n, m are the numbers of carbon and hydrogen in the formulae respectively).

Fig. 18 shows the mass spectra of the pyridine extract of sediment of depth 2.8 m from MLB in both negative (Fig. 18a) and positive (Fig. 18b) ionisation modes. Fig. 19 displays the mass spectra obtained by FT-ICR-MS analysis of the pyridine extract of
MLF, of depth 3 m in negative (Fig. 19a) and positive (Fig. 19b) ionisation modes. Although the individual spectra vary somewhat among the samples (in terms of which peaks are dominating), mainly due to the relative dominance of more easily ionisable compound peaks (such as fatty acid type formulae), the spectra shown in Fig. 18 and 19 are generally representative of all other samples from MLB and MLF, especially for the occurrence of compounds containing N. Many of the N-containing compounds are assignable to proteinaceous materials (Fig. 18b), dictated by the NMR studies discussed above that indicate their presence. Close inspection of the peaks that contain the elements CHNO identifies, in addition to proteinaceous materials, molecular formulae that I believe, based on NMR studies above and discussions below, belong to a homologous series of amides having the specific formulae $C_xH_yNO$ and $C_xH_yN_2O_2$ (or as sodium adducts in Fig. 18b and 19b: $C_xH_yNONa$ and $C_xH_yN_2O_2Na$) with the latter being quite prevalent in Fig. 19a, ($x$ and $y$ are the number of carbons and hydrogens respectively in the formulae). The solids $^{15}$N NMR data indicate that all N is amide. The HMBC data indicate that proteins/peptides and alkyl amides are the only two possible component amides. The foundation of my belief is that the above CHNO formulae are more consistent with alkyl amides than proteins.

Two component peaks (Elemental formulae: $C_{22}H_{43}NO$ and $C_{22}H_{45}NO$) from these series are shown in expanded regions in the insets to Fig. 18a and 18b (m/z of 336.327066 and 338.342887 for negative and m/z of 360.323608 and 362.339397 for positive ionisation). Additionally, since the magnitudes of the peaks at m/z of 336.327066 (Fig. 18a) and 360.323608 (Fig. 18b) are so high, their $^{13}$C isotopologues can also be observed (m/z of 337.330449 for negative and 361.326975 for positive), indicating their single charge (difference is equal to mass of one neutron, 1.003). While
there are many differences between these spectra, there are several peaks which, although their magnitudes may vary, are observed in spectra from both lakes. Fig. 19, that of MLF, also identifies several alkyl amides and diamides and is similarly analysed.

Fig. 19. FT-ICR-MS of the pyridine extract from MLF sediment, depth 3 m. They are analysed by (a) negative ionisation mode and (b) positive ionisation mode. Insets show expanded spectra of peaks with formulae of CHNO (not all CHNO containing peaks are shown) and other formula types. In the table legend the respective symbols for various formula types are listed, and subscripts n and m refer to numbers of carbon and hydrogen in the formulae respectively.

I considered that addition of ammonium hydroxide to render the sample basic during the preparation of the extracts for negative ionisation could potentially lead to the artificial formation of amides in the ESI process. However, the presence of many of the same formulae in positive ionisation analysis, where no ammonium hydroxide is added,
confirms that these compounds are naturally present within the sediment.

4.3.3.2 Confirmation of ionisation efficiency of alkyl amide. To confirm that alkyl amides are readily ionised by the ESI process and detected, I analysed alkyl amide standards (one of these, behenamide is shown as an example) in both negative (Fig. 20a) and positive ionisation modes (Fig. 20b). The elemental formula corresponding to behenamide ($C_{22}H_{43}NO$) is successfully identified in both ionisation modes. These data were only externally calibrated slightly increasing the error of formulae assignments ($< 2.5$ ppm), however, since these are standards I am confident in my assignments. Additionally, isotopologues of behenamide ($^{13}C_{12}C_{21}H_{43}NO$ and $^{13}C_2^{12}C_{20}H_{43}NO$) are identified which further confirm my successful formula assignments. Therefore, I feel that assignment of alkyl amides to MLB and MLF identified formulae is warranted, however, their exact structure (branched or normal) is unknown by this analysis. In addition, mass spectrometry/mass spectrometry (MS/MS) was attempted on the MLB/MLF sample extracts and alkyl amide standards to obtain structural information directly from these mass spectra; however, I was unable to fragment these compounds sufficiently to identify their structure. Therefore, I decided to investigate alternative techniques that are capable of confirming their structure.
4.3.4 Confirmation of alkyl amide structure by GC-MS analysis

GC-MS analysis of the MLB extracts (Fig. 21) shows the presence of ten different alkyl amides, observed in the selected ion chromatogram (m/z 59 characteristic of ion fragmentation for alkyl amides, C₅H₅NO); they correspond to saturated and unsaturated n-alkyl amides with C₁₄ – C₂₂. Alkyl amide standards were analysed and used as comparisons for peak identifications in chromatograms of MLB samples. This GC-MS analysis verifies that compounds from homologous series of CₓHᵧNO in FT-ICR-MS are long chain alkyl amides. The relative distributions of individual homologues from GC-MS do not match the distributions gleaned from FT-ICR-MS. However, lack of correspondence is partly related to differences in ionisation efficiencies and the fact that
the GC-MS technique was not fully optimised, especially with regard to observing the higher molecular weight homologues. In FT-ICR-MS data from both lakes these alkyl amides are observed to a sediment depth of 3 m, in MLB a horizon that approximately corresponds to an age of 2000 yrs BP (Knicker and Hatcher, 2001; Orem et al., 1986) and in MLF, corresponding to an age greater than 5000 yrs BP (Filley et al., 2001). While I understand these ages are in reference to the carbon component of these sediments, I believe, due to the prevalence of alkyl amides throughout my depth profile that they are produced rapidly and remain in situ throughout burial. It is remotely possible that these molecules could undergo diffusion throughout the sediment, however, based on past studies regarding the examination of the pore waters from MLB there are many trends that are depth dependent suggesting that there is little diffusion through the sediment. In addition the solubility of these compounds in water is low; therefore, I expect that after formation of these compounds in the pore water/sediment that they will precipitate out of the pore water into the sediment.
Fig. 21. GC-MS selected ion chromatogram (m/z 59) of MLB extract from sediment of depth 2.8 m. Label numbers refer to the carbon number (number on left) and the number of unsaturation points or double bond equivalent (number on right).

4.3.5 Categorisation of MLB and MLF sediment extracts molecular formulae assignments

Examination of the large numbers of peaks and assigned formulae reveals the existence of compounds that share some common elemental compositions, thus allowing me to subdivide the formulae into classes according to their heteroatom contents. Fig. 22 shows the classification of the different types of formulae observed in the mass spectra from both lakes. The total number of peaks assigned to each formula type was calculated and then converted to a percentage of all formulae identified within the mass spectrum. The formula type CHO is included in this Figure for comparison purposes with all the N-containing formulae types. Fig 4.8 shows that CHO is the dominant formula type at all
sediment depths. I believe that most of these peaks are fatty acids which usually display high ionisation efficiencies and tend to dominate such spectra when present. Apart from CHO, the next greatest number of formulae are those of CHNO (when the other formula types are included in this Figure that constitute the total percent in this sample, CHNO is still the second greatest contributor after CHO - data not shown).

![Graph showing percentage of formula types from MLB at different depths.](image-url)

Fig. 22. The percentage of formula types from MLB at (a) 0.52 m and (b) 2.8 m and MLF (c) 0 cm and (d) 3.0 m analysed using negative ionisation mode FT-ICR-MS.

4.3.6 Identification of CHNO formulae observed

Closer identification of CHNO containing formulae in MLB for a selected depth (2.8 m) and MLF (3 m) is displayed in Table 6 listing the formulae and possible structures (third column). The elemental formulae identified for more shallow depths varies from these two but, since there does not appear to be a pattern, these are not included here. Most formulae belong to homologous series of peaks. These series are
described in this Table and consequently identified (first column) as alkyl amides based on their elemental formulae. Primary amides with long chain alkyl groups (singly unsaturated or saturated) are the most probable structures, since aliphatic functionalities dominate the solid-state $^{13}$C NMR spectra (Fig. 15) and the GC-MS data confirms this for several of the formulae. The proposed alkyl amides are either completely saturated or have one degree of unsaturation in the alkyl chain (Table 6). An additional degree of unsaturation is observed only in MLB spectra. Both alkyl amides and alkyl diamides are observed in MLB, whereas in MLF only alkyl amides are found. Possible reasons for this could be due to the inherent differences in the lakes, such as the age of the organic matter, the sources, the extent of degradation, the amount of ammonia available, or the location of the lake. The insinuation as to the most important of these reasons is speculative at this point, but is being further studied in my group.
Table 6

Series of alkyl amides which are observed in mass spectra of the pyridine extracts from the sediments of MLB and MLF at depths of 2.8 m and 3 m respectively using FT-ICR-MS in negative ionisation mode.

<table>
<thead>
<tr>
<th>General formula for each series</th>
<th>Peaks observed in the mass spectrum belonging to series</th>
<th>Suggested structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{x+y}H_{2(x+y)}NO$</td>
<td>MLF: $x+y = 16-22$ even only, MLB: $x+y = 13-23$</td>
<td>![Structure 1]</td>
</tr>
<tr>
<td>$C_xH_{2x}NO$</td>
<td>MLF: $x = 15-21$ even only, MLB: $x = 13-30$</td>
<td>![Structure 2]</td>
</tr>
<tr>
<td>$C_xH_{2x}N_2O_2$</td>
<td>MLB: $x = 18-40$</td>
<td>![Structure 3]</td>
</tr>
<tr>
<td>$C_{x+y}H_{2(x+y)-2}N_2O_2$</td>
<td>MLB: $x+y = 21, 32-35$</td>
<td>![Structure 4]</td>
</tr>
<tr>
<td>$C_{x+y}H_{2(x+y)+2}N_2O_2$</td>
<td>MLB: $x+y+z = 13, 15, 19, 28, 32$</td>
<td>![Structure 5]</td>
</tr>
</tbody>
</table>

* Calculated using included atoms of the suggested structures.

$X, Y, Z = (\text{CH}_2)^{x,y,z}$

Identifying the individual formulae in a single mass spectrum can provide much information about the constituents of sediment extracts, however, the relative importance of these alkyl amides in a mass spectrum should also be considered if their role in N sequestration is important. It should be noted that the efficiency with which an ion charges in the ESI source and its behaviour in the ICR cell have large impacts on the
magnitude of the resulting peak in the mass spectrum (Wozniak et al., 2008). However, if instrument parameters are carefully maintained when analysing samples of similar matrices it is possible to make general statements regarding the magnitude of the peaks in each spectrum. Linking the magnitude of a particular peak and its concentration in the original sample is currently not possible, but the magnitude can indicate whether a peak is changing in concentration relative to the rest of the sample. This is a useful assessment to make when comparing similar samples such as the pyridine extracts in this study. In Fig. 23 I present the calculated contributions that alkyl amides make to the CHNO formula type for negative ionisation mode analysis (more sensitive than positive ionisation mode) for a depth profile in MLB sediment extracts. The data are weighted according to their magnitude, and calculated as a percentage of alkyl amides/diamides compared to the total for all CHNO containing formulae classified in Fig. 22, for each horizon throughout the depth profile of this MLB core (0.22 – 3 m). The magnitude weighted percentage generally increases (excluding 0.22 m depth) throughout the profile from 19.12 % (0.52 m) to 80.95 % (3 m). The main reason why the horizon corresponding to 0.22 m does not follow this pattern is due to the large magnitude of m/z 336.327002 assigned the formula C_{22}H_{43}NO, a singly unsaturated alkyl amide. When examining the broadband spectra of all these depths it is clear that this peak increases in magnitude relative to the rest of the mass spectrum with increasing depth. This Figure demonstrates that formulae assigned as alkyl amide/diamide increase in relative magnitude when compared to all CHNO containing compounds for sediment depths below the 0.22 m sample, confirming their relative importance for this formula type. The increasing dominance of alkyl amides over other more labile CHNO containing compounds such as proteins/peptides can possibly be indicative of the relative
recalcitrance of these alkyl amide types of molecules.

![Magnitude weighted percentage of alkyl (di)amide to CHNO compounds](image)

Fig. 23. Magnitude weighted percent of alkyl amides in FT-ICR-MS negative ionisation data of depths 0.22 m to 3.0 m depth of pyridine sediment extracts from MLB.

The relative contribution of alkyl amides to the total of all peaks observed in the spectra cannot be judged to reflect an accurate quantitative measure of importance. However, it is important to mention that the abundance observed for these alkyl amides/diamides is relatively high considering the overall dominance of fatty acids which ionise well in ESI. Fig. 24 displays the ESI-FT-ICR-MS spectrum of fatty acids and alkyl amides when analysed in negative ionisation mode. Despite the fact that the fatty acid concentration is a factor of one hundred lower than that of the alkyl amides, several of the fatty acid peaks (C_{22}H_{43}O_2^- and C_{24}H_{47}O_2^-) are larger than those of some alkyl amides (C_{14}H_{28}NO^- and C_{18}H_{34}NO^-). Fig. 24 demonstrates that fatty acids can dominate mass spectra despite having several orders of magnitude lower concentration, explaining
their predominance in several of the negative ionisation spectra for the MLB sediment extracts. Additionally, the magnitudes of the alkyl amide peaks in the standard (Fig. 24) differ considerably, seemingly depending only the length of alkyl chain; most notably for $C_{22}H_{42}NO^-$ and $C_{22}H_{44}NO^-$. The reason why small changes in structure make such a large difference in ionisation efficiency such as in this case is still unknown for this ion source.

Fig. 24. FT-ICR-MS spectrum of fatty acid and alkyl amide mixed standard analysed using negative ionisation mode. Mass spectrum is expanded off scale vertically in order to show the lower magnitude peaks.

4.3.7 Possible route of formation of alkyl amides

This study identifies, for the first time in organic-rich and thermally immature sediments, a homologous series of alkyl amides and diamides. Since alkyl amides are observed throughout both lakes and their respective cores, they probably are incorporated rapidly and preserved for many thousands of years (carbon dating indicates that both lakes were formed about 10,000 yrs BP). Although speculative, it seems reasonable to suggest that the several chains could form a micelle which would serve to protect and preserve the N functionality in these highly organic sediments. They could also be
incorporated within the more recalcitrant organic matter such as algaenan which was suggested for peptide preservation (Knicker, 2004; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001). Another explanation for their survival is that the microorganisms in the anoxic sediments do not have the enzymes to degrade alkyl amides.

To my knowledge alkyl amides have not been considered as significant components of recently deposited and thermally unaltered organic-rich sediments (Derenne et al., 1993; Rushdi and Simoneit, 2004). Derenne et al. (1993), suggested that alkyl amides could be present in refractory isolates of certain algae, known as algaenan, based on indirect evidence from pyrolytic degradation of organic matter. Additionally alkyl amides have been observed in sediments impacted by hydrothermal conditions (Rushdi and Simoneit, 2004) and biomass burning (Simoneit et al., 2003). The implication is that high temperature alteration is involved in the formation of alkyl amides. I have no reason to believe that the alkyl amides in MLB are sources from biomass burning. Several meteorological studies have identified that trace elements measured in Bermuda could be sourced from North America, Europe and Africa (Arimoto et al., 1995; Moody et al., 1995). In these locations biomass burnings occur frequently. However, biomass burning also produces condensed aromatic compounds that are often observed in soils, sediments, aerosols, and waters (Dittmar and Koch, 2006; Hammes et al., 2007; Hockaday et al., 2006; Hockaday et al., 2007; Wozniak et al., 2008) and I was unable to find evidence of black carbon, a common biomarker for biomass burnings in my FT-ICR-MS spectra (Fig. 18). In the case of MLF, biomass burnings could potentially be a more substantial source to the sediment organic matter, however, black carbon was not observed in its mass spectra (Fig. 19) and so is deemed unlikely to be a major source to the sediment organic matter. It is unlikely that the alkyl amides
observed in both MLB and MLF are thermogenic.

The biological production of primary amides through a lipase-catalysed reaction of esters with ammonia is well known (de Zoete et al., 1996). However, they could also be produced abiologically *in situ* in the sediments. It is commonly known that algae naturally produce long chain alkyl esters at high concentrations which become incorporated in sedimentary organic matter (Burdige, 2007). The reaction of esters with N-containing nucleophiles is therefore probable since the carbonyl carbon is an excellent electrophile (Amrani et al., 2007). A well established reaction in many industrial and biological processes, for the production of amides, is by amidation of esters (Basha et al., 1977; de Zoete et al., 1996; Zabicky, 1970). Since copious amounts of ammonia, an excellent nucleophile, are produced by the degradation of organic matter in anoxic sediments common to both lakes (approximately 4 mmol/L in pore waters of MLB (Hatcher et al., 1982; Orem et al., 1986), it is very likely that the abundant esters undergo amidation readily (Amrani et al., 2007).
Fig. 25. FT-ICR-MS spectra analysed using negative ionisation mode of sunflower oil (a) is before reaction, and (b) after reaction with ammonium carbonate. The insets show the expanded region containing one of the alkyl amides that was formed during the reaction.

To test the feasibility of natural amidation of long chain alkyl esters, sunflower oil, which is composed almost entirely of glyceryl esters, was reacted with ammonium carbonate close to environmental conditions (temperature and concentration increased to decrease reaction time). Fig. 25 shows the mass spectra of the oils before and after reaction, analysed using negative ionisation mode with FT-ICR-MS. An expanded region is shown in the insets to Fig. 25a and b representing what is observed at other alkyl amide masses throughout the mass spectrum. When comparing Fig. 25a (before reaction) and Fig. 25b (after reaction) it is immediately obvious that a new peak in Fig. 25b at m/z 336.326943 now dominates the mass spectrum. This new peak can be
assigned an elemental formula of \(C_{22}H_{43}NO\). This formula is exactly that which was found in the two lake sediment mass spectra, both positive and negative ionisation modes. This new peak appears only after the reaction and so the only possible source for this compound is the reaction of the esters present in the oil with ammonium carbonate. This experiment lends credence to my suggestion that the alkyl amides observed in the sediments studied can be abiotically produced by this amidation reaction.

4.4 Conclusions

Solid-state \(^{15}\)N NMR spectra of thermally immature natural organic matter buried in sediments examined in numerous other studies (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001), indicates the ubiquitous presence of amide functionality. Previous reports have sought to explain the existence of amides in MLB as contributions from proteinaceous remains of living organic matter that have somehow been preserved (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; Nguyen and Harvey, 2001). My new observations suggest that abiotic amidation is also a likely explanation for the existence of amide functionality in many sediments. Although N functionality in these two lakes is comprised of both proteins/peptides and alkyl amides, it is possible that alkyl amides could dominate in some other sedimentary systems where the proteinaceous remains are subjected to more extended degradation. The occurrence of ammonia in sediments co-existing with natural organic matter from plants that contain substantial amounts of alkyl triglycerides (algae) is common. Therefore, this amidation reaction should be an important pathway for N-immobilisation and possible sequestration. Clearly, a reassessment of previous hypotheses regarding preservation of N-containing
compounds is required before the impact of sequestration in sediments on global cycles can be fully understood.
CHAPTER 5

NITROGEN GEOCHEMISTRY OF SEDIMENTS AND PORE WATERS OF MANGROVE LAKE, BERMUDA STUDIED BY FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY: ABIOTIC PRODUCTION OF ALKYL AMIDES AND DIAMIDES

5.1 Introduction

Nitrogen containing organic matter (NCOM) is continually cycled throughout marine and fresh water environments in the sediments, the water above it and the water trapped in the sediments (pore water, PW) (Baldock et al., 2004; Chapelle, 1995). In addition, sediments are major pools of refractory organic matter (OM) throughout the world (Hedges and Oades, 1997). The OM preserved in these environments remains sequestered for decades due to the relative recalcitrance of these compounds compared to others originally present in the deposition system, however, the reason for their recalcitrance is still poorly understood (Burdige, 2007; Hedges and Keil, 1995). There have been many studies investigating the molecular character of the NCOM present in sediment, marine/fresh water and PW environments individually, but relatively few studies investigate all three pools of the same system (Bellanger et al., 2004; Hatcher et al., 1982; Schmidt et al., 2009; Schuster et al., 2003). Lacustrine environments provide an ideal location to conduct a study of NCOM due to the high OM concentration (compared to open ocean) in the three OM pools and the relatively enclosed environment, enabling a more comprehensive study (Meyers and Ishiwatari, 1993). Understanding the
interactions of the OM of various pools in a lacustrine system can assist in explaining the multifaceted cycling of important nutrients within a lacustrine environment, e.g., nitrogen, whose chemistry is extremely complex and difficult to explain on a global scale (Berman and Bronk, 2003; Herbert, 1999; McCarthy et al., 1997; Zehr and Ward, 2002).

Mangrove Lake, Bermuda (MLB) presents an ideal environment for a study involving molecular characterisation of the N in a variety of OM pools. This well studied lacustrine environment has been investigated for decades detailing the chemistry of the dissolved organic matter (DOM) of the Lake water (LW), the PW DOM and the sediment OM (SOM) present (Boudreau et al., 1992; Hatcher, 1978; Hatcher et al., 1982; Knicker and Hatcher, 1997; Knicker and Hatcher, 2001; McKee and Hatcher, 2010; McKee and Hatcher, 2011; Orem et al., 1986; Spiker and Hatcher, 1984; Zang and Hatcher, 2002; Zhong et al., 2011). Past studies revealed that much of the OM is derived from autochthonous algae in the LW, and much of their original structure is preserved in the sediments anaerobically (Hatcher et al., 1982; Knicker and Hatcher, 2001). Consequently, molecular characterisation of the sediment reveals that it is composed mainly of aliphatic moieties, as shown by nuclear magnetic resonance (NMR) spectroscopy (Knicker and Hatcher, 1997; McKee and Hatcher, 2010). Further molecular studies reveal that the SOM and PW DOM are composed of a mixture of carbohydrates, peptides and aliphatic structures (Knicker and Hatcher, 2001; McKee and Hatcher, 2010; Orem et al., 1986). Most important for this investigation is a previous study of this location is the revelation of a new class of compounds not previously discussed in immature sediments, namely, alkyl amides and diamides that seem to be selectively preserved in the sediment (McKee and Hatcher, 2010). In addition, their proposed mechanism of production via an amidation reaction involves the abiotic reaction of
naturally present esters and ammonia probably in the aqueous phase. The present study is an extension of my previous work in which I seek to relate the occurrence of alkyl amides in the sediments to their formation in PW through abiotic processes.

In order to facilitate this type of study an analytical technique capable of analysing the different pools of OM needs to be employed. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is one such method that has been used extensively to study a variety of OM pools on the molecular level. These include: marine/fresh water DOM (Hertkorn et al., 2006; Koch et al., 2005; Kujawinski et al., 2009; Reemtsma et al., 2008; Sleighter and Hatcher, 2008; Sleighter et al., 2009), PW DOM (D'Andrilli et al., 2010; Schmidt et al., 2009; Tremblay et al., 2007), rainwater DOM (Altieri et al., 2009), soil extracts (Kramer et al., 2004; Ohno et al., 2010), sediment extracts (McKee and Hatcher, 2010; McKee and Hatcher, 2011; Zhong et al., 2011), water soluble atmospheric OM (Kroll et al., 2011; Mazzoleni et al., 2010; Wozniak et al., 2008), and meteorite OM (Schmitt-Kopplin et al., 2010a). The wide variety of studies this technique has been utilised for presents an ideal tool to examine the N contents of the LW DOM, PW DOM and SOM.

The overall goal of this study is to investigate the NCOM in MLB by examining in molecular detail the PW, LW and sediments, with the ultimate objective of assessing the importance of alkyl amide and diamides in this system. Initially, I will examine trends in the bulk N measurements to establish observable changes in the N geochemistry that are salient to the production of alkyl amides. The molecular investigation of the NCOM in this system will be facilitated by FT-ICR-MS analysis of the three OM pools, the LW, PW and sediment. It will be used to assess the relative importance of molecular N in this system, and determine the likely structural sources of the identified N. The last
stage of this N examination will be directed towards the alkyl amide and diamide
contents of the LW, PW and sediment, to further understand their formation via
amidation by monitoring their concentration in this system in addition to key reactants,
i.e., ammonia.

5.2 Materials and Methods

5.2.1 Collection of samples from MLB

All samples used in this study were collected during November 2009 from MLB.
The exact location of collection is shown in Fig. 26, reached by construction of a cataraft
on the edge of the Lake. This location was chosen for two main reasons: first, the depth
of the water was shallower than other locations in the Lake (< 2 m), facilitating the
collection of a continuous sediment core, and second, it is in the centre of the Lake where
OM concentration is highest (Hatcher, 1978). Cores were extruded from the sediment by
use of five 15 cm diameter by 5 ft long plexiglass core barrels, attached together by pre-
threaded stainless steel couplings for secure and air tight fittings. Attached to the deepest
core barrel was a stainless steel coupling with sharp teeth that was able to penetrate the
sediment with minimal perturbation of the sediment surface. A piston was positioned at
the top of the barrels to maintain a vacuum in the barrels at all times and minimise loss of
material from the bottom. The barrels were guided into the sediment by suspension of
the piston and barrels from an A Frame on top of the raft (Fig. 27). Due to the depth of
the overlying LW the top core barrel was filled only with LW which was discarded after
its removal from the Lake.
Fig. 26. A map showing the location where samples were collected (shown with an X) in MLB in November 2009. The inset map is that of Bermuda, annotated with a square denoting the exact location of the Lake.

Two sediment cores were collected from the Lake on two separate days (3 days apart), from approximately the same position in the Lake. One complete filled core barrel was extruded from the sediment on the first day of collection (Core 1); this Core contained approximately 125 cm of sediment from the topmost region of the deposit. On the second collection day, a longer core section with a total of 330 cm of sediment contained within was recovered in three segmented sections of the coring device (subcores 2a, b and c). Sediments were removed from the Lake by gradually lifting the barrel sections up using the supporting A Frame. Upon their complete removal from the sediment but while still suspended in the LW, core barrel sections were uncoupled and capped using rubber stoppers and sealing electrical tape.
In the case of Core 2, core barrel sections were separated carefully working from the top down, with immediate sealing of the sediment filled barrels with the stoppers to prevent exposure to air. At all times, core barrels were stored upright to ensure PW did
not redistribute through the core. After collection, sediment core barrels were transported back to the Bermuda Institute of Ocean Sciences (BIOS) laboratory and subsampled as quickly as possible.

Core 1 was subsampled approximately 4 hours after collection from the Lake and Core 2a was subsampled approximately 2 hours after collection. Due to time constraints Cores 2b and 2c required storage overnight upright with the core barrels carefully sealed before their subsampling the following morning. Subsampling of the sediment into different depths was accomplished by pushing sediments up through the top of barrel into a N$_2$ environment by use of a Glove Bag™ (Glas-Col). Wet sediments were collected and stored under N$_2$ in Teflon® bags sealed with Clip-N-Seal® closures (Welch Fluorocarbon). When subsampling the sediment from the core barrel, the outer 1 – 2 cm was discarded and all tools used to divide the depths were cleaned between each depth to ensure minimal contamination between subsamples. Wet sediments were sampled by depth increments, with 0 cm representing the sediment/water interface. For Core 1, the initial 0 – 10 cm of core was subsampled every 1 cm; from 10 – 125 cm the core was subsampled every 5 cm. For Core 2, 0 – 10 cm was subsampled every 1 cm (Core 2a), 10 – 93 cm every 5 cm (Core 2a), and 93 – 330 cm every 10 cm (Core 2b and 2c). The reason for these different horizons is because I predict that there will be more substantial changes in the OM of the sediment and PW in the shallower depths, so closer sampling increments were collected in order to enable their visualisation. All wet sediments were placed inside a N$_2$ filled Glove Bag, and stored in a refrigerator before their work-up either the same or the following day.

LW was collected from the side of the raft on the same day as Core 2. All effort was made to exclude visible algae from the sample. Water was collected in pre-acid
washed, 2 L polypropylene bottles and stored in the dark until arrival at the BIOS where they were transferred to a refrigerator. Select samples were filtered through a 0.22 μm Millipore™ filter, into thoroughly rinsed 2 L sample bottles (3 x rinse bleach, 3 x rinse MilliQ, 3 x rinse sample). All glassware used was precombusted, and contamination from human contact was minimised by use of nitrile gloves. All LW samples were immediately frozen and transported back to Old Dominion University (ODU) via the R/V HSBC Atlantic Explorer.

PW samples were removed from all sediments that contained enough to be extruded in Core 1, and from select depths in Core 2. At lower depth (deeper horizons) it was often not possible to remove any PW because of the inherent compact nature of the sediment. PW was separated from the sediment by centrifugation. Although a variety of methods have been employed to remove PW from sediments, centrifugation is the most efficient for the highly sapropelic nature of MLB sediment (Adams, 1994; Orem et al., 1986). At all times, samples were processed and stored in an inert environment (N₂), to minimise oxidative reactions occurring in the PW. Glove Bags were employed (Fig. 27) to remove the PW from the sediment, by using the following general procedure: 1. place all items necessary for separation into the Glove Bag including Teflon bags filled with sediment (4 at a time) (see Fig. 27), 2. purge Glove Bag 3 times with N₂, 3. fill centrifuge tubes (FEP with Tefzel ETFE closure, Nalgene®, 1 per sample) with sediment sample and cap securely, 4. remove centrifuge tubes from the glove bag and centrifuge for 12 – 24 minutes at 3000 – 3400 rpm, 5. remove centrifuge tubes and place back in Glove Bag, followed by purging, 6. remove PW from centrifuge tubes by pipetting with a precombusted glass pipette into a precombusted 20 mL amber glass jar with Teflon® lined cap, 7. remove all items from glove bag and store appropriately. Sediments
remaining in centrifuge tubes were transferred and stored in separate polyethylene bags that were pre MilliQ and acid washed. Centrifuge tubes were rinsed thoroughly with MilliQ water between samples to reduce contamination and dried before reuse. All PW samples were stored at 4 °C from the time of collection under N2. The pH was measured of PW samples when possible using colorpHast indicator strips (EM science), with an accuracy of ± 0.2, while still in the Glove Bag. All samples were labelled according to their core and depth from the surface.

5.2.2 Preparation of samples for analyses

It was necessary to prepare the collected samples for analysis by a variety of techniques. For the collected sediments, all the remaining sediments, after removal of sediment for PW extraction, were frozen at the BIOS. They were transported back in coolers packed with ice packs, and remained frozen during the air transportation back to ODU. Upon arrival at ODU, samples were immediately placed in the freezer and stored at -14 °C. All sediments were lyophilised for more permanent storage at 4 °C, until their analysis. Whole unextracted dry sediments were used for elemental analysis of N. For FT-ICR-MS analysis, dry sediments were pyridine (Fluka) extracted, a method that has been shown effective to remove a large, representative fraction of SOM (McKee and Hatcher, 2010; McKee and Hatcher, 2011; Zhong et al., 2011). The protocol described briefly is: approximately 50 mg of dry sediment is added to 1 mL of pyridine and incubated for 3 days on an orbital shaker (VWR model 57018-754). Samples are 0.6 μm filtered (Whatman) using a glass syringe and stainless steel needle and filter holder, a procedure that effectively stops the extraction procedure. Pyridine extracts are stored at 4 °C, until their analysis a few days later. Previous studies of pyridine extracts of MLB
sediments demonstrated that storage at 4 °C does not substantially alter the molecular OM contents (data not shown).

As mentioned above, LW was frozen at the BIOS and transported back to ODU before defrosting for its analysis. For total nitrogen (TN) analysis, 15 mL of filtered water was used for analysis without further workup. In order to prepare for FT-ICR-MS analysis, the salts present need to be removed to prevent competition with sample ions for charge in the electrospray (ESI) source (Cech and Enke, 2001). A variety of methods have been used in past studies to remove salts and concentrate DOM in aqueous samples including ultrafiltration (Schwede-Thomas et al., 2005; Simjouw et al., 2005), reverse osmosis coupled with electrodialysis (Gurtler et al., 2008; Koprivnjak et al., 2006; Vetter et al., 2007) and solid phase extraction (SPE) (Dittmar et al., 2008; Fontanals et al., 2005; Kim et al., 2003b; Louchouarn et al., 2000; Morales-Cid et al., 2009). SPE polymer cartridges have been shown to yield high amounts of DOM recovery for small sample volumes compared to more traditional techniques (D’Andrilli et al., 2010; Dittmar et al., 2008). Two SPE filters were tested in this study specifically, PPL and Plexa, with filters that are both composed of styrene divinyl benzene polymers (Varian). PPL was determined to recover (by TOC/TN analysis) the most DOM from the LW sample (data not shown) and so was used for all further extractions. The protocol used for PPL extraction followed that of previous studies, and briefly involves 1. acidification of the sample with formic acid (Fluka), 2. 1 cartridge volume rinse with methanol (MeOH, Sigma-Aldrich) (3 mL), 3. addition of sample to the cartridge, 4. rinse with MilliQ water (6 mL), 5. dry with N₂ for 30 minutes, and 6. extraction of collected DOM on the cartridge with MeOH (3 mL) (Dittmar et al., 2008; Fontanals et al., 2005). For the LW sample, 40 mL of sample was extracted onto the PPL disc and used for FT-ICR-MS
PW samples were prepared for a variety of analyses. Ammonia concentration was determined immediately upon my return to ODU for all PW samples collected. For each analysis 50 μL was used of each PW sample. The rest of the samples were frozen at -14 °C and stored under N\textsubscript{2} to prevent the degradation of DOM. For FT-ICR-MS analysis select PW samples (depths of 0 – 1, 4 – 5, 10 – 15, 20 – 25, 40 – 45, 65 – 70, 120 – 125 and 320 – 330 cm from both cores) were defrosted immediately prior to PPL extraction, and due to their higher concentration and limited sample volume only 1 – 2 mL was extracted of each sample onto the PPL cartridge. This was determined to provide sufficient material by testing with one PW sample that was additionally centrifuged at ODU upon my return yielding additional PW (10 – 15 cm, Core 1) (data not shown). The remaining PW of all samples was refrozen at -14 °C. MeOH extracts obtained by PPL extraction were then used for FT-ICR-MS analysis. For TOC (used for extraction efficiency determination) and TN analysis of the same select original PW samples, preparation of samples was conducted in a N\textsubscript{2} filled Glove Bag where 200 μL of the sample was added to 1.8 mL N\textsubscript{2} purged MilliQ water in a 2 mL pre-combusted amber vial, with a PTFE septum lid. Prepared samples were analysed on the same day as preparation but in the interim stored at 4 °C to maintain sample integrity.

Following analysis by FT-ICR-MS, the MeOH PPL extracts of the LW and PW samples were stored at -14 °C for their TOC and TN concentration determination, conducted a few days later.

5.2.3 TOC and TN measurements of DOM

TOC (for extraction efficiency determination) and TN concentration was
determined for the LW and select PW samples as described above. Measurements were made on a TOC-CPH with TNM attached analyser (Shimadzu). Automatic injection was used for the analysis of the 15 mL defrosted and pre-filtered LW sample and manual injection was used for PW samples. The LW sample was acidified and sparged prior to catalytic oxidation on a platinum catalyst at 720 °C and detection by a nondispersive infrared detector, determining non purgeable TOC and TN concentrations. Concentrations of LW TOC and TN were calculated using a 5 point calibration curve based on potassium hydrogen phthalate for TOC and potassium nitrate for TN, averaging 3 measurements for each data point. MilliQ water was run between each sample to confirm that there was no contamination.

For the PW sample TOC and TN determinations, manual injection of samples and standards was employed to ensure a completely air free analysis and allow lower sample volume requirements. For each determination, 100 μL was drawn into a gas-tight syringe (Hamilton) through the vial septum and using the manual injection kit provided by Shimadzu, samples were directly infused onto the catalyst. Although errors (< 3 % for 5 determinations) were slightly larger than with automatic injection (< 2 % error for automatic) it was still possible to obtain average values that are close to what would be expected for automatic injection. A 5 point calibration curve was employed using the same standards as for automatic injection for total carbon (TC), inorganic carbon (using sodium hydrogen carbonate and sodium carbonate, IC) and TN concentration analyses, determining a value from an average of 5 runs. In order to ascertain the TOC concentration for sample analyses, IC concentration is subtracted from TC concentration. MilliQ water was analysed between all samples to ensure there was no contamination and only when measurements reached background levels was a new sample analysed.
In addition to analysing the LW and PW original samples, TOC and TN concentrations were determined for the MeOH extracts of these same samples obtained by PPL extraction for extraction efficiency calculation. This was achieved after analysis of the extracts by FT-ICR-MS and removal of the same samples stored from the freezer. MeOH was removed from the extract by blowing down to dryness with N\textsubscript{2}. To ensure that all MeOH was completely removed from the extract, samples were lyophilised overnight. The solid residue resulting from the PW extracted samples was then dissolved in 15 mL of MilliQ water for analysis of TOC and TN contents. Automatic injection was utilised as described above for the original LW sample. These measurements were used to determine the PPL extraction efficiency of DOM. The resulting average extraction efficiency for all samples is 69 % OC recovery and 77 % N recovery of DOM in the PW and LW samples.

5.3.3 Determination of ammonia concentration in PW

The concentration of ammonia was determined for all PW samples collected. As mentioned above, these measurements were determined upon my immediate arrival at ODU following collection of the samples, since they remained refrigerated throughout this time. The manual phenol hypochlorite method was utilised in this study to determine ammonium (NH\textsubscript{4}\textsuperscript{+}) concentration for all samples (Solorzano, 1969), using a spectrophotometer for detection of the product indophenol blue. Reagents used for NH\textsubscript{4}\textsuperscript{+} determination were prepared and remade as necessary according to their stability during the three day period required for measurements. Standards and the MilliQ water blank were prepared at the beginning of each day. Artificial seawater was used to dilute standards of ammonium chloride to concentrations of the expected range of the samples.
All analyses were conducted in triplicate and averages were recorded for both standards and samples. Due to the large concentration range of samples, it was necessary to dilute them by varying amounts to ensure all spectrophotometer readings were in the linear range of Beer's Law (i.e., absorbance < 1 AU). In order to ensure that PW samples experienced minimal contact with air, artificial seawater used for dilution purposes was purged with N₂ prior to its addition to each sample in the N₂ filled Glove Bag. Diluted PW samples were then removed from the Glove Bag and reagents were added as quickly as possible. Unfortunately, it was not possible to add the reagents to the diluted PW samples in the Glove Bag due to the size constraint of the Bag. Errors associated with NH₄⁺ measurements were calculated as a standard deviation of the three measurements for each of the samples.

5.3.4 Elemental analysis (EA) of N in the sediment

The percent of N was determined of selected dried sediments (same depths as the PW samples) that were analysed in triplicate, and used to calculate an average value, and standard deviation. TN percentages were calculated by use of a 5 point calibration curve of nicotinamide (Sigma-Aldrich). This was completed by using a Thermo Finnigan Flash EA 1112 Series elemental analyser (CE Instruments, U.K.).

5.3.5 FT-ICR-MS analysis of LW DOM, PW DOM and sediment extracts

All of the liquid pyridine sediment extracts were diluted 10-fold with MeOH and the LW and PW extracts were analysed without dilution. All samples were adjusted to slightly basic pH with ammonium hydroxide immediately prior to analysis by negative ionisation MS. All FT-ICR-MS analyses were performed on a Bruker Daltonics 12 Tesla
Apex Qe 175 FT-ICR-MS at the College of Sciences major instrument cluster facility. The sediment and DOM extracts were directly infused into the Apollo II ESI ion source at a rate of 120 µL/hr using a syringe pump. Before addition of samples, the instrument was externally calibrated using polyethylene glycol which was accumulated for 1.0 sec in the hexapole before being transferred to the ICR cell and 20 transients were co-added. DOM extracts were analysed using a 0.5 sec accumulation time in the hexapole with 500 co-added transients. Sediment extracts were analysed using a 0.3 – 0.5 sec accumulation time, adjusted to produce the highest quality spectra possible, with addition of 300 co-added transients. ESI voltages were altered for each sample as needed. A 4 MWord time domain was used for the free induction decay. Zero-filling was completed once on the summed FID signal and then Sine-Bell apodised before Fourier transformation and magnitude display using the Bruker Daltonics Data Analysis software v. 1.3™.

Internal calibration was achieved using internal fatty acids and homologous series found within samples by Kendrick mass defect analysis (Kendrick, 1963; Sleighter et al., 2008). Molecular formulae were assigned using the Molecular Formula Calculator (Molecular Formula Cal. v.1.0, (NHMFL, 1998)) with a maximum error allowed of 1 ppm. Unrealistic formulae were removed using a conservative set of rules described in Wozniak et al. (2008) developed for Matlab (v. 7.4.0, The Mathworks Inc., Natick, MA).

5.3 Results and Discussion

5.3.1 Bulk analysis

An initial examination of the NCOM contents on the bulk level provides an important background for molecular level examination. The bulk amount of N in the PW
and sediment are displayed in a depth profile in Fig. 28. A depth profile gives an insight into the OM chemistry of the MLB sediment environment over time (approximately 5,000 years) (Hatcher, 1978). When comparing the N profiles for the PW and sediment it can be readily observed that there is little correlation between them in terms of general trends with increasing depth. The sediment appears to have more oscillations in concentration than the PW, which steadily increases in concentration in all depths except the deepest. This is obviously due to the varying presence of inorganic shells that were found inhomogeneously distributed throughout the core. On the other hand the sediment N % does not seem to increase as a function of depth; this result is similar to past study of this sediment, where it has been observed that N % remains approximately 2 % to a depth of 9 m (Hatcher, 1978). The TN concentration for the LW is 36 μM, significantly lower than the PW concentration, probably because much of the N in the LW system is being utilised by algae observed to be growing profusely in the waters. These results confirm that N is in much higher concentration in both the PW and the sediment than in other systems, as also shown in earlier studies of this system (Berman and Bronk, 2003; Emerson and Hedges, 2008; Hatcher, 1978; Wetzel, 2001). The concentration of ammonia in the PW increases over the entire depth profile with oscillations throughout. A probable reason for this is that remineralisation occurs throughout the core, producing NH$_4^+$ so that it increases in concentration progressively. The ammonia depth profile is not very similar to the concentration of TN, except for the fact that it increases with depth overall, indicating that the concentration of ammonia is affected by a variety of processes occurring in the sediment environment; similar to earlier studies (Hatcher et al., 1982; Orem et al., 1986). It is possible that when measuring the concentration of ammonia, the unavoidable exposure of the PW samples to air when adding the reactants (releasing
some of the ammonia as gas) could have adversely affected the results, making any real
trends present in their concentration difficult to differentiate. Nevertheless, it appears
that a large fraction of the TN concentration in the PW is composed of ammonia,
especially at shallow horizons.

Fig. 28. Depth profiles of TN concentration of the PW (mM), N % of dry sediment
weight and NH$_4^+$ concentration in the PW (mM) in the samples collected from MLB in
November 2009. The error bars reflect the standard deviations determined from
replicates (error bars for PW are smaller than the points).

5.3.2 FT-ICR-MS analysis of the sediment, PW and LW extracts

Nine different depth horizons were chosen for FT-ICR-MS analysis of the
sediment and PW, spanning the depths from 0 – 1 cm to 320 – 330 cm. The same
horizons were analysed for the two different sample types (PW and sediment) so that
direct comparisons can be made between samples of the same depth. Samples from both
collected cores are used for this analysis to cover the entire range of depths collected.
One sediment depth common to both cores (65 – 70 cm) was analysed from each core to
assess whether the two cores are comparable in terms of their molecular composition.
Following the close examination of the two sediment extracts of 65 – 70 cm, it is determined that the sample from Core 1 matched that of Core 2a almost exactly (data not shown, Core 2a sample used here). It should be noted that for both sample types analysed by FT-ICR-MS, the extraction protocols have been shown in past studies to remove the highest amount of OM possible to date from these types of OM pools (Dittmar et al., 2008; McKee and Hatcher, 2011). Therefore, I am confident that for the two different sample types, the extracts from the sediment, LW and PW are representative of the whole, original samples OM composition; making comparisons between them reasonable. For the mass spectra displayed of the PW and sediment extracts, only 1 mass/charge (m/z) region (336 – 337 Da) is shown since a narrow mass range demonstrates the differences and similarities between the samples most effectively. This 2 Da region is shown as an example only for evaluating the relative differences between the samples and is generally representative of the kinds of changes that are generally observed in the remainder of the spectral regions.

Fig. 29 displays the expanded mass spectra for four of the sediment extracts analysed by FT-ICR-MS as examples for the other sediments analysed. Each of the labelled peaks has been assigned elemental formulae and these are listed in Table 7. Labelling of peaks incorporates all of the spectra in this study so that easier comparison is possible between them. Evident in Fig. 29 is the significant changes in the number and relative magnitude of peaks in this depth profile. The MS of 0 – 1 cm displays only 4 peaks in this narrow mass range, however, when progressing through the core the MS of 320 – 330 cm contains 9 different peaks.
Table 7

Data pertaining to the pyridine extracted sediments and PPL extracted PW and LW DOM analysed by FT-ICR-MS negative ion mode of the spectra of the 336 – 337 Da mass range of a depth profile from MLB. Peaks are labelled in mass spectra, and their exact mass and formulae are displayed below.

<table>
<thead>
<tr>
<th>Peak number</th>
<th>Exact mass</th>
<th>Formula</th>
<th>Peak number</th>
<th>Exact mass</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>336.091117</td>
<td>[C_{16}H_{18}NO_5S]^-</td>
<td>19</td>
<td>337.132647</td>
<td>[C_{14}H_{25}O_7S]^-</td>
</tr>
<tr>
<td>2</td>
<td>336.108876</td>
<td>[C_{16}H_{18}NO_7]^-</td>
<td>20</td>
<td>337.140510</td>
<td>[C_{16}H_{21}N_6O_2]^-</td>
</tr>
<tr>
<td>3</td>
<td>336.132238</td>
<td>[^{12}C_{20}^{13}CH_{19}O_4]^-</td>
<td>21</td>
<td>337.146113</td>
<td>[^{12}C_{20}^{13}CH_{22}NOS]^-</td>
</tr>
<tr>
<td>4</td>
<td>336.145261</td>
<td>[C_{17}H_{22}NO_6]^-</td>
<td>22</td>
<td>337.147904</td>
<td>[C_{18}H_{25}O_7S]^-</td>
</tr>
<tr>
<td>5</td>
<td>336.153367</td>
<td>[^{12}C_{17}^{13}CH_{23}O_6]^-</td>
<td>23</td>
<td>337.151274</td>
<td>[C_{18}H_{29}O_4S_2]^-</td>
</tr>
<tr>
<td>6</td>
<td>336.189752</td>
<td>[^{12}C_{18}^{13}CH_{27}O_3]^-</td>
<td>24</td>
<td>337.165662</td>
<td>[C_{18}H_{25}O_6]^-</td>
</tr>
<tr>
<td>7</td>
<td>336.229509</td>
<td>[^{12}C_{16}^{13}CH_{35}O_4S]^-</td>
<td>25</td>
<td>337.169033</td>
<td>[C_{18}H_{29}O_6S]^-</td>
</tr>
<tr>
<td>8</td>
<td>336.327189</td>
<td>[C_{22}H_{42}NO]^-</td>
<td>26</td>
<td>337.180918</td>
<td>[C_{22}H_{25}O_3]^-</td>
</tr>
<tr>
<td>9</td>
<td>337.042118</td>
<td>[C_{12}H_{17}O_7S_2]^-</td>
<td>27</td>
<td>337.184289</td>
<td>[C_{18}H_{29}O_3]^-</td>
</tr>
<tr>
<td>10</td>
<td>337.059876</td>
<td>[C_{12}H_{17}O_6S]^-</td>
<td>28</td>
<td>337.202048</td>
<td>[C_{18}H_{29}O_5]^-</td>
</tr>
<tr>
<td>11</td>
<td>337.075133</td>
<td>[C_{16}H_{17}O_6S]^-</td>
<td>29</td>
<td>337.205418</td>
<td>[C_{18}H_{33}O_5S]^-</td>
</tr>
<tr>
<td>12</td>
<td>337.078503</td>
<td>[C_{13}H_{21}O_6S_2]^-</td>
<td>30</td>
<td>337.223177</td>
<td>[C_{18}H_{33}O_7]^-</td>
</tr>
<tr>
<td>13</td>
<td>337.092891</td>
<td>[C_{16}H_{17}O_8]^-</td>
<td>31</td>
<td>337.235748</td>
<td>[C_{18}H_{29}N_6O_2]^-</td>
</tr>
<tr>
<td>14</td>
<td>337.096262</td>
<td>[C_{13}H_{21}O_8S]^-</td>
<td>32</td>
<td>337.238433</td>
<td>[C_{20}H_{33}O_4]^-</td>
</tr>
<tr>
<td>15</td>
<td>337.104124</td>
<td>[C_{15}H_{17}N_2O_7]^-</td>
<td>33</td>
<td>337.272133</td>
<td>[C_{18}H_{33}N_6O]^-</td>
</tr>
<tr>
<td>16</td>
<td>337.111518</td>
<td>[C_{17}H_{21}O_4S]^-</td>
<td>34</td>
<td>337.311204</td>
<td>[C_{22}H_{41}O_2]^-</td>
</tr>
<tr>
<td>17</td>
<td>337.122751</td>
<td>[C_{16}H_{21}N_2O_4S]^-</td>
<td>35</td>
<td>337.330622</td>
<td>[^{12}C_{21}^{13}CH_{42}NO]^-</td>
</tr>
<tr>
<td>18</td>
<td>337.129277</td>
<td>[C_{17}H_{21}O_7]^-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 29. Pyridine extract mass spectra of 4 sediment depths collected from MLB in 2009, analysed by FT-ICR-MS negative ion mode. Numbered peaks (Table 7) are scaled to the same relative magnitude and consequently some peaks are displayed off scale.
Several peaks in Fig. 29 are only observed in one or two depths, this includes peaks 14, 17, 21, 23, 25, 32 and 33, whose assigned formulae contain N, O and S in varying amounts. Since the OM that is deposited in this system is inhomogeneous in nature it is expected that there are differences in the molecular composition of the compounds present. These peaks therefore, could be present due to a multitude of reasons including original differences in deposition conditions, and transformations that have occurred to the OM during its preservation. In contrast, peaks 27, 29 and 30 are present in all 4 samples and in approximately the same relative magnitude throughout. The assigned formulae are \([C_{19}H_{29}O_3S]\), \([C_{16}H_{33}O_5S]\) and \([C_{16}H_{33}O_7]\) for peaks 27, 29, and 30, respectively. Two of these three constant peaks are S-containing which seems reasonable considering that although S concentration was not measured in this study, a past study of this system revealed that due to the anaerobic character of deposition, the PW of MLB is high in sulphide concentration (Orem et al., 1986). If some of this S is incorporated into the OM of the sediment through substitution reactions with O, then the formulae of peaks 27 and 29 are likely a result of this. The reason(s) why these particular peaks are seemingly preserved in the sediment are not clear and is an avenue for future study which is beyond the scope of the present study.

Because the focus of this study is upon N-containing formulae, of particular note is peak 8. The relative magnitude of this peak increases gradually through the core, so that in the MS of 320 – 330 cm it dominates the spectrum. The formula of this peak \([C_{22}H_{42}NO]\) corresponds to an alkyl amide previously identified in MLB (McKee and Hatcher, 2010). The \(^{13}\text{C}\) isotope of this amide is assigned to peak 35 whose presence is correlated to that of peak 8.

The same mass range (336 – 337 Da) is examined in detail for the PW extracts at
the same corresponding depth intervals as the sediments, and the resulting spectra for the 4 depths are displayed in Fig. 30. These spectra at first glance, look very different from those of the sediment, despite the fact that they are from the same corresponding depths. Similar to Fig. 29, peaks are labelled according to their assigned formulae (1 – 35) in Fig. 29, and those corresponding formulae are displayed in Table 7. In general, it appears that more peaks are observed for the lower depths than those most shallow, especially at low mass defect. Similar to Fig. 29 of the sediment, the MS of depth 0 – 1 cm, has far fewer peaks than the three lower depths. Many of the peaks in these MS are observed throughout the core, with few differences by depth except the shallowest, including peaks 9, 11, 12, 14, 16, 18, 24, 28, 29 and 34, containing N, O and S in their formulae. In addition, the relative magnitude of these peaks does not seem to vary systematically with depth, indicating that these particular formulae seem to be preserved between 20 – 25 cm to 320 – 330 cm in the PW. Out of these listed peaks only peak 29 is observed in the sediments suggesting that generally, the molecular composition of sediments are different than those of PW.
Fig. 30. PW DOM extract mass spectra (336 – 337 Da) of a depth profile from MLB sediment collected in 2009 and analysed by FT-ICR-MS negative ion mode. Numbered peaks are assigned formulae (listed in Table 7). * indicates an unassigned formula.
Peak 8 dominates the spectra for all four depths with an increase in relative magnitude through the core, and corresponds to the formula $[C_{22}H_{42}NO]^-$. Due to its high relative magnitude in all four spectra the $^{13}$C isotope can also be observed (peak 35). It therefore, appears that peaks 8 and 35 are common to both the PW and sediment MS, as suggested above for the sediment peaks, these peaks likely correspond to a structure of alkyl amide. Based on the information from Fig. 29 and 30, the formula $[C_{22}H_{42}NO]^{-}$ is observed in both OM pools, and since this study investigates the LW OM also, I will now examine the corresponding MS to ascertain its general characteristics and whether peak 8 is observed.

Following examination of the sediment system, the LW DOM is also compared to assess the molecular differences between them (Fig. 31). The LW MS is more similar in number and distribution of peaks to the PW samples than the sediments samples. This observation is reasonable since the PW and LW are both in aqueous phase and it is expected that there is some exchange between the two OM pools. There are more peaks identified in the 336 m/z range than the sediment or PW samples in the LW sample, with most of them being unique to this sample (peaks 2, 5 and 6). Peaks 5 and 6 are $^{13}$C isotopologues, and the reason these peaks are not observed in the PW and sediment samples is because the corresponding $^{12}$C peak of peaks 5 and 6 are too low in relative magnitude. Comparing the LW spectrum to the PW spectra at lower depth it appears that there are fewer peaks with low mass defect in the LW sample than the PW spectra, possibly indicating that the OM in the LW is more oxygen-rich (Table 7). Many of the peaks in the 337 m/z region are similar to those in the PW samples, and this is probably due to the fact that the OM in the PW is being exchanged with the LW OM.

Of particular interest in this study is the N characteristics of this system, and
similar to the sediment and PW spectra peak 8 and its isotopologue are present. Peaks 8 and 35 can be observed in all spectra of this study, one of the few peaks that are ubiquitous between all three sample types. As a result of the MS examination of this mass range I will assess whether it is an important component of this system overall by examining the bulk MS properties of these MS.

Fig. 31. LW DOM PPL extracted mass spectra from MLB analysed by negative ion FT-ICR-MS. Peaks are numbered, and their assigned formulae are shown in Table 7.

5.3.3 Formula type classification

In order to understand the relative importance of a class of compounds it is useful to categorise the assigned formulae into their respective formula types. Therefore, assigned formulae for samples in this study were parsed according to their N, O, S and P contents and the results are displayed in Fig. 32. This Figure shows that the downcore formula type classifications for the aqueous and sediment samples are very differently distributed. In the LW and for many of the PW samples, CHO and CHOS containing formulae dominate, together accounting for about 70 % of all assigned formulae. CHO containing formulae often dominate FT-ICR-MS spectra due to the ubiquity of many structures that fall in this classification, particularly fatty acids (Sleighter et al., 2008).
The MS spectra of Fig. 30 and 31 indicated that S is a major component of the DOM, and the formula type classification seems to confirm this. It is not unexpected that this observation is made considering the high amounts of hydrogen sulfide found in PW, mainly the results of sulfate reduction in an anaerobic system fuelled by the advection of sulfate-rich seawater through the porous bedrock (Hatcher et al., 1982).

Another difference between the two sample types is the depth profile trend. In the sediment, all depth horizons seem to have a similar distribution of formula types, whereas in the PW samples, the percentage of CHO containing formulae decreases while CHOS containing formulae increases. This depth profile for the PW perhaps indicates that the CHO containing formulae are incorporating S with depth as the amounts of hydrogen sulfide increase and more time evolves. It is interesting to note that the formula type classification of the LW is more similar to the PW than the sediment, giving further evidence that there is a substantial mixing between the DOM of the PW and LW but not between the sediment and PW. The dominating N-containing formula types of the LW DOM are CHNO and CHNOS, formulae that could be attributed to peptides and alkyl amides, as has been suggested as dominant for NCOM of the sediment in a previous study of this Lake (McKee and Hatcher, 2010).
Fig. 32. Formula type classifications as determined by N, O, S and P contents, calculated as a percent of total number of formulae assigned for PW and LW DOM and sediment extracts analysed by FT-ICR-MS negative ion mode.

The sediment formula classification looks, at first glance, very different from the DOM samples since CHO containing formulae are not dominant and those containing N,
i.e., CHNOS and CHNO are more dominant (30 - 40 % of total formulae). This indicates that there is a larger diversity of formula types containing N present in the sediment than in the DOM samples. The increase in total N (Fig. 28) in the PW is not well reflected in the formula distributions, perhaps suggesting that much of the additional N contribution at lower depths is from inorganic forms (ammonia, nitrates, nitrite). However, N-containing formulae still contribute 10 - 15 % of all assigned formulae in the PW and LW samples, representing a significant fraction of the MS peaks. The significant fraction that comprises N-containing formulae (both sediment and DOM) in this system means that they play an important role in the chemistry of this Lake. However, the structural source of this N pool is still unknown.

5.3.3 Assessing the sources of N-containing formulae

The dominating N-containing formulae, CHNO and CHNOS, could have many sources including peptides, aminosugars and alkyl amides. In order to determine the structural source of the N-containing formulae I plotted the data on an N/H versus O/C diagram. This plot has been demonstrated in a past study (Chapter 3) to portray the N structural characteristics of an OM sample. Fig. 33 shows the resulting diagrams for the LW sample and two depths of the PW and sediment samples as examples for the other depths in this study (0 - 1 and 320 - 330 cm). The locations of the formulae displayed in the Figure are distributed according to their N, O, C and H contents. Consequently, the plot is divided into distinct regions which reflect the most likely sources of the N-containing formula. This Figure shows that the N contents of these 5 samples are very different, due to a variety of likely origins. The LW sample contains formulae that can be assigned likely structures of peptides, lipids and aminosugars, consistent with the idea
that surface waters of MLB are highly productive algal factories in which soluble metabolic products of primary productivity are being supplied to the LW. In addition, many of the data points are vertically aligned on the plot indicating homology of formulae by saturation series, probably due to common formation pathways. Most of the formulae of the LW sample contain more N than O, since most of the data points lie above the slope of 0.5, (# N = # O in formulae for a slope of 0.5). When comparing the LW sample to the PW sample of 0 – 1 cm, a clear difference in their N contents can be observed. Almost all of the formulae in the PW sample are N-poor lipids; in addition, many of these formulae lie in a homologous series (CH₂) determined by lines connecting points through the origin.

The PW plot of 320 – 330 cm is very different from that of 0 – 1 cm. The average O/C of the formulae in the 320 – 330 cm PW sample is much larger, centred on 0.4, as opposed to 0.1 for the 0 – 1 cm sample. When examining the PW samples as a function of depth, formulae become more O enriched, but retain approximately the same amount of N (generally 1 per formula). In addition it seems that, from this plot that is based on distribution of points, the N type remains about the same for the PW samples, formulae consistent with N-poor lipids. However, the amount of C in these N-poor lipid formulae is being reduced with depth.
Fig. 33. Plots of N/H versus O/C for two sediment samples, two PW samples and one LW sample from MLB. All samples are extracts analysed by FT-ICR-MS negative mode analysis. The plots are divided into distinct regions linking the N-containing formulae to likely source material.

Comparing both of the PW samples to the LW sample, it is immediately obvious that the N-containing DOM of the shallow PW sample is quite different in the character
of N-containing formulae, since its distribution of data is located in a different region. Perhaps the reason for this difference is that, once N-containing OM is deposited in the sediment system, the DOM changes dramatically, so that the shallow PW is very different in nature to that of the LW above it despite its similar formula type classification (see Fig. 32). Perhaps the LW DOM has undergone substantial remineralisation, similar to the PW DOM of 320 – 330 cm depth, thus explaining their similarity in N structure. Previous studies show that the upper sediment horizons of MLB are quite rich in carbohydrates and proteins and have not undergone much diagenesis (Knicker et al., 2001). As they decompose over time, they release soluble DOM to the PW and it is these substances that dominate the distribution. The existence of N-poor lipids is perhaps characteristic of freshly-deposited organic matter. At deeper horizons, most of the labile sedimentary organic matter has disappeared or been reduced in abundance. The DOM is accordingly reflective of this and dominated by more recalcitrant organic matter composed of more oxygen-containing species.

The sediment plots are more similar in data point distribution for N-containing formulae than the DOM samples, except that the near surface sample appears to be more enriched in peptides and other labile materials as expected based on the NMR studies of Knicker et al. (2001). The main regions the peaks are distributed within for the two sediment samples are N-poor lipids, N-rich lipids and peptides. In general, there are fewer peaks in these three regions for the lower depth sample than the shallow, and is probably because many of the compounds present in the shallow sample have been degraded through diagenesis (Hatcher et al., 1983). The increase in relative magnitude of peaks was observed in the MS of Fig. 29 and appears to be consistent throughout the MS. Therefore, those sample peaks and attendant structures that remain and are identified in
the sample from 320 – 330 cm can be considered more recalcitrant in nature. In several locations of the plots for both sediment samples, homologous series (CH2) of structural types can be identified since data points fall on lines trending towards 0. This suggests that many of the formulae identified in the sediment samples are related in series, and probably in formation source/mechanism, similar to the homologous series in the DOM samples.

5.3.4 Occurrence and formation of alkyl (di)amides

The common N-containing structure type is N-poor lipids for all plots of Fig. 33. In this class of compounds are formulae that can be identified as alkyl (di)amide, i.e., with the general formula of C_{x}H_{y}N_{x}O_{x} where x is 1 or 2. Following this classification, the consistent peaks (peaks 8 and 35) observed in the MS of the LW, PW and sediment samples (Fig. 29, 30 and 31) fall into this group. The alkyl (di)amides comprise a significant fraction of the N-poor lipid formulae. In the previous work (McKee and Hatcher, 2010) evidence was provided that these structures can be produced via amidation of acyl glyceride esters by reaction with ammonia, possibly, but not proving, their origin in MLB. The essential ingredients for amidation are abundantly present in MLB (Hatcher et al., 1982), so I can hypothesise that one might observe a relationship between the production of alkyl amides and the availability of ammonia and acyl glycerides.
To this end, the abundance of alkyl amides in samples is assessed by firstly summing those formulae that fall into the general formula of identified alkyl amides and diamides identified above and comparing their total to the total number of CHNO containing formulae, displayed as a percentage. The results for the LW, PW and sediment samples are displayed in Fig. 34 as a function of depth. The plot shows that alkyl amides and diamides dominate as a large percentage of the CHNO containing formulae present in the PW samples and the sediment. For comparison, the average amount of N in each assigned formula is also displayed in Fig. 34 for the PW and sediment samples. Throughout the core, the average amount of N in assigned formulae
of the sediment is higher than observed for formulae of the PW samples. Since the sediment samples contain a large number of structurally diverse N-containing formulae, as shown in Fig. 33, the relative proportion of alkyl amides and diamides, when comparing to all CHNO-containing formulae in the sediment, is reduced. The PW data for N-containing structures is not as diverse, and, consequently, the alkyl amides dominate over other formulae (Fig. 33).

Additionally plotted in Fig. 34 is the concentration of ammonia in the PW by depth, as duplicated from Fig. 28. The trend as a function of depth displayed in the % amides plot, generally indicates an increase for both PW and sediment samples, mirroring the trend for the ammonia concentration in the PW. However, the percent of alkyl amides and diamides in the LW is very low, only about 3 % of all CHNO containing formulae. This is not surprising since Fig. 33 displays formulae in LW that can be attributed to a wide variety of structural entities which “dilute” the relative amounts of alkyl amides. These other compounds are the labile N-containing substances like peptides which are continuously being degraded and entrained within the PW in the top 100 cm of the sediment system. It is likely that when N-containing OM is deposited in the sediment system, diagenesis occurs immediately, so that the OM is altered significantly as shown in Fig. 33 by the large difference in N structural types of the PW samples. As a result, ammonia is produced in high concentration in the PW. In parallel, amidation occurs, utilising much of the ammonia produced in the PW, resulting in the dominance of alkyl amides and diamides in the shallowest depth, with small amounts precipitating out into the sediment. As diagenesis of other more labile N-containing OM increases (reflected by the percentage calculation) with depth, the concentration of ammonia increases, and, consequently, the alkyl (di)amide percent increases. Since it is
likely that amidation occurs in the PW, as the alkyl (di)amides are produced they gradually precipitate out of the PW due to the low solubility of these compounds and they become incorporated in the sediment. In the PW and sediments at the depth of 320 – 330 cm there is almost an equal percent of alkyl (di)amides in the PW and sediment, about 25 % of all CHNO containing formulae for each pool and 50 % of the N-pool in the entire sediment system. This mirroring of alkyl (di)amide percent in the PW and sediment with the ammonia concentration gives substantial credence to the previously proposed production pathway of amidation and indicates that this reaction is occurring throughout the depth core of this study.

5.4 Conclusions

The goal of this study was to examine the molecular level N chemistry of OM deposited throughout an anaerobic lacustrine environment, by studying the LW, PW and sediments. FT-ICR-MS analysis allowed me to study in detail those formulae that contain N. When studying the N characteristics of these samples, it became obvious that the N present in the PW, LW and sediment is derived from completely different structural entities, probably because of differences in compound solubility and remineralisation stage. One of the few common N structural types is N-poor lipids, and this was investigated further by examining one particular formula type, alkyl amides. When I examined in detail the abundances of alkyl amides relative to the overall N contents of the PW and sediment and compared it with the concentration of ammonia with depth, I am able to clearly demonstrate a correlative relationship which proves the importance of amidation as a formation pathway for these compounds.
6.1 Introduction

Dissolved organic nitrogen (DON) is a major component of marine and freshwaters, comprising up to 50 % of soluble N (Wetzel, 2001). Due to their widespread prevalence in many organisms, proteins are often major constituents of the DON pool. However, due to their high lability, proteins undergo rapid decomposition; most significantly through hydrolysis (Herbert, 1999; Mulholland and Lee, 2009; Roth and Harvey, 2006). The resulting peptides and free amino acids are subsequently removed from the DON pool by direct uptake (if < 600 Da for bacteria) (Alexander, 1973), or by extracellular degradation (Arnosti, 2011; Obayashi and Suzuki, 2008; Pantoja and Lee, 1999; Pantoja et al., 1997). A small fraction is preserved, and enters the sediment where it can be further remineralised or sequestered (Fogel and Tuross, 1999; Keil and Kirchman, 1994; Nguyen and Harvey, 1997; Nguyen and Harvey, 2001; Tanoue et al.,

This text is being prepared for submission to Marine Chemistry.

As a means to explain why certain peptides and amino acids survive for unexpectedly longer periods of time in natural waters and sediments, a number of hypotheses have been suggested to explain peptide preservation. These include: hydrophobic and non-covalent interactions (Nguyen and Harvey, 2001), selective preservation (Hatcher et al., 1983), protection by encapsulation into organic matter (OM) (Knicker and Hatcher, 1997; Zang et al., 2000) and mineral matter (Gordon and Millero, 1985; Mayer, 1994a; Mayer, 1994b), chemical structure (Liu et al., 2010) and formation of adducts to OM (Hsu and Hatcher, 2006; Kirchman et al., 1989; Liu et al., 2010; Rillig et al., 2007). It is probable that a combination of these mechanisms is important for peptide preservation, although it is not currently clear which these are.

Of particular interest in my current study of peptide preservation is the one that suggests that the inherent differences in chemical reactivity are due to differences in the sequence of individual amino acid residues (e.g., chemical structure). Chemical structure was investigated by Liu et al. (2010) when they examined differences in peptide removal rates in natural water for two tetrapeptides: alanine-valine-phenylalanine-alanine (AVFA) and serine-tryptophan-glycine-alanine (SWG). Their results suggest that peptide length is a key factor in the rate of removal in the water but not sequence. Although their results are informative, they are not shown to be universal and a more extensive study is required involving a wider range of chemical characteristics of amino acids and length of the peptide chain.

Although there is almost complete ubiquity of peptides in natural waters there have been few studies aimed at investigating whether differences in peptide preservation exist at different study sites. The wide variety of chemical characteristics, micro-
organism community and environmental conditions suggests that peptide preservation should differ depending on location. The presence or absence of oxygen in natural systems is generally considered to greatly affect the amount of preservation of OM, and there have been several studies to confirm this (Gray et al., 2002; Hartnett et al., 1998; McCarthy et al., 1996). Past studies suggest that there are distinct differences in peptide removal rates between fresh and marine waters (Liu et al., 2010; Roth and Harvey, 2006). The implications of these studies are limited since they only study locations with similar characteristics, i.e., estuarine environments (Roth and Harvey, 2006) or along one transect (Liu et al., 2010). Environmental conditions such as exposure to anthropogenic sources, natural water type (e.g., riverine vs lacustrine systems) and OM composition have not been thoroughly considered with respect to their effect on peptide removal rates.

The overall objective of this study is to investigate some of the possible reasons why some peptides seem to be selectively preserved over others in natural water by examining removal rates of a series of synthesised peptides which have defined characteristics. Firstly, I will examine the effect on removal rates in aqueous systems as a factor of peptide length and sequence and secondly, I will compare removal rates for select tetrapeptides in two brackish waters, one riverine and the other lacustrine; Elizabeth River (ER), VA and Mangrove Lake, Bermuda (MLB). The micro-organism community and level of anthropogenic influence in these two natural waters are very different, presumably leading to significant differences in their peptide removal rates. High performance liquid chromatography (HPLC) will be utilised in this study to monitor peptide concentrations by UV-Vis detection. Ribulose-1, 5-bisphosphate carboxylase/oxygenase (RubisCO) is ubiquitous in natural waters; I, therefore, chose sequences that are components of this protein and are likely released upon its
remineralisation in natural systems.

6.2 Materials and Methods

6.2.1 Preparation of peptide standards

For all syntheses the highest quality reactants available were used. Fluorenylmethyloxycarbonyl chloride (FMOC) amino acids were purchased from Protein Technologies. Resins were purchased from Advanced Chemtech (alanine, glycine) and Nova Biochem (valine).

I used FMOC solid phase peptide synthesis (SPPS) on an automated solid phase peptide synthesiser (PS3, Protein Technologies, AZ), to produce all of the peptides used in this study; a method that reproducibly produces peptides of high purity (Chan and White, 2000; King et al., 1990; Merrifield, 1963; Stewart, 1976). For more details of the protocol see Liu et al. (2010). Briefly, FMOC protected amino acids were selectively deprotected at the amino group and added stepwise to a polymeric support resin (Wang in my study, with either alanine, glycine or valine pre-attached) (Fields et al., 1991; Wang, 1973). Following attachment of all amino acids to the resin, the FMOC protecting groups and the solid resin were removed using a mixture of anisole (Fluka), thioanisole (Fluka) and trifluoroacetic acid (Fluka), known as the cleavage cocktail, without disrupting the newly formed peptide bonds (Chan and White, 2000). The same procedure as that described in Liu et al. (2010) was used in this study to isolate the unpurified peptide briefly involving removal of the cleavage cocktail, extraction into aqueous solvent, and lyophilisation. Four different peptides were synthesised at 0.4 mmol concentrations using this method including alanine-valine-phenylalanine-alanine (AVFA), alanine-
arginine-glutamic acid-glycine (AREG), alanine-phenylalanine-arginine-valine (AFRV)
and alanine-alanine-phenylalanine-arginine-valine (AAFRV).

For FT-ICR-MS analysis the relevant collected peaks were directly analysed on a
Bruker Daltonics Tesla Apex Qe 175 FT-ICR-MS at the College of Sciences major
instrument cluster (COSMIC) facility at Old Dominion University (ODU). The collected
fractions were acidified to pH 6 using formic acid (Sigma-Aldrich) and directly infused
into the Apollo II electrospray (ESI) ion source at a rate of 120 μL/hr using a syringe
pump. Polyethylene glycol was used as an external calibrant and was accumulated for
1.0 sec in the hexapole before being transferred to the ICR cell and 20 transients were co-
added. The collected fractions were analysed using a 1.0 sec accumulation time in the
hexapole with 10 co-added transients. A 4 MWord time domain was used for the free
induction decay (FID). Zero-filling was completed once on the summed FID signal and
then Sine-Bell apodised before Fourier transformation followed by magnitude calculation
using the Bruker Daltonics Data Analysis software v. 1.3™. Resulting mass/charge (m/z)
data for each peptide were confirmed as being correct by comparing to their exact mass
(See Appendix B).

For 1 H NMR analysis, a 400 MHz Bruker Avance spectrometer (Bruker Biospin,
Inc.) located at the COSMIC facility was utilised. The pulse programme included water
suppression for optimal spectral quality. Spectra were matched to simulated ones
generated by a 1 H NMR predictor (ACD Labs v. 9.0) (See Appendix B).

6.2.2 Sampling locations and collection

Two different water samples were used in this study, ER and MLB (See Fig. 35
for a map). ER is a major tributary of the Chesapeake Bay, surrounded by large numbers
of industrial companies on either side of the river, and its natural sources include groundwater from the Columbia and Yorktown aquifers and the Great Dismal swamp (Charette and Buesseler, 2004; Helms et al., 2008; Minor et al., 2007; Shafer et al., 2004). 2.5 L of water was collected in a precombusted amber glass bottle from the side of the dock of the ODU sailing center (30 cm depth) at 7.11 pm on 30th July 2010. At 7.02 pm low tide occurred for this location at 0.49 feet. The previous evening, a large rain storm (4.23 inches) occurred, contributing to the OM dissolved in the river. Water not used for the initial incubation studies described below was stored frozen at -14 °C. An additional 500 mL of ER water needed for tests on the effects of filtration was collected on 30th August 2010 in a precombusted glass bottle; conditions were similar to those of the July collection.

MLB is a brackish lake, with its primary source of OM originating from algae and its source of water is through permeation of the bedrock and rainfall (Hatcher, 1978; Hatcher et al., 1982; McKee and Hatcher, 2010). It is located in a site distant from any anthropogenic sources, except a nearby golf course (see Fig. 35 for location). Water from a depth of 30 cm was collected in November 2009 from a home-built raft in 2.5 L pre acid-washed polypropylene bottles. Unfiltered samples (those used for this experiment) were immediately frozen upon return to the Biological Institute of Ocean Sciences (BIOS) approximately 1 hour after collection. There was no recorded rain in Bermuda during collection time and the air temperature was within the average for the season. Frozen samples were transported back to ODU by the R/V HSBC Atlantic Explorer and were stored at ODU at -14 °C and defrosted overnight at 4 °C on 29th July 2010.
Fig. 35. Maps showing the locations of water collected in MLB and ER and used for peptide incubations. The surrounding area maps around the sampling location are inset into the main maps. White boxes on the surrounding area maps show the locations of the expanded region and the X on the expanded maps denotes the exact locations of sampling.
Measurements were made in situ for water temperature, pH and salinity at both sites using a waterproof pH and temperature tester (pHep 5, Hanna Instruments) and portable refractometer (Ultraprecision Brix handheld Goldberg refractometer, Reichert Analytical Instruments); results are listed in Table 8. For total organic carbon (TOC), dissolved organic carbon (DOC), total nitrogen (TN) and dissolved nitrogen (DN) measurements, 2 x 10 mL of water were removed of each sample, one was used for unfiltered water measurements (TOC and TN), the other was filtered using a 5 mL sterile polypropylene syringe (NORM-JECT) and 0.22 μm PVDF sterile filter (Fisherbrand) setup, to obtain DOC/DN numbers. Two 10 mL aliquots of each sample were additionally placed in sterile centrifuge tubes, 2 drops of formaldehyde (Fisher) were added to each tube and the samples were stored at 4 °C, until use for bacteria concentration determination.

6.2.3 Bulk measurements of ER and MLB waters

6.2.3.1 TOC, DOC, DN and TN measurements As stated above, samples of the water collected were prepared and set aside for analysis of TOC, DOC, DN and TN. Measurements were made on the same day as the degradation experiments were started so it was unnecessary for them to be stored. Samples were acidified and sparged prior to catalytic oxidation on a platinum catalyst at 720 °C and detection by a nondispersive infrared detector, determining non purgeable TOC and TN (Shimadzu TOC-V CPH with TNM attached). Concentrations were calculated using a 5 point calibration curve for all samples, using an average of 3 measurements for each concentration. Relative errors associated with replication of injections were 1.7 % for TOC and 7.5 % for TN. MilliQ water was analysed between samples to confirm that there was no contamination, and
was used to determine the background signal for the instrument of 33 \( \mu \text{M} \) for TOC and 0.8 \( \mu \text{M} \) for N measurements. Table 8 lists the resulting concentrations.

Table 8

Bulk measurements of the ER and MLB water recorded at the place of sampling and in the laboratory.

<table>
<thead>
<tr>
<th>Temp. (^{\circ} \text{C} )</th>
<th>Salinity</th>
<th>pH</th>
<th>TOC(^2) (( \mu \text{M} ))</th>
<th>DOC(^3) (( \mu \text{M} ))</th>
<th>TN(^2) (( \mu \text{M} ))</th>
<th>DN(^3) (( \mu \text{M} ))</th>
<th>([\text{Bacteria}] ) (Cells mL(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER 28.0</td>
<td>23</td>
<td>8.9</td>
<td>433 340</td>
<td>24 16</td>
<td>5.94 x 10(^6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML 22.3</td>
<td>32</td>
<td>8.0</td>
<td>490 443</td>
<td>40 36</td>
<td>4.18 x 10(^6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) The temperature is that of the water.

\(^2\) The concentration is determined of whole unfiltered water.

\(^3\) The concentration is determined of 0.2 \( \mu \text{m} \) filtered water.

6.2.3.2 Bacteria measurements  Samples for bacteria concentration determinations were prepared as stated above. Bacterial concentrations (cells mL\(^{-1}\)) were measured on formalin-preserved whole water (before sterile filtration) using an adaptation of the 4, 6-diamidino-2-phenylindole (DAPI) staining protocol (Porter and Feig, 1980) followed by bacterial enumeration on a Zeiss Standard 16 epifluorescence microscope using an Omega bandpass, DAPI filter set and x 1,000 magnification (Table 8).

6.2.4 Incubation of peptides

Prior to collection of the ER water, stock solutions of the peptides were made at their maximum concentration ranging from 0.001 – 0.002 M. Stock solutions were prepared on the same day as collection and were stored at 4 \( ^{\circ} \text{C} \) prior to their use. For each incubation bottle (precombusted, 60 mL amber bottle) 40 mL of ER and MLB water were added to each immediately prior to the addition of the peptides. Final
concentrations of individual peptides were 10 µM in each incubation, except in a few experiments where concentrations were reduced due to limitations on the quantity of product available. Although the concentration of the peptides added to the natural waters in this study is significantly higher than natural levels, I assume that the removal pathways determined here, are similar for natural levels. Previous experiments proved that peptide incubations were very reproducible and so only one incubation bottle was prepared for each experiment (data not shown).

In addition each peptide was incubated for four weeks in MilliQ water alone as an experiment blank, and time points were taken and monitored by HPLC. Results showed that all peptides were stable throughout the time period, with no change in their concentration and no appearance of additional peaks in HPLC chromatograms.

All incubations were placed on an orbital shaker (VWR model 57018-754) that operated at 80 rpm, except a sample of AVFA which was placed on the bench to undergo degradation in a static mode. I wanted to determine if use of the orbital shaker was required. Incubations were carried out at room temperature (~ 24 °C) for 4 weeks and subsamples were harvested for analysis at specific time points. For each time point, a 2 mL aliquot was removed from each incubating bottle using an acid washed pipette tip, filtered using a 0.22 µm PVDF sterile filter and a 5 mL sterile polypropylene syringe, and placed in a precombusted 2 mL vial. All time point samples were frozen at -14 °C immediately after collection, except 36, 48, 52 and 60 hours that were stored at 4 °C, because their HPLC analysis was planned for the following day. Previous experiments showed that storage of samples for up to 2 weeks at 4 °C did not significantly alter the peptide concentration monitored by HPLC (data not shown).
6.2.5 Monitoring of peptide concentrations by HPLC

In order to facilitate a more realistic study, low concentrations of peptides are necessary; however, this is difficult when monitoring them in real time. Therefore, previous studies have attached fluorescent tags or isotopically label the peptides in order to monitor them at lower concentrations (Hollibaugh and Azam, 1983; Obayashi and Suzuki, 2005; Obayashi and Suzuki, 2008; Pantoja and Lee, 1999; Pantoja et al., 1997). These methods are problematic however, due to the interference of the tags with the lability of the peptides (Stevenson, 1994). In order to avoid these difficulties, HPLC was used by Liu et al. (2010) to monitor moderately low concentrations (10 μM) of tetrapeptides without the requirement of a tag by using UV-Vis detection. The study by Liu et al. (2010) was able to successfully detect the presence of peptides in natural water to a concentration of 1 μM, and although this concentration is significantly higher than natural quantities of peptides in natural water, it can be viewed as a model for processes occurring at lower, undetectable, concentrations.

Peptide concentrations were therefore, monitored by HPLC during incubations. The same HPLC and column was used throughout the experiment (described above) so that retention times remained constant throughout. The mobile phases were 0.05 M NaH₂PO₄ and MeOH. The MeOH was ramped on a gradient that was optimised for each run for time and optimal separation. Flow rate was 1 mL/min for all incubations except in cases where additional resolution was needed. Concentrations of each peptide were determined by comparison to standards analysed on the same day. Wavelengths monitored by the photodiode array (PDA) detector were 210 nm for all peptides, except AREG (200 nm). Duplicate runs resulted in high levels of agreement (peaks were reproduced within 5%).
6.2.6 Rates of removal calculation

Due to the fact that there was no lag time for any of the incubations in this study first order kinetics was applied to calculation of the removal rates for each peptide.

6.2.7 FT-ICR-MS analysis of ER and MLB waters

Frozen water samples stored after the incubations were thawed to room temperature and prepared for FT-ICR-MS analysis. Salts need to be removed, because they outcompete the sample peaks for charge in the ESI source (Brown and Rice, 2000; Constantopoulou et al., 1999). The use of solid phase extraction (SPE) cartridges has been demonstrated as a successful method to remove higher quantities of OM than more traditional techniques (Dittmar et al., 2008; Fontanals et al., 2005; Morales-Cid et al., 2009). Therefore, SPE extraction cartridges were utilised in this study, in particular PPL, a styrene divinyl benzene polymer (Varian). The procedure used for PPL extraction followed that described by Dittmar et al. (2008), except MilliQ water, rather than aqueous 0.01 M hydrochloric acid, was used to rinse the cartridge following extraction of the OM onto the disc. Exactly 10 mL of each water sample was passed through the cartridge and collected in 1 cartridge volume of MeOH (3 mL). The MeOH extracts were then analysed by FT-ICR-MS without further dilution, except the 1-drop addition of ammonium hydroxide in order to adjust the pH to 8 for negative ion mode analysis. Instrumentation for the FT-ICR-MS was the same as described above. External calibration was facilitated with polyethylene glycol by accumulating for 1.0 sec in the hexapole before being transferred to the ICR cell where 20 transients were co-added. MeOH extracts of ER and MLB samples were analysed using a 0.5 sec accumulation time in the hexapole with 500 co-added transients. A 4 MWord time domain was used
for the FID. Zero-filling was completed once on the summated FID signal and then Sine-Bell apodised before FT followed by magnitude calculation using the Bruker Daltonics Data Analysis software.

Internal calibration was accomplished using naturally occurring fatty acids within the sample and homologous CHO series found in the samples by Kendrick Mass Defect analysis (Hughey et al., 2001; Sleighter et al., 2008). Due to the ultrahigh resolution of the spectra it is possible to assign elemental formulae to the m/z values obtained using a Molecular Formula Calculator (Molecular Formula Cal. v.1.0, (NHMFL, 1998)). Unrealistic formulae are removed using a conservative set of rules described in Wozniak et al. (2008) developed for Matlab (v. 7.4.0, The Mathworks Inc., Natick, MA).

6.3 Results and Discussion

6.3.1 Removal rate considerations and comparison to past studies

In this study I primarily use the term “removal rate” to describe the rate of decrease in concentration of the peptides added to ER and MLB waters. There are a variety of reasons for the decrease in peptide concentration in natural waters including extracellular hydrolysis, direct uptake and adduction to existing organic and inOM in the water (Amosti, 2011; Hsu and Hatcher, 2005; Nguyen and Harvey, 2001; Nikaido and Vaara, 1985). Since initially, I am not sure which of these mechanisms dominates to reduce the concentration for the synthesised peptides in this study, the term removal is used to encompass all of the above mentioned mechanisms to infer a decrease in the amount of observed peptide in the water.

One of the peptides used in this study, AVFA, is the same as that used in the
study by Liu et al. (2010). In the past study AVFA was incubated in waters collected from the James River, Virginia to the Chesapeake Bay, along a transect, and in ER water. In my study I incubated AVFA in water from the same location in ER as reported by Liu et al. (2010). Using data of Liu et al. (2010) and the additional data in this study, comparisons can be made between three different seasons, winter 2008, spring 2008 and summer 2010. During the winter of 2008, AVFA is removed from ER water at a rate of 0.18 d^{-1}; in spring 2008 it is 0.48 d^{-1} and during the summer of 2010 the rate is 6.36 d^{-1} (Table 9). It appears that during the summer, AVFA is removed from ER water much more rapidly than during winter or spring. This seasonal difference in peptide removal from natural waters is similar to that observed by Pantoja and Lee (1999), when summer has the highest rate of removal compared to winter, probably due to the increased microorganism activity during this season.

Although much of the experimental setup for this study is the same as that of Liu et al. (2010), there is one main difference. The main difference is that incubations in this study are conducted on an orbital shaker, designed to mimic the advective mixing that occurs in natural water. For comparison to the study of Liu et al. (2010) a dual incubation experiment was conducted of AVFA on both the orbital shaker (motion) and the lab bench (static). The resulting removal rates displayed in Table 9, demonstrate that there is no difference for removal of AVFA from ER water when incubated with simulated advective mixing. Since the aliquots removed for HPLC analysis from each incubation are taken at several hour time intervals, this is not a surprising result because diffusion could easily account for the similarity in removal rates. This similarity of removal rate of AVFA from ER water between the two incubation conditions enables a reliable comparison with the study of Liu et al. (2010) because all other conditions are
Table 9

Information regarding the incubations of peptides in collected ER and MLB water analysed by HPLC. Peptides are listed in the first column, the location of the water sample is in the second, and the details of the incubation bottle are in the third column (+ means incubated in the same bottle). The last columns are the initial concentration of the peptide in the water sample, the calculated removal rate using first order kinetics and finally the regression ($R^2$) of the calculated rate. Unless stated in the incubation column, incubations in this study use unfiltered water and are conducted on the orbital shaker.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Location</th>
<th>Incubation</th>
<th>Initial Conc (μM)</th>
<th>Rate (d$^{-1}$)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVFA</td>
<td>ML</td>
<td>Single</td>
<td>10.5</td>
<td>1.10</td>
<td>0.96</td>
</tr>
<tr>
<td>AVFA</td>
<td>ER</td>
<td>Single</td>
<td>10.7</td>
<td>6.36</td>
<td>0.93</td>
</tr>
<tr>
<td>AVFA</td>
<td>ER</td>
<td>Single-static</td>
<td>9.0</td>
<td>5.62</td>
<td>0.86</td>
</tr>
<tr>
<td>AVFA</td>
<td>ER</td>
<td>Single-frozen</td>
<td>11.9</td>
<td>3.86</td>
<td>0.99</td>
</tr>
<tr>
<td>AFRV</td>
<td>ER</td>
<td>Single</td>
<td>12.9</td>
<td>8.64</td>
<td>1.00</td>
</tr>
<tr>
<td>AAFR V</td>
<td>ER</td>
<td>Single-Unfiltered water</td>
<td>13.1</td>
<td>13.2</td>
<td>0.91</td>
</tr>
<tr>
<td>AAFR V</td>
<td>ER</td>
<td>Single-Filtered water</td>
<td>12.7</td>
<td>3.50</td>
<td>0.99</td>
</tr>
<tr>
<td>AREG</td>
<td>ER</td>
<td>Single</td>
<td>10.2</td>
<td>5.03</td>
<td>0.75</td>
</tr>
<tr>
<td>AFRV</td>
<td>ML</td>
<td>Dual, AFRV + AVFA</td>
<td>17.6</td>
<td>1.67</td>
<td>0.81</td>
</tr>
<tr>
<td>AVFA</td>
<td>ML</td>
<td>Dual, AFRV + AVFA</td>
<td>12.2</td>
<td>1.33</td>
<td>0.88</td>
</tr>
<tr>
<td>AFRV</td>
<td>ER</td>
<td>Dual, AFRV + AAFR V</td>
<td>4.3</td>
<td>4.53</td>
<td>0.41</td>
</tr>
<tr>
<td>AAFR V</td>
<td>ER</td>
<td>Dual, AFRV + AAFR V</td>
<td>5.1</td>
<td>5.27</td>
<td>0.80</td>
</tr>
<tr>
<td>AFRV</td>
<td>ML</td>
<td>Dual, AFRV + AAFR V</td>
<td>11.6</td>
<td>2.53</td>
<td>0.87</td>
</tr>
<tr>
<td>AAFR V</td>
<td>ML</td>
<td>Dual, AFRV + AAFR V</td>
<td>6.3</td>
<td>2.67</td>
<td>0.86</td>
</tr>
<tr>
<td>AVFA</td>
<td>ER</td>
<td>Quad, AFRV + AVFA + AREG + AAFR V</td>
<td>20.1</td>
<td>32.1</td>
<td>0.83</td>
</tr>
<tr>
<td>AFRV</td>
<td>ER</td>
<td>Quad, AFRV + AVFA + AREG + AAFR V</td>
<td>14.4</td>
<td>7.90</td>
<td>0.85</td>
</tr>
<tr>
<td>AREG</td>
<td>ER</td>
<td>Quad, AFRV + AVFA + AREG + AAFR V</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>AAFR V</td>
<td>ER</td>
<td>Quad, AFRV + AVFA + AREG + AAFR V</td>
<td>30.9</td>
<td>14.9</td>
<td>0.78</td>
</tr>
</tbody>
</table>

ND-not detected

6.3.2 Incubations of AFRV and AAFR V

Previous studies of peptide removal from natural waters have primarily examined either tetrapeptides, shorter peptides, or entire proteins (Liu et al., 2010; Mulholland and Lee, 2009; Nunn et al., 2003; Pantoja and Lee, 1999; Pantoja et al., 1997; Pantoja et al.,
2009; Roth and Harvey, 2006). A study by Pantoja and Lee (1999) found that an 8-amino-acid peptide composed of only alanine was hydrolysed gradually until it was undetectable, with the shorter peptides detected over several hours. This study used a fluorescent tag in order to sensitively detect the peptides. Little is known of the effects this has on this stepwise hydrolysis mechanism. Additionally, since alanine was the only amino acid in the peptide sequence, it is unclear whether stepwise hydrolysis occurs for other amino acids for a larger peptide (i.e., more than three residues). As part of this study I am interested in investigating peptide length effects further without interference of a fluorescent tag and with inclusion of amino acids other than alanine. I therefore, included AAFRV into my study to assess whether there are significant differences with the addition of one amino acid to AFRV. Fig. 36 displays the concentration plots of AAFRV and AFRV removal from ER water. Fig. 36a is the plot of AFRV and AAFRV when incubated together in the same bottle. Both peptides are removed from the water within 24 hours, this is faster than that observed for previous studies and is probably due to the fact that these incubations are conducted during the summer (Liu et al., 2010). The removal rates are 4.53 d⁻¹ for AFRV and 5.27 d⁻¹ for AAFRV indicating that the peptides are being removed from ER water at a similar rate. AFRV increased in concentration in the water in the first 12 hours, leading to a low R² value for its rate. The removal of AFRV from the water (data not shown) follows a consistent trend, similar to AAFRV in this incubation. The reason for this increase in AFRV concentration when examining the dual incubation of AAFRV and AFRV is likely due to the microbial removal of A from AAFRV. I therefore, examined the single incubation of AAFRV (Fig. 36b) using water collected from ER in August. Because the rate of removal was precipitous and this would likely mask the formation of an intermediate product, I decided to use filtered ER
water (0.2 μm filter size) for degradation experiments. Filtration served to remove most of the microbial populations to slow the degradation. The result from the unfiltered August water experiment is similar to the July incubations conducted with unfiltered water; AAFRV is removed rapidly within 18 hours. When examining the chromatogram in detail for the August experiment in which filtered water was used, an additional peak is observed. This peak, identified as AFRV, increased in intensity throughout the incubation, (measured as percent of maximum area). Thus, AFRV is produced during the removal of AAFRV from filtered ER water collected in August. This explains why the concentration of AFRV increases when incubated together with AAFRV in the July experiments.

![Fig. 36. Plot of concentration against time for the incubation of AFRV and AAFRV in ER water. Displaying, (a) the dual incubation of AFRV and AAFRV together and (b) AAFRV single incubation, with 0.2 μm filtered, and unfiltered ER water. The data labelled as “AFRV in filt. ER” in (b) is plotted as % maximum against time.]

There is a striking difference between the rates of removal of AAFRV in unfiltered and filtered ER water, 13.2 d⁻¹ and 3.50 d⁻¹ respectively. AAFRV is removed
from ER water approximately ten times faster in unfiltered ER water than filtered. This indicates that organisms larger than 0.2 μm in size are primarily responsible for the conversion of AAFRV to AFRV followed by its final removal, this would include generally phytoplankton but also most bacteria (Wetzel, 2001). It is apparent that organisms smaller than 0.2 μm such as viruses and microbacteria, are capable of removing AAFRV while converting it to AFRV but this process is slow in comparison to experiments conducted with no filtration. Interestingly, no AFRV was detected in the unfiltered ER water experiment with AAFRV. AAFRV has a molecular weight of 562 Da, and AFRV has a molecular weight of 491 Da. It has been suggested in previous studies that 600 Da represents the cutoff size for direct uptake by cells (Nikaido and Vaara, 1985). The 600 Da cutoff for uptake is approximate only and since the molecular weight of AAFRV is close to this size the results of the AAFRV and AFRV incubation could be evidence of two different processes removing these two peptides from ER water. AAFRV could be removed by initial extracellular hydrolysis removing the first alanine, to produce AFRV. AFRV could then be directly taken up the microbes. These results complement those found for longer peptides composed of alanine and suggests that perhaps alanine is preferentially removed due to its smaller size (compared to valine) (Pantoja and Lee, 1999). A larger set of peptides is required in order to confirm this hypothesis which is subject for future work.

6.3.3 Competition for peptide removal in natural water

In natural waters there is a continuous cycle of production and removal of peptides. It has been suggested in previous studies that certain amino acids, in particular, those that are polar and charged, are more labile than others (Liu et al., 2010; Roth and
Harvey, 2006). I wanted to answer the question of whether the amino acid sequence of a peptide has any impact on the peptide removal rate, and if so is its removal from natural water samples in competition with other, more labile peptides. As part of this study I synthesised four peptides with different sequences, consequently with varying hydrophilicities, as defined by Hopp and Woods (1981). Hydrophilicity varies from -1.2 to +1.4 for the four peptides in the increasing order of: AVFA → AFRV → AAFRV → AREG (Hopp and Woods, 1981). Fig. 37 displays three different concentration plots of the four peptides when incubated singularly in separate bottles and when incubated together in water collected from ER. The plot of the four peptides incubated separately shows that AAFRV is removed much more rapidly (13.2 d⁻¹) than the other three (5.02 – 8.64 d⁻¹). This could be due to the effect of the two-stage removal via AFRV and that removal of the A from AAFRV is very rapid, as suggested in the previous section. AFRV seems to be removed slightly faster (8.64 d⁻¹) than AVFA (6.36 d⁻¹) and AREG (5.03 d⁻¹). Despite their different chemical characteristics (i.e., hydrophilicity) there is minimal difference in the removal rates of the three tetrapeptides when incubated singly.
A more realistic situation for a natural water system is when peptides are incubated together in the same bottle. In Fig. 37 I display the concentration plot of all the peptides incubated together (labelled quad incubations ER). Unfortunately, due to the fact that AREG has a very similar retention time to that of the dissolved organic matter (DOM) present in the water I was unable to detect AREG after the initial addition to the water in this incubation bottle. The removal rates for the three peptides are 14.9 d\(^{-1}\) (AAFRV), 7.90 d\(^{-1}\) (AFRV), and 32.0 d\(^{-1}\) (AVFA). The concentration of AVFA is significantly higher than AAFRV and AFRV at the beginning of the incubation experiment. AFRV has the slowest removal rate of the three peptides, and this is probably due to the fact that AAFRV increases the concentration of AFRV upon its removal. Single and quad incubations of the three peptides are similar suggesting that...
there is little competition regarding removal from ER water, either because they are removed by different mechanisms or the removal mechanism is not limited. Liu et al. (2010) saw a similar effect for AVFA, AFRV and AAVF, suggesting that competition between peptides of similar length does not occur with respect to their removal from ER water.

To assess whether peptide non-competitive removal in ER water also occurs in another water system such as MLB water I compared the single incubation of AVFA to doubly incubated AFRV and AVFA, as shown in Fig. 37. The limited amounts of peptides precluded the examination of a larger set of experiments with more peptides. The general pattern of concentration change of AVFA from MLB water is very similar to that observed for ER water for AVFA single and dual incubations. Comparing AFRV and AVFA in their dual incubation, their plots mirror each other almost exactly, with the exception of their initial concentrations. The removal rates for AFRV and AVFA in these incubations are almost identical: AVFA single is 1.10 d⁻¹, dual is 1.33 d⁻¹ and AFRV dual is 1.67 d⁻¹. Similar to when present in ER water, MLB water does not appear to have any competition to remove peptides from the system, likely due to the same reasons presented above for ER water. However, there is a difference in removal rates in ER and MLB waters for the same peptides, e.g., single AVFA incubation in ER and MLB water the rates are 6.36 d⁻¹ and 1.10 d⁻¹ respectively. In the next section I will investigate possible reasons for ER and MLB peptide removal rate differences.

6.3.4 Comparison of removal rates of peptides in water from two locations: ER and MLB

As mentioned above I found a significant difference in the peptide removal rates incubated in water from the two locations. Using AVFA single incubations as an
example, the removal rate for AVFA is $6.36 \text{ d}^{-1}$ and $1.10 \text{ d}^{-1}$ in MLB and ER, respectively, a 6-fold difference in removal rate. One of the reasons for this difference could be due to the fact that MLB water was frozen after collection and then defrosted to be utilised for these biodegradation experiments. In order to evaluate whether this is the reason for the removal rate difference in water from the two locations, I froze ER water upon its collection to simulate the conditions under which biodegradation experiments were conducted for MLB water. Two weeks after freezing the ER water the frozen water was thawed and used for a separate incubation of AVFA. The removal rate of AVFA in frozen ER water is listed in Table 9 as $38.6 \text{ d}^{-1}$. This rate is much faster than the rate of removal of AVFA in unfrozen ER water and is probably because the water was not immediately frozen upon collection but after storage at $4 ^\circ C$ for 4 days. Unfortunately, it was not anticipated that this experiment would be required and it only became clear upon analysis of the samples by HPLC that there was a significant difference in removal rate between ER and MLB water. At this stage (i.e., after 4 days) the remaining ER water was frozen, however, the storage time at $4 ^\circ C$ clearly had a substantial effect on the removal rate of AVFA. The delay in freezing the ER water meant that there was probably an increase in micro-organism activity due to “bottle effects” (Crompton, 2006). This explains why I observed such a large disparity between the unfrozen and the frozen ER water AVFA removal rates. However, despite this difference it is clear that freezing of samples for this experiment does not reduce the ability of the micro-organism community to remove peptides from the water. I suggest that the reason for the different removal rates of AVFA in ER and MLB water is due to the indigenous chemistries of each water sample and will attempt to shed light on this. The difference in peptide removal rates for ER and MLB collected water is similar for all peptides (Table 9).
One of the reasons for the difference in removal rates in collected ER and MLB water could be due to the varying composition of OM in the water samples. If the peptides are being removed from the water by association or reaction with OM already present in the water, then the molecular composition of the DOM could explain why there are striking differences in the removal rates of peptides in ER and MLB water.

Advanced analytical techniques such as FT-ICR-MS, can investigate the molecular level characteristics of ER and MLB DOM. The resulting MS spectra are displayed in Fig. 38. The whole spectra are displayed and insets show the expanded mass region, 353 Da. The full mass range spectra show that a large number of compounds are ionised in these two samples. Expanded MS are often more informative than the full spectra because peaks assigned to specific elemental formulae can readily be compared; Fig. 38 demonstrates this. Exactly 15 (MLB) and 29 (ER) different assigned elemental formulae (labelled a – u) are observed in this expanded region which displays only 1 Dalton.

It should be noted that several more than the 15 peaks could be identified for MLB, however, their signal to noise ratio was below the threshold of 4 and so are not included in the m/z lists used to determine the elemental formulae. ER and MLB DOM spectra have many of the same assigned formulae, although they could have very different structural compositions; it is not possible to determine the structure of an ion from the formula alone. Most of the formulae are CHO containing, which is expected since these formulae charge most easily in an ESI source and the dominant components of CHO compounds (fatty acids) are ubiquitous in DOM (Sleigher et al., 2008). The differences between the two expanded mass range spectra are peaks a, b and c. Peaks a and b are observed only in the ER DOM spectrum, with formulae of \([C_{15}H_{13}O_{10}]^+\) and
[C$_{17}$H$_{13}$O$_7$], respectively, whereas peak c ([C$_{15}$H$_{17}$N$_2$O$_6$S]) is observed only in the MLB DOM spectrum. Peaks a and b are low mass defect, oxygen-rich formulae, most likely derived from tannins/lignin. Peak c cannot be readily assigned to a particular structural source, since it is not peptidic (a peptide sequence cannot be assigned) and the presence of N suggests it is not derived from lignin, its structural source therefore, remains unknown.

Although examining specific mass regions is useful, it is difficult to assess the general nature of the OM, in particular to seek reasons for removal rate differences of peptides in water from these two locations. Several diagrams have been used in past studies of OM to facilitate recognition of differences between samples (Hughey et al., 2001; Kendrick, 1963; Kim et al., 2003a; Sleighter and Hatcher, 2007). I will use one such diagram that has successfully been used in past studies, the van Krevelen diagram.
Fig. 38. Mass spectra of DOM from ER and MLB water PPL extracted and analysed by negative ionisation mode in FT-ICR-MS. Expanded spectra of m/z 353 are displayed with whole spectra inset. Identified peaks (a – u) formulae are listed in the table.
The van Krevelen diagram has been used for examination of the general characteristics of OM samples in many studies (Kim et al., 2003a; Sleighter and Hatcher, 2008). The organisation of structural types into general regions facilitates a rapid comparison between samples. The van Krevelen diagram therefore, is used here in the comparison of DOM from ER and MLB waters so that I am able to deduce possible reasons for the removal rate differences of peptides incubated in these waters (Fig. 39).

The van Krevelen diagram is divided into 7 regions, and is indicated on the Figure. There seems to be little difference in the lipid, protein/peptide and aminosugar characteristics of these two samples. The peptide/protein area similarity is perhaps a little surprising considering the differences in removal rates of the peptides in ER and MLB water in this study. This suggests that despite the difference in removal rates, the resulting formulae assigned to peptides/proteins are ultimately similar. Of course it is possible that the structures of the formulae that fall in the protein/peptide region are very different, however, as stated above this technique is not able to differentiate structures without further analysis. There seems to be a small difference in the cellulose region.

The most significant difference between the van Krevelen diagrams of these two samples is in the lignin, tannin, and condensed hydrocarbon regions. Tannins are produced only by higher plants, many of which are the source of DOM in ER water which drains the Dismal Swamp (Helms et al., 2008; Minor et al., 2007; Sleighter and Hatcher, 2008). Similarly, lignin is a highly abundant component of higher plants (Killops and Killops, 2005; Stenson et al., 2003) In contrast to DOM from ER where a significant source of the water is from higher plants, MLB water DOM is primarily algal derived (Hatcher et al., 1982). Condensed hydrocarbons are generally associated with anthropogenic inputs and fire related natural events (Hammes et al., 2007). In addition, ER is situated in a
densely populated area with known anthropogenic sources and fire events likely producing condensed hydrocarbons (Shafer et al., 2004). Formulae consistent with these condensed hydrocarbons are observed in Fig. 39.

Fig. 39. van Krevelen diagrams plotting H/C against O/C of DOM from ER and MLB waters analysed by negative ionisation FT-ICR-MS. There are regions that are divided on the plots and these are labelled I – VII. The names of these regions are listed below the Figure.

The association of peptides through adduction to existing OM has been suggested to occur with various structural components including lignin and tannins (Rillig et al., 2007). Because Fig. 39 indicates that lignin- and tannin-like molecular formulae are present in much higher amounts in DOM from ER water than MLB water, it is logical to presume that this difference could explain the differences observed in removal rates of peptides between the two sites. MLB water DOM contains few of the tannin and lignin-based structural components as determined by FT-ICR-MS analysis leading me to believe that there are multiple mechanisms occurring to remove peptides from these waters. One of which is common to both locations and is probably extracellular hydrolysis followed
by direct uptake. The observation that AAFRV is converted to AFRV, as discussed above, supports this process. Another is association with lignin- and tannin-like substances which bind through Michael reaction chemistry to form adducts that modify peptide removal rates. Although there is no observed change in retention time for the peptides observed by HPLC in ER water incubations, it is possible that following the association with lignin- and tannin-like structures the peptides could not be detected due to such a significant change in their chemical characteristics, resulting in an apparent reduction of concentration. This hypothesis is unresolved here but will be discussed in Chapter 7. Examination of some of the bulk measurements of ER and MLB waters may assist in other explanations for removal rate differences.

The measurements recorded when the water samples were collected are displayed in Table 8. There is a marked difference in the temperatures of the water, 28.0 °C for ER water and 22.3 °C for MLB water. As discussed in the previous section (6.3.1) the removal rate of AVFA in ER water varies according to the season. When water temperatures are lower it takes longer for AVFA to be removed from the water. The removal rate differences of peptides in ER and MLB waters could be simply due to the fact that they were collected during different seasons (ER-summer, MLB-fall/winter). However, even during the winter experiments in ER water, the removal rates of peptides are still faster than in MLB water, indicating that the temperature of the water at the time of collection is not the only factor governing the removal rates of peptides in these two locations.

The reason why temperature of the water could have significant impact on the removal rate of peptides is that the micro-organism community varies in type and number according to the time of year (Kuznetsova and Lee, 2001). Unfortunately, I did not
investigate the types of bacteria and phytoplankton in this study however, the concentration of bacteria in each water sample was measured (Table 8). The concentration of bacteria in the water collected from the two locations is very similar, suggesting that the type of organism present in the water is a more important factor for determining removal rate differences of peptides than simply the number of bacteria present.

Another reason for removal rate differences in ER and MLB water could be salinity differences. At the time of collection ER water had a salinity of 23 and MLB water was 32. This is a substantial difference, however, Liu et al. (2010) found that for water samples from an upper James River to lower Chesapeake Bay transect the salinity ranges from 0 to 32 and there was no difference for any of the transect samples in the removal rate of AVFA, suggesting that salinity is not a major factor governing removal of AVFA and likely not for the other tetrapeptides and pentapeptide in this study.

The concentration of OM in the water collected is also different for the two locations. TOC concentration which includes DOC and particulate organic carbon is similar for the two sites; 433 µM for ER water and 490 µM for MLB water. A larger difference is observed in the DOC concentration values of 340 µM for ER water and 432 µM for MLB water. This means that ER water has more particulate OM than the MLB water collected. Some of the particulate OM could be from organisms that are removing the peptides in the water from the two locations, however, the concentration of bacteria in the two samples is similar and so if the particulates are composed of living matter, they are primarily larger organisms. Being that ER water has more particulate OM, and that larger organisms could be responsible for their removal, this would explain why ER water has a faster removal rate of peptides than in MLB water. DN is much lower than
TN for water collected from ER than MLB, following the same pattern as for C. N is in general at a higher concentration in water collected from MLB than in ER water possibly meaning that MLB N is more bioavailable. Consequently, since N is commonly a limiting nutrient in water systems (Vitousek and Howarth, 1991; Zehr and Ward, 2002) and ER collected water has a lower amount of N, the introduction of a new source of N, such as the peptides in this study, will result in faster peptide removal rates in MLB water than ER water.

An additional reason for the difference in the removal rates of peptides in MLB and ER water is through adsorption of the peptides to mineral matter (Gordon and Millero, 1985; Mayer, 1994a; Mayer, 1994b). Since ER water contains a higher quantity of inorganic matter than MLB water, it is possible that the newly introduced peptides could become sorbed to these structures, serving to reduce their measured concentration. It is clear, however, that there are many possible reasons for the differences in removal rates of peptides in water collected from ER and MLB. It is probable that several of these are important, producing the large difference in peptide removal rate in these two locations, e.g., association with tannins, lignin and inorganic matter and extracellular hydrolysis. However, I am unable to discern which is the most likely at this stage. Future studies of peptide removal rates and mechanisms from natural waters will include more sites and a more comprehensive study of location differences so that the mechanism can be determined.

6.4 Conclusions

Despite their ubiquitous presence in natural waters, the reasons why some
peptides are preserved are not well understood. This study sought to understand some of these reasons by measuring the removal rate of synthesised peptides in natural water designed to examine impacts of peptide chemical structure, length, and water location. Four peptides were used in this study derived from the ubiquitous protein, RuBisCO and incubated in collected ER and MLB water. It appears that chemical structure has little impact on the removal rate of tetrapeptides in collected ER and MLB water as determined by the peptides studied here, further studies are required to assess the wider applicability of these results. However, length of peptide appears to be a key factor, especially when close to 600 Da, thought to be the limit for direct uptake by bacteria (Nikaido and Vaara, 1985). AAFRV, close to the 600 Da limit, is not removed directly but instead, one alanine is removed producing AFRV, before complete removal from the water. Results from comparisons of the two locations are that ER water has more rapid removal of peptides than in MLB water. Although there are many possible reasons why this occurs, I am unable to suggest which is/are the most important; bacteria/phytoplankton type and adduction to aromatic species (e.g., quinones) seems the most probable. This study uncovers several interesting reasons for peptide preservation in brackish waters. This information can be used as the basis for several additional studies to assist in understanding some of the complex mechanisms that are occurring in natural waters upon peptide release.
CHAPTER 7

EFFECT OF MICHAEL ADDUCTION ON PEPTIDE PRESERVATION IN NATURAL WATERS

7.1 Introduction

Proteins are key structures of many organisms and as a consequence they can be a major fraction of the dissolved organic nitrogen (DON) pool in marine and fresh water systems. Most proteins and peptides released into aqueous systems are rapidly removed mainly through hydrolysis, however, a small fraction is preserved in the DON and sediment ON (SON) pools (Fogel and Tuross, 1999; Herbert, 1999; Keil and Kirchman, 1994; Nguyen and Harvey, 1997; Nguyen and Harvey, 2001; Roth and Harvey, 2006; Tanoue et al., 1996; Tanoue et al., 1995; Tappin et al., 2007; Zang et al., 2001). The reasons why these proteins and peptides are not immediately hydrolysed are not well understood. A number of hypotheses have been suggested in past studies and these include: hydrophobic and non-covalent interactions (Nguyen and Harvey, 2001), selective preservation (Hatcher et al., 1983), protection by encapsulation into organic matter (OM) (Knicker and Hatcher, 1997) and mineral matter (Mayer, 1994a;

This text is being prepared for submission to Marine Chemistry.

Mayer, 1994b), chemical structure (Liu et al., 2010; McKee et al., 2011) and formation of
adducts to OM (Hsu and Hatcher, 2006; Kirchman et al., 1989; Liu et al., 2010; Rillig et
al., 2007). The adduction to quinones in aqueous systems is one of the proposed
mechanisms for preservation of peptides in OM systems (Hsu and Hatcher, 2006) and
will be investigated in more detail in this study. In most environmental systems
conducive to the likely formation of such adducts, it is apparent that the likely adduct
receptors are quinones derived from the oxidative degradation of polycyclic aromatic
hydrocarbons (PAHs).

PAHs are common anthropogenic and natural contaminants released into the
aqueous environment (Bolton et al., 2000; Samanta et al., 2002). There have been
several studies suggesting that PAHs form adducts with OM, serving to reduce PAH
bioavailability as well as that of the OM (Accardi-Dey and Gschwend, 2002; Burgos and
Pisutpaisal, 2006; Means et al., 1982; Novoszad et al., 2005; Parris, 1980). Peptides are
one type of OM that has been successfully linked to several adducts of PAHs, most
notably quinones (Briggs et al., 2003; Hsu and Hatcher, 2005; Tandon and Maurya,
2009). The Michael reaction is a common reaction that occurs between quinones and
primary amines (Bittner et al., 2002). Hsu and Hatcher (2005) demonstrated that the
peptide glycine-glycine-glycine-arginine reacted with 1, 4-naphthoquinone (Nap) via a
Michael reaction and suggested that these types of adducts can also form with humic
acids and can thereby possibly serve to protect the peptide from degradation. However,
there has been no direct evidence to date that this reaction affects the remineralisation of
a peptide once bound to a quinone.

In this study I investigated the ability of Michael adduction to impact the removal
of tetrapeptides from natural water using a model quinone: Nap. In order to assess
whether adduction affects the preservation of peptides I will synthesise a series of peptide adducts and compare their removal in collected water. Water from two sites is used in this study, the first is Mangrove Lake, Bermuda (MLB) and the second is Elizabeth River, VA (ER). The large difference in aqueous environment, i.e., lacustrine versus riverine, enables wider applicability of the findings in this study to other sites, but lack of time does not permit a more extensive examination of all environmental situations that would impact peptide preservation/degradation. In addition, a previous study (Chapter 6) has revealed that there are some notable differences in the rate at which peptides are removed from water collected at these two sites. The rates of removal of specific peptides in the riverine system are about 3 – 6 times faster than those from the lacustrine system.

In the current study, three different peptides were synthesised with a quinone attached covalently, including alanine-valine-phenylalanine-alanine (AVFA), 1, 4-naphthoquinon-2-yl-alanine-valine-phenylalanine-alanine (Nap-AVFA), alanine-arginine-glutamic acid-glycine (AREG), 1, 4-naphthoquinon-2-yl-alanine-glutamic acid-glycine (Nap-AREG), alanine-phenylalanine-arginine-valine (AFRV) and 1, 4-naphthoquinon-2-yl-alanine- phenylalanine-arginine-valine (Nap-AFRV). These Nap adducts are then added to each of the two natural water samples and incubated over a period of 4 weeks. The temporal concentrations of these three peptide adducts are monitored by high performance liquid chromatography (HPLC) directly and compared to similar experiments with the un-adducted peptides to assess whether adduction to Nap has any effect on their rate of removal from the water.
7.2 Materials and Methods

7.2.1 Preparation of peptide adduct standards

For all syntheses the highest quality reactants available were used. Fluorenylmethyloxy carbonyl chloride (FMOC) amino acids were purchased from Protein Technologies. Resins were purchased from Advanced ChemTech (alanine, glycine) and Nova Biochem (valine).

Nap adducts to peptides were produced by initially forming an adduct with alanine, synthesising 1, 4-naphthoquinon-2-yl-alanine (Nap-A). I used a method based upon that of Bittner et al. (2002), which describes a synthetic pathway for production of a related amino acid adduct, 1, 4-naphthoquinon-2-yl-glycine. The method is modified as follows: approximately 20 mmol of unprotected A (Acros) was dissolved in water (MilliQ) and added to a hot 40 mmol ethanol (Fisher) solution of Nap (Acros). This solution was stirred at room temperature for 24 – 72 hours, after which the solution was removed by filtration, and the solid brought to dryness by rotor evaporation. Purification of Nap-A used a silica gel column (grade 62, 60 – 200 mesh, 150 Å, Sigma-Aldrich) eluting the product with a 9:1 v/v mixture of methylene chloride (Acros)/methanol (Fisher) with average yields of 4 %.

Attachment of Nap-A to other amino acids to yield the peptide adduct was facilitated by the use of FMOC solid phase peptide synthesis (SPPS) on an automated solid phase peptide synthesiser (PS3, Protein Technologies, AZ) to produce all of the peptides and peptide adducts used in this study; a method that reproducibly synthesises peptides of high purity (Chan and White, 2000; King et al., 1990; Merrifield, 1963; Stewart, 1976). The details of the synthesis are described elsewhere (Chapter 6, Liu et
al., 2010). All of the protocol is the same for synthesis of the Nap adducted peptides as for the un-adducted peptides except the following steps. Following stepwise addition of the amino acids to the resin the last step involves the addition of the purified Nap-A, without any prior deprotection step. The workup procedure was exactly the same as for the un-adducted peptides, except that the last step was omitted because of low solubility. Thus, the Nap-adducted peptides and un-adducted peptides were purified in chloroform by HPLC directly using a Shimadzu HPLC with a C_{18} (Grace, Apollo 5 μ, 150 x 10 mm) column attached to a photodiode array (PDA) detector equipped with a fraction collector. Mobile phases were 0.05 M NaH_{2}PO_{4} (Fisher) which was kept constant while the organic phase, methanol, was ramped on a gradient. Due to their unique characteristics, the gradient and flow rate was optimised for each peptide. Pure peptides and adducts were collected using the fraction collector and the structure of each was confirmed by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) and nuclear magnetic resonance (NMR) analysis (see Appendix B). Peptide and adduct yields were generally ~ 50 %.

For FT-ICR-MS analysis the relevant collected peaks were directly analysed on a Bruker Daltonics Tesla Apex Qe 175 FT-ICR-MS at the College of Sciences major instrument cluster (COSMIC) facility at Old Dominion University (ODU). The collected fractions were acidified to pH 6 using formic acid (Sigma-Aldrich) and directly infused into the Apollo II electrospray (ESI) ion source at a rate of 120 μL/hr using a syringe pump. Polyethylene glycol was used as an external calibrant and was accumulated for 1.0 sec in the hexapole before being transferred to the ICR cell and 20 transients were co-added. The collected fractions were analysed using a 1.0 sec accumulation time in the hexapole with 10 co-added transients. A 4 MWord time domain was used for the free
induction decay (FID). Zero-filling was completed once on the summed FID signal and then Sine-Bell apodised before Fourier transformation followed by magnitude calculation using the Bruker Daltonics Data Analysis software v. 1.3™. Resulting mass/charge (m/z) data for each peptide and peptide adduct were confirmed as being correct by comparing to their exact mass.

For \(^1\text{H} \) NMR analysis, a 400 MHz Bruker Avance spectrometer (Bruker Biospin, Inc.) located at the COSMIC facility was utilised. The pulse programme included water suppression for optimal spectral quality. Spectra were matched to simulated ones generated by a \(^1\text{H} \) NMR predictor (ACD Labs v. 9.0).

7.2.2 Sampling locations

Two different water samples were used in this study, ER and MLB. ER is a major tributary of the Chesapeake Bay, surrounded by large numbers of industrial companies on either side of the river, and its sources include groundwater from the Columbia and Yorktown aquifers and the Great Dismal Swamp (Charette and Buesseler, 2004; Helms et al., 2008; Minor et al., 2007; Shafer et al., 2004). MLB is a brackish lake, with its primary source of OM originating from algae and its source of water is through permeation of the bedrock and rainfall (Hatcher, 1978; Hatcher et al., 1982; McKee and Hatcher, 2010). It is located in a site distant from any anthropogenic sources, except a nearby golf course. The full details of the water collected from these two locations is described in Chapter 6 and are not repeated here.

7.2.3 Incubation of peptides and adducts

The setup and conditions of the incubations of the peptides and adducts discussed
in this study are similar to those of Chapter 6. Exceptions to these conditions are that, during preparation of the Nap-adducts, dissolution of stock solutions, due to the limited solubility, was assisted by use of sonication for 1 – 2 hours. Since Nap peptide incubations have not been previously shown to be reproducible, two of the adducts were incubated in two separate bottles (Nap-AFRV and Nap-AVFA). The reproducibility of chromatograms and resulting calculated concentrations suggest that Nap adduct experiments are reproducible, as will be shown below.

Abiotic control experiments were also prepared using mercuric chloride (HgCl₂), an effective treatment that prevents microbial activity (Lee et al., 1992; Trevors, 1996; Tuominen et al., 1994). The previous studies of Liu et al. (2010) demonstrated that HgCl₂ effectively quenched all microbial activity over their two week experiment using a concentration of 180 μM. In this study I did not repeat their control on experiments involving the degradation of individual peptides but confirmed that the activity of bacteria administered Nap adducted peptides can also be quenched with HgCl₂. In addition each peptide and adduct was incubated for four weeks in MilliQ water alone, and time points were taken and monitored by HPLC. Results showed that all peptides and adducts were stable throughout the time period, with no change in their concentration and no appearance of additional peaks in HPLC chromatograms.

7.2.4 Monitoring of peptide concentrations by HPLC

Peptide and Nap-peptide concentrations were monitored by HPLC during incubations. Full details of this are described in Chapter 6. The wavelengths monitored by UV-Vis were 210 nm for the un-adducted peptides and 264 nm for the Nap-peptides. Duplicate runs resulted in high levels of agreement (peaks were reproduced within 5 %).
Newly formed peaks during Nap peptide incubations were collected using the fraction collector attached to the HPLC. Single incubations for the three Nap peptides were used for collection at two different time points, (12 and 168 hours). Due to the need for use of a high volatility buffer in FT-ICR-MS analysis, the aqueous solvent was changed to sodium bicarbonate (0.5 M) with no changes to the original gradients. Although retention times varied slightly among the different experiments; comparison to standards was sufficient to deduce the identity of each peak. Peaks were collected during a single run of each time point (12 and 168 hours) for the three Nap adducts and analysed by FT-ICR-MS as described above. For tandem MS analysis of the peaks, collision induced dissociation (CID) was employed on a m/z range of 1 Da centred on the mass spectral peak of interest. Collision voltage was altered until optimal fragmentation occurred. Exactly 10 co-added transients were accumulated in the ICR cell for each successfully fragmented peak.

7.2.5 Rates of removal calculation

Due to the fact that there was no lag time for any of the incubations in this study first order kinetics was applied to calculation of the removal rates for each peptide and adduct.

7.3 Results and Discussion

The previous chapter (Chapter 6) shows that a significant difference exists in the rates of removal of peptides between water samples collected in the ER and those from MLB. While there could be many explanations for the observed differences in rates of
removal, one that appears likely is the fact that ER contains substances in the dissolved organic carbon that are possible excellent Michael receptors. This can presumably enhance the formation of Michael adducts with peptides leading to a change in the observed rate of removal. The discussions presented below focus on experiments conducted in ER water however, although the changes occur at different rates in MLB water, the overall results are the same.

7.3.1 Incubation of peptides and peptides adducts

The HPLC chromatograms for selected time points for one of the peptides (AFRV) and adduct (Nap-AFRV) is shown in Fig. 40. AFRV and Nap-AFRV are incubated singly and together in one bottle (dual), to compare if there are any differences in removal for AFRV and Nap-AFRV in the incubations. In the case of AFRV the removal rates for the single and dual incubations are 8.64 and 13.1 d\(^{-1}\) respectively (Table 10). Although their rates differ, the chromatograms for the two incubations for AFRV are similar. After 24 hours of incubation, AFRV is greatly reduced in concentration in both incubations and after 48 hours I am unable to detect it, implying that the reason for rate differences of single and dual incubations is simply due to differences in initial concentrations. This similarity in removal rates suggests that AFRV removal in collected ER water is not significantly affected by the presence of Nap-AFRV.
Table 10

Information regarding the incubations of peptide and peptide adducts in collected ER and MLB water analysed by HPLC. Peptides/adducts are listed in the first column, the location of the water sample is in the second, the details of the incubation bottle are in the third column (+ means incubated together), followed by the initial concentration of the peptide in the water sample, the calculated removal rate using first order kinetics and finally the regression ($R^2$) of the calculated rate.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Location</th>
<th>Incubation</th>
<th>Initial Conc (μM)</th>
<th>Rate (d⁻¹)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVFA</td>
<td>ML</td>
<td>Single</td>
<td>10.5</td>
<td>1.10</td>
<td>0.96</td>
</tr>
<tr>
<td>AVFA</td>
<td>ER</td>
<td>Single</td>
<td>10.7</td>
<td>6.36</td>
<td>0.93</td>
</tr>
<tr>
<td>AFRV</td>
<td>ER</td>
<td>Single</td>
<td>12.9</td>
<td>8.64</td>
<td>1.00</td>
</tr>
<tr>
<td>AREG</td>
<td>ER</td>
<td>Single</td>
<td>10.2</td>
<td>5.03</td>
<td>0.75</td>
</tr>
<tr>
<td>Nap-AVFA</td>
<td>ER</td>
<td>Single-Repeat 1</td>
<td>8.8</td>
<td>8.97</td>
<td>0.91</td>
</tr>
<tr>
<td>Nap-AVFA</td>
<td>ER</td>
<td>Single-Repeat 2</td>
<td>11.7</td>
<td>11.9</td>
<td>0.99</td>
</tr>
<tr>
<td>Nap-AFRV</td>
<td>ER</td>
<td>Single-Repeat 1</td>
<td>10.1</td>
<td>10.4</td>
<td>0.94</td>
</tr>
<tr>
<td>Nap-AFRV</td>
<td>ER</td>
<td>Single-Repeat 2</td>
<td>11.9</td>
<td>12.1</td>
<td>0.96</td>
</tr>
<tr>
<td>Nap-AREG</td>
<td>ER</td>
<td>Single</td>
<td>19.5</td>
<td>14.8</td>
<td>0.90</td>
</tr>
<tr>
<td>AVFA</td>
<td>ER</td>
<td>Dual, AVFA + Nap-AVFA</td>
<td>9.4</td>
<td>6.62</td>
<td>0.91</td>
</tr>
<tr>
<td>Nap-AVFA</td>
<td>ER</td>
<td>Dual, AVFA + Nap-AVFA</td>
<td>9.7</td>
<td>6.86</td>
<td>0.87</td>
</tr>
<tr>
<td>AVFA</td>
<td>ML</td>
<td>Dual, AVFA + Nap-AVFA</td>
<td>12.7</td>
<td>2.65</td>
<td>0.98</td>
</tr>
<tr>
<td>Nap-AVFA</td>
<td>ML</td>
<td>Dual, AVFA + Nap-AVFA</td>
<td>11.6</td>
<td>2.57</td>
<td>0.97</td>
</tr>
<tr>
<td>AFRV</td>
<td>ER</td>
<td>Dual, AFRV + Nap-AFRV</td>
<td>13.0</td>
<td>13.1</td>
<td>0.89</td>
</tr>
<tr>
<td>Nap-AFRV</td>
<td>ER</td>
<td>Dual, AFRV + Nap-AFRV</td>
<td>14.8</td>
<td>11.3</td>
<td>0.88</td>
</tr>
<tr>
<td>AFRV</td>
<td>ML</td>
<td>Dual, AFRV + Nap-AFRV</td>
<td>10.5</td>
<td>2.04</td>
<td>0.98</td>
</tr>
<tr>
<td>Nap-AFRV</td>
<td>ML</td>
<td>Dual, AFRV + Nap-AFRV</td>
<td>10.3</td>
<td>2.10</td>
<td>0.98</td>
</tr>
<tr>
<td>AREG</td>
<td>ER</td>
<td>Dual, AREG + Nap-AREG</td>
<td>8.7</td>
<td>8.96</td>
<td>0.88</td>
</tr>
<tr>
<td>Nap-AREG</td>
<td>ER</td>
<td>Dual, AREG + Nap-AREG</td>
<td>8.8</td>
<td>6.74</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Fig. 40. HPLC chromatograms of AFRV and Nap-AFRV incubations in collected ER water. AFRV and Nap-AFRV were incubated singly in their own individual bottle and together (dual). Two different wavelengths are displayed, 210 nm for the AFRV chromatograms and 264 nm for the Nap-AFRV chromatograms. The three peaks labelled on Nap-AFRV incubation chromatograms are: peak 3 is Nap-AFRV, peaks 1 and 2 are two additional peaks observed during the experiment.

Nap-AFRV chromatograms, however, are significantly different from those of AFRV. For Nap-AFRV in the dual incubation, three peaks are observed: peak 3 is Nap-AFRV and peaks 1 and 2 are additional new peaks not previously observed in standards incubated in MilliQ water. At the start, 0 hours, Nap-AFRV is the dominant peak in the chromatogram, with only small amounts of peak 2 observable. This new peak must form
immediately upon addition of the Nap-AFRV to the water samples to be incubated. After 24 hours, Nap-AFRV is significantly removed, peak 2 increases in intensity, and another new peak, peak 1, grows in. After 48 hours only peak 1 is detectable, it remains after 168 hours of incubation, and is finally removed from the water by 500 hours (chromatogram not shown). The chromatograms representing singly incubated Nap-AFRV are almost identical to those of Nap-AFRV in the dual incubation experiment, suggesting that the presence of the AFRV adduct is required in the incubation bottle to produce the additional two peaks, 1 and 2. Moreover, the removal rates for each of the replicate experiments for Nap-AFRV single incubations and the dual incubations are very similar, demonstrating that this is a reproducible process (Table 10). The evolution of peaks 1 and 2 is also reproducible. A control experiment in which Nap-AFRV is incubated in MilliQ water for the same length of time as ER water incubations shows that the Nap-AFRV concentration remains constant and no additional peaks are observed (data not shown). This proves that ER water is responsible for the production of peaks 1 and 2. The same experiment conducted on MLB water shows the appearance of the same two peaks, albeit at a longer time interval characteristic for all MLB incubations.

The similarity of Nap-AFRV incubations in ER and MLB collected waters, including the appearance of two additional peaks as the concentration of Nap-AFRV rapidly diminishes, suggests that a similar process of Nap-AFRV removal is occurring in the two different environments. A similar transformation may be occurring for the other peptide adducts studied: Nap-AVFA and Nap-AREG. The concentration graphs of the dual incubations of Nap-AVFA + AVFA and Nap-AREG + AREG in ER water are shown in Fig. 41 along with the data for incubations of Nap-AFRV + AFRV re-plotted for comparison purposes. The removal rates for each of the peptides and adducts are very
similar for each incubation bottle and can be clearly observed in graphical form in Fig. 41 and quantitatively in Table 10. The emergence of two additional peaks in the incubations of Nap-AFRV (peaks 1 and 2), is also observed in those of Nap-AVFA (these are labelled peaks 3 and 4) and Nap-AREG (labelled peaks 5 and 6) dual incubations. It is clear from Fig. 41 that peaks 2, 4 and 6 follow a similar trend for all three adducts; increasing to a maximum concentration after 30 – 40 hours and then rapidly decreasing in concentration followed by complete removal after 50 – 80 hours. A second peak (peak 1, 3 and 5, respectively) emerges in each of the incubations after 24 hours and continues to increase in concentration, with several fluctuations, until removal after 150 hours for Nap-AFRV and 500 hours for Nap-AREG and Nap-AVFA. Due to the fact that peaks 1, 3, 5 and 2, 4, 6 have similar production and removal patterns for all three adducts it is logical to believe that they may involve similar transformations of the precursor adduct. To investigate the processes taking place, all the peaks were collected using the fraction collector attached to the HPLC, and analysed further.
7.3.2 Identification of emergent peaks

Each of the collected peaks is analysed by electrospray ionisation (ESI)-FT-ICR-MS, including, when possible, use of tandem MS to fragment the compounds and aid in their identifications. I use accurate m/z values (good to 1 ppm of mass) found for each of the collected eluates from the HPLC. This enables the assignment of an exact formula for each one of the peaks observed in HPLC. To facilitate structure determination for
each compound, I use tandem MS (CID) to fragment the peak and ascertain the elemental formulae for each of the fragments to a precision of 1 ppm. The success of tandem MS varied for each compound collected from the HPLC fraction collection, however, using all of the information available and considering likely changes to the structure of the peptide adduct I am able to determine the structure of each of the collected peaks. Proposed formation pathways are suggested for each peak based on likely chemical alterations, however, the exact mechanism and level of involvement by microorganisms present in the water in producing these compounds is unknown at this stage and beyond the scope of this work. Initially I will discuss the proposed structure of peak 3, whose structure was completely determined based on the fragmentation of the molecular ion.

Peak 3, collected from the incubation of Nap-AVFA was determined to have a m/z value of 588.240234 Da. Tandem MS results for peak 3 (Fig. 42) identifies numerous neutral mass losses and fragment ions that enable determination of a structure for this new compound. The first step to determining the structure of peak 3 is the identification of a likely formula. Each of the fragmentation peaks observed in Fig. 42, can also be assigned formulae with a high level of precision (Fig. 43). Upon examining the fragmented ions formulae and comparing them to the formula options for the molecular ion, the only formula that contains consistently higher amounts of C, H, N and O than the determined fragment formulae is C$_{26}$H$_{33}$N$_{7}$O$_{9}$. 
Fig. 42. Tandem MS of peak 3 during the ER water incubation of Nap-AVFA. Peaks 1 – 8 are identified.

Armed with this knowledge I can determine the exact mass of the fragments lost in the CID process and am able to determine an overall structure of peak 3 (Fig. 43). For example, fragment 1 immediately determines that there is at least one amine group easily lost. Fragment 2, shows an additional loss of a hydroxyl group, leading me to the conclusion that the quinone is in the reduced form, this type of loss has been observed for hydroquinone MS fragmentation previously (Mason and Liebler, 2000). Fragments 3 and 4 enable me to determine that an alanine initially attached in Nap-AVFA loses the methyl group becoming glycine. Fragments 5 and 6 indicate that a carboxylic acid group is attached to the end of a chain, indicating hydrolysis of one of the peptide bonds.
Fig. 43. Each of the suggested fragments displayed in Fig. 42 is indicated on the peak 3 structures. The assigned formula of each of the fragments is identified in addition to its loss formula from the molecular ion.
Fragment 7 allows me to identify the presence of a complete alanine molecule and an additional nitro group, speculated to be attached on the ring. This piece of information leads me to believe that two molecules of Nap-AVFA are necessary to form peak 3 since of the two alanine molecules present in the molecular ion, one is difficult to access suggesting that an additional Nap-AVFA is involved. The exact positions of the nitro and the glycine groups on Nap cannot be determined from the mass spectral data. However, I speculate that they are attached at the positions indicated in Fig. 43, due to the electron density of the ring. Fragment 8 allows me to determine that phenylalanine and alanine are still attached but can be removed from the overall structure. Consequently, it appears that valine is completely removed upon generation of this structure. Fragment 8, allows me to confirm the above proposal of the involvement of two molecules of Nap-AVFA, since 3 alanine molecules are required to form peak 3. All of the fragments are formed by removing certain sections of the peptide backbone, leaving Nap intact, this seems plausible based on past examination of fragmentation patterns (Tang et al., 1993). When cumulating the tandem MS data, the structure proposed in Fig. 42 explains all of the above listed fragments created during CID of peak 3 (Fig. 43). It is important to state that the structure proposed is not proven, because full confirmation really requires NMR and insufficient quantities of the pure compound could be recovered for this to be done.

Having determined the likely structure of peak 3, I am able to suggest mechanisms to explain its formation from Nap-AVFA. The proposed structure for peak 3 and the original Nap-AVFA structure are shown in Fig. 44. It is obvious that substantial modification of the original Nap-AVFA occurs. Nap-AVFA is divided into three sections specifying the locations where bonds are broken and reformed to produce the
proposed structure for peak 3. Section 1 contains alanine attached directly to Nap (the synthesis assured this), and the additional amine is likely sourced from valine, a process that involves hydrolysing the amide bond between alanine and valine. Two different transformations occur concerning section 1, involving two separate molecules of Nap-AVFA. In the first transformation, section 1 remains unaffected as shown in Fig. 44 (section 5). In the second molecule of Nap-AVFA, section 1 is detached from Nap and reattached at position 8, on the benzene ring, ortho to the quinone, via a Michael reaction (section 4). This is coupled with the removal of the side chain methyl group.

A reaction occurs, to attach nitrate at position 4 of the benzene ring, para to the attachment of section 4. Nitrate and nitrite have been found to react with PAHs in aqueous environments and could therefore, be reacting in a similar way with Nap-AVFA (Mack and Bolton, 1999). The source of the nitrate/nitrite added to peak 3 is not clear, but I speculate that it is already present in the water. The concentration of nitrate/nitrite has been measured as 9.7 µM in ER water, which is typical of an urban sourced river (Charette and Buesseler, 2004; Mulholland et al., 2008). Although this concentration is low, the electron poor character of the quinone could induce nitrate/nitrite to react with it, thereby producing peak 3. Section 2 of Nap-AVFA is lost completely in the formation of peak 3, the reason for this is not clear. Section 3 is detached from section 2, by hydrolysis of the amide bond between valine and phenylalanine, and then reattached to the meta carbon of the quinone via a Michael reaction (section 6). Likely due to the addition of the various groups around Nap, the quinone stays in a reduced form; the reason for this is unclear but could be due to stabilisation via resonance or rapid oxidation in the presence of dissolved OM.
Fig. 44. Structure of Nap-AVFA (top) and the proposed structure of peak 3 (bottom), formed during the degradation of Nap-AVFA in ER water collected by HPLC and analysed by positive ion mode FT-ICR-MS. Both structures are divided into sections, according to proposed changes to the Nap-AVFA structure.

Unfortunately it was not possible to fragment either of the structures of peaks 1 and 5, however, since all of the proposed changes are simple in nature, I used the structure and transformations of peak 3 as a guide to determine the most likely structures.
of peaks 1 and 5. Peak 1, formed during the incubation of Nap-AFRV, is determined to have a m/z of 659.268671, corresponding to a molecular formula of $[\text{C}_{26}\text{H}_{40}\text{N}_{10}\text{O}_{7}\text{Na}]^{+}$. The proposed structure of peak 1 is shown in Fig. 45, along with the structure of Nap-AFRV, annotated with the points of change. The suggested changes are determined primarily by examining the structure of Nap-AFRV. Due to the identification of the formula of peak 1, I can compare the formulae of these two molecules and assess the loss in C, H, N and O contents. The overall loss to form peak 1 from Nap-AFRV is $\text{C}_7\text{H}$ accompanied by the addition of $\text{N}_3\text{O}_2$. In order to maintain a similar H content in peak 1, it seems reasonable to suggest that the benzyl side chain of phenylalanine could be lost. In addition there are several bacteria that selectively remove and degrade aromatic structures in water, and I suggest that they are acting on Nap-AFRV (Leahy et al., 2003; Smith, 1990; Widdel and Rabus, 2001). The question of why these bacteria do not also attack Nap is not clear, but this could be either because of steric hindrance at the site of attack or the strength of the bonds attaching Nap to the rest of the peptide. These same reasons could also explain the stability of phenylalanine present in peak 3 which remains attached to the backbone peptide structure of Nap-AVFA throughout its transformation. In order to explain the addition of N and O to the structure of Nap-AFRV, I considered small molecules that are composed only of these atoms as possible additions. In Fig. 42 and 7.4, it is determined that a nitro group is added to Nap. Therefore, it seems reasonable that nitrate/nitrite might also be an additive to Nap-AFRV, thereby accounting for the addition of the N and O in the formula of peak 1. Thus far, in the suggested transformations to Nap-AFRV to form peak 1, there has been a loss of $\text{C}_7\text{H}_7$ and the addition of $\text{NO}_2$. The remaining difference in the formulae is $\text{N}_2\text{H}_6$. I suggest therefore, that one other change to Nap-AFRV to produce peak 1 is the addition of two amine
groups to Nap. I postulate that these are added via a Michael addition reaction. Ammonia reacts readily with quinones via a Michael addition reaction and so the addition of amines to produce peak 1 is not unexpected. The remaining two hydrogen’s can be accounted for by the quinone being in the reduced form. Since peak 3, is found to be in a reduced form, it is reasonable to suggest that peak 1 might also contain a reduced quinone for similar reasons.

Fig. 45. Structure of Nap-AFRV and proposed structure of peak 1, formed during the incubation of Nap-AFRV in collected ER water. The peaks were monitored and peak 1 was collected by HPLC and analysed by FT-ICR-MS positive ion mode. The changes during this transformation are labelled and annotated on both structures.
The m/z resulting from the analysis of peak 5, formed during the incubation of Nap-AREG is 637.305159 Da, with the assigned formula of \([C_{26}H_{41}N_{10}O_9]^+\). I propose that similar changes occur to Nap-AREG to produce peak 5 as those of Nap-AFRV to produce peak 1. These are displayed on the structures in Fig. 46. Most of the peptide remains intact, with no changes to the backbone or sidechains. All of the changes occurring to Nap-AFRV to produce peak 5 involve Nap. Three amine groups are added to Nap postulated to be via a Michael addition reaction. This is the only way that I can reasonably explain the increase in N in the formula. The quinone remains in the reduced form, similar to peaks 1 and 3, likely because of the addition of the amine groups and the assigned formula contains additional hydrogens that are not possible in the oxidised form. I postulate that the source of the amines is from the water itself, naturally dissolved in the water, since peak 5 is only formed when incubated in ER water.
Fig. 46. Structure of Nap-AREG and proposed structure of peak 5 formed during the incubation of Nap-AREG in ER water. Peak 5 was collected by HPLC and analysed by FT-ICR-MS positive ion mode. The changes to peak 5 from Nap-AREG are annotated on the peak 5 structure.

The changes that occur to Nap-AVFA, Nap-AFRV and Nap-AREG to produce peaks 3, 1, and 5 respectively seem to involve predominately Michael addition to Nap, and various alterations to the peptide chain. The reasons why these particular reactions occur is unknown, it is clear however, that these newly formed compounds are more stable and resistant to degradation than the original adduct and the peptide alone.
Other compounds formed as intermediates during the incubations of these three Nap-adducts, labelled peaks 2, 4 and 6. Following analysis of peaks 2, 4 and 6 for each of the adducts by MS, the determined exact mass was exactly the same for all three adducts (453.336256 Da). This is highly unexpected since the peptide structures for all three adducts are very different except the first alanine bonded to Nap. After detailed investigation of a structure that could be formed from all three peptide adducts, I was able to determine one as shown in Fig. 47. Tandem MS analysis of the collected peaks, produces the same fragmentation pattern for each of the adducts; suggesting that they are the same structure. The fragments produced by tandem MS of all three peaks is H$_2$O and H$_4$O$_2$. These fragments unfortunately only assist in demonstrating that there are two hydroxyl groups that can be easily lost, this could be explained with a reduced quinone form. This piece of information suggests that a quinone is still present in the molecule, this is reasonable considering Nap is preserved to form peaks 1, 3 and 5. The only formula that seems plausible that can be assigned with the m/z observed for peaks 2, 4 and 6 is [C$_{24}$H$_{45}$N$_4$O$_4$]$^+$. This formula consists of almost complete saturation of H, which initially seems difficult to explain due to the unsaturation levels of Nap-AVFA, Nap-AFRV and Nap-AREG. The highly saturated nature of this formula means that almost all of the bonds in the molecule are saturated, and since a quinone is a component of this structure (as determined from the fragmentation), only 2 further bonds can be saturated. Since this molecule is formed from all three of the adducts in this study, the resulting structure of peaks 2, 4 and 6, should contain the structural component that is common to all three original adducts. This component is Nap-A.
Fig. 47. The proposed structure of peak 2, 4 and 6 formed during the incubation of Nap-AFRV, Nap-AVFA and Nap-AREG respectively in ER water. Changes from the original structures of Nap-AVFA, Nap-AFRV and Nap-AREG are annotated on the structure.

Considering the assigned formula, [C_{24}H_{45}N_{4}O_{4}]^{+}, the N and O contents of this formula could be due to two alanine residues. I suggest that one alanine remains intact and attached to the quinone throughout the transformations to the original adducts. An attachment of a second alanine attached via a Michael adduction similar to what has been suggested for peaks 1, 3 and 5 could occur, using a second molecule of the adduct and therefore, fully accounting for the N and O contents of the proposed formula. Following this information the rest of the produced molecule must be completely saturated in nature, therefore, it is not possible for Nap to remain intact. Although speculative, I suggest that
the benzene ring is removed via a 2 step process forming phthalic anhydride, by firstly opening the ring, followed by oxidation and hydrolyxation reactions. This type of reaction has been found to occur under hydrothermal conditions (Onwudili and Williams, 2007). Even though the incubations in this experiment are at conditions very different from hydrothermal conditions, it takes 12 – 24 hours for this compound to be detected in the incubations in this study, allowing this reaction to occur at a slower rate. The rest of the molecule must be composed of saturated alkyl chains since this is the only remaining part of the above formula not accounted for. The alkyl chains that constitute the majority of this structure are most likely derived from the original peptide chain, with the sidechains and N/O all removed through a series of hydrolysis, elimination and substitution reactions. The exact details of the mechanism(s) involved in these reactions for each of the peptide adducts is unknown at this stage. Indeed, even the suggested structure is not proven due to the lack of NMR data.

7.4 Conclusions

It is clear that the adduction of Nap, greatly changes the removal pathway of these peptides from natural water. Since this experiment is preliminary in nature, my main goal was to determine whether Nap adduction has any effect on the removal of peptides from water, having established that there are substantial differences, future studies can direct attention to understanding the exact mechanisms involved in the removal of Nap peptides from ER and MLB water. Although during the incubations of Nap-AVFA, Nap-AFRV and Nap-AREG the Nap adducts themselves are removed from natural water at the same rate as the peptide alone, two new compounds are formed from each of these
adducts. Following tandem MS and careful consideration I am able to propose structures for each of the new compounds formed during Nap incubation. Unfortunately, I cannot determine the concentration of any of the new compounds in the water due to the fact that the small amount collected by HPLC was used completely during analysis by MS. It seems unlikely due to the large difference in structures that the two sets of peaks formed during the incubations of the three adducts are related. Therefore, two separate mechanistic pathways occurred to reduce the concentration of the adducts and form their respective degradation products. Nevertheless, it is clear that Michael adduction of peptides to Nap, dramatically changes the way it is degraded in natural water thereby proving that this pathway is a reasonable mechanism for large components of original peptides to be preserved in natural water. The more significant finding is that some of the peptide structural elements are preserved even though the original intact peptide is lost. This partially explains why amide bonds are detected (by nuclear magnetic resonance) in ancient buried sediments (del Rio et al., 2004; Hatcher et al., 1983; Knicker and Hatcher, 1997) or in aged water samples (Aluwihare et al., 2005; McCarthy et al., 1997) while the yields of hydrolysable amino acids are significantly more diminished than would be suggested from the NMR data. Extensive modifications of the peptides in such resulting structures ensures that hydrolysis, commonly used for the recovery of amino acid residues, yields products that are unrecognizable as peptide derivatives.
CHAPTER 8

CONCLUSIONS AND FUTURE DIRECTIONS

8.1 Conclusions and summary of research

The work of this dissertation was focussed upon investigating amide N in natural systems; in particular I was interested in understanding how its structure influenced its preservation and formation pathways. This compound class has been found to be recalcitrant in several systems but the reasons why this occurs are still not well understood (Hedges et al., 2000; Hedges and Keil, 1995; McCarthy et al., 1997). In order to investigate this subject I employed a variety of strategies developing a more accurate understanding of the molecular composition of organic N with the goal of assessing how this influences the preservation and formation routes of amide N.

To facilitate this process I used advanced analytical techniques, namely nuclear magnetic resonance (NMR) spectroscopy and Fourier transform ion cyclotron mass spectrometry (FT-ICR-MS) to investigate the molecular characteristics of N contents of sediments, and natural waters of a number of sites including most importantly, Mangrove Lake, Bermuda (MLB). In order to study the N-containing organic matter (NCOM) of the sediment I needed to develop a method to analyse them, due to fact that liquid samples are required for the electrospray (ESI) source. It is of primary importance when extracting a subsection of the sediment to ensure that a representative fraction is examined to ensure that all discussions pertaining to the whole sediment are valid. In Chapter 2, I investigated the ability of five solvents of varying chemical properties
(chloroform, aqueous base, dichloromethane, pyridine and water) to remove the largest and most representative of the NCOM of the MLB sediment. Bulk and NMR analysis of the five sediment extracts showed that the organic solvents were most effective at removing NCOM, and in addition, they removed functionalities of the sediment that dominate its solid matrix. MS analysis of the sediment extracts, demonstrated that methanol and pyridine were most effective, however, pyridine seemed best suited for my study due to its ability to remove more NCOM that could be visualised by MS. I then employed the pyridine extraction protocol to examine three additional sediments; one organic-rich sediment, from Mud Lake, Florida (MLF) and two organic-poor sediments, Bayou Grande, (BG), Pensacola, Florida and samples from the South Pass of the Mississippi Delta. The results of their MS comparison enabled me to observe clear differences in their N composition, demonstrating the effectiveness of pyridine to remove the most important N-containing components of any sediment. The results of this Chapter are pertinent to any MS analysis of a solid sample, whether it is sediment, or another type of solid sample (e.g., soil, degraded wood, fossilised wood etc.). A simple extraction with pyridine can remove a substantial portion of the OM, and indeed this approach has already been employed for other sample types (Koenig et al., 2010; Salmon et al., 2011; Zhong et al., 2011).

Having established a successful method for analysing sediment NCOM, I applied this technique throughout the rest of my dissertation. In particular, I was interesting in determining the optimum method for visualising the extremely complex dataset produced by MS analysis of an OM sample with particular reference to N-containing formulae. In Chapter 3 I examined several different diagrams to assess whether they were capable of providing specific information regarding the N-containing formulae; in particular, the
type of formula and probable structural nature of MLB pyridine extracted sediment. Although the commonly used van Krevelen diagram proved useful for separating N-containing formulae, it is not specific for these types of formulae and so limits its usefulness for examining N-containing compounds in particular. Kendrick mass defect analysis was also investigated but it is unable to ascertain the structural nature of formulae and so is not overly useful for this study. A plot of double bond equivalent/C versus O/C was able to display series homology but structural nature was not clearly elucidated, as well as the ability to specifically examine N-containing formulae. The best choice of display for specifically examining N-containing formulae was determined to be that of elemental N/H versus O/C. This diagram was demonstrated to specifically only plot those formulae that contained N and O, and able to display series homology and structural relationships. This diagram is unique to the community of scientist employing such diagrams in that it is specifically examining only those formulae that contain N. The N/H versus O/C diagram can easily be applied to any OM sample analysed by FT-ICR-MS to assess the general characteristics of the NCOM within it without pre-sorting of data for N.

In Chapter 4, I applied the pyridine extraction protocol developed in Chapter 2 to examine amide N in the sediment of MLB to assess whether all of the amide N was composed of peptides/proteins as has been previously suggested for this sediment (Knicker and Hatcher, 1997; Knicker and Hatcher, 2001). By using the 2D NMR technique of heteronuclear multiple bond correlation, I was able to differentiate two different types of amide N, one that could be correlated to peptides/proteins and another that was due to a newly discovered class of amide compounds in this sediment, alkyl amides. This was particularly of note since these types of compounds have been
previously identified as resistant to hydrolysis (the commonly used technique to analyse peptide/protein contents of OM samples) as a component of some algae species; (Derenne et al., 1993). Using FT-ICR-MS and I was able to identify alkyl amides and diamides throughout the sediments of MLB and MLF, and prove their structures as long chain primary amides by analysis of the extracts by gas chromatography mass spectrometry. This finding is very significant since this is the first time alkyl amides and diamides have been observed in sedimentary OM, not subjected to hydrothermal alteration, in particular in lacustrine sediments. The presence of this new class of compounds in sediment OM could be a major finding if they are produced through a simple abiotic reaction that could be active in other systems. I therefore, examined the possible formation of these compounds in another lacustrine environment. Glyceride esters are a common component throughout lacustrine systems, mostly from the algae that are dominant OM sources in the MLB and MLF systems (Hatcher, 1978). In addition, ammonia is highly prevalent in these systems, due to the anaerobic conditions of diagenesis of the OM (Hatcher, 1978). In Chapter 4 I proposed that amidation could be occurring between the glyceride esters and ammonia and proved that this reaction was possible by incubating sunflower oil with ammonium over 4 weeks, with the result of forming alkyl amides.

In Chapter 5, I wanted to demonstrate that amidation could clearly be discerned as a depth-dependent process in MLB to produce the alkyl amides. Therefore, using new fresh samples collected during the course of my dissertation in MLB, I analysed the NCOM of the lake water, the pore water and the sediment. I wanted to assess whether alkyl amides are present in the water of this Lake since it is likely that amidation occurs in the aqueous phase and whether I could link the amount of alkyl amides present to
amidation as a formation pathway. FT-ICR-MS analysis of the three OM pools of MLB revealed that their N-containing formulae are generally very different in number, structural diversity and source material (as determined from a N/H versus O/C plot developed in Chapter 3). However, the common structure group for the three OM pools was N-poor lipids, and within this class of compounds, alkyl amides are a significant component. After further investigation, I identified a trend of increasing dominance of alkyl amides and diamides with depth, which mirrored that of the ammonia concentration in the pore water. This observed trend lends significant credence to the abiotic production of alkyl amides and diamides by amidation fuelled by the increasing ammonia in the pore waters, and therefore, could be considered to also be an important process in other systems. Amidation could be a major pathway for N incorporation into OM, in any system which has high concentrations of ammonia and the presences of esters.

Another focus of my dissertation was examining the preservation routes in natural waters of another class of amides, peptides and proteins. To accomplish this I synthesised a series of peptides (AVFA, AFRV, AAFRV and AREG) and incubated them in MLB lake water and Elizabeth River (ER), VA riverine water. I compared the rate of removal, presumably via abiotic degradation of the peptides in Chapter 6 by measuring their change in concentration by high performance liquid chromatography (HPLC). From my results it seems that chemical structure i.e., the residue sequence plays little role in determining the rate of removal of the tetrapeptides in this study. This result is similar to a previous study, when the order of amino acids in the same sequence was found to have no impact on removal rate in natural water (Liu et al., 2010). Peptide chain length seems to play a more prominent role in peptide removal rate, since a distinct difference in removal rate was observed between a pentapeptide and tetrapeptide (AAFRV and
AFRV); whose only difference in structure was one alanine. Upon closer examination of the chromatograms of AAFRV, it became apparent that a new peak was being formed, and the retention time of this new peak exactly matched that of AFRV. This gives substantial evidence to the fact that AAFRV degrades to AFRV, before its complete removal from the natural waters. In addition, since AAFRV is very close to the molecular weight cut-off of direct uptake by microbial cells (~ 600 Da), it is probable that the transformation of AAFRV (562 Da) to AFRV is facilitated by extracellular hydrolysis (Arnosti, 2011; Nikaido and Vaara, 1985). In addition, no extra peaks are observed in the chromatograms of the tetrapeptides in this study (AREG, AVFA and AFRV) providing further evidence for this hypothesis. Past studies of peptide removal in natural waters has demonstrated that length of chain is a key factor in determining the rate and that hydrolysis seems to be directed towards certain bonds first (Liu et al., 2010; Pantoja and Lee, 1999). My results therefore, support this suggestion based on the fact that specifically only the A-A bond is cleaved in the removal of AAFRV from natural water probably because it is the smallest amino acid at the end of the chain (versus V).

In Chapter 6, I incubated the synthesised peptides in two different aqueous environments, one lacustrine (MLB) and the other riverine (ER). The observed difference in removal rate of the same peptides incubated in the two different waters was significant, with a $3 - 6 \, \text{d}^{-1}$ increase in rate for ER water incubations. All of the incubations in this study showed this difference in rate, suggesting that the water environment was the cause of the rate difference. Although I am unable to suggest a definitive reason why this occurs it is probably because there are different microbial communities present in the two water systems and the significantly different OM composition in the two collected waters, with ER water containing more lignin and tannin
like structures, that could be associating with the synthesised peptides, thereby reducing their observed concentration in the water (Hsu and Hatcher, 2005; Rillig et al., 2007). These differences in observed removal rates for these two locations show that environment of peptide degradation plays an important role in its removal rate, and is probably why there is such a large difference in observed rates in past studies (Hollibaugh and Azam, 1983; Kuznetsova and Lee, 2001; Kuznetsova et al., 2004; Liu et al., 2010; Mulholland and Lee, 2009; Pantoja and Lee, 1999; Pantoja et al., 1997; Pantoja et al., 2009).

In Chapter 7, I extended my study of peptide removal rates in natural waters to investigate one of the proposed hypotheses for peptide preservation, Michael adduction to quinones (Hsu and Hatcher, 2005; Hsu and Hatcher, 2006). After synthesising three tetrapeptides using a novel approach (AVFA, AFRV and AREG) adducts to 1, 4-naphthoquinone (Nap) I incubated them in MLB and ER waters and monitored their concentration using HPLC. Comparing the adduct and non-adducted peptide incubations I observed no difference in rate of removal, even when incubated together in the same bottle. However, in the incubations of the Nap-adducted peptides, I observed a series of additional peaks that grew in with time during the course of the experiments. After collecting these new peaks using the fraction collector attached to the HPLC, and analysing them by FT-ICR-MS with the use of tandem MS I was able to provide further insight regarding each of the new peaks formed from the three Nap adducted peptides. In each of the incubations two peaks were observed for each of the adducts, and labelled 1 – 6. Comparing their emergence patterns with time, the two new peaks seem to follow similar trends for each of the adducts suggesting similar structures and production pathways. After in-depth analysis I was able to propose structures for each of the peaks
observed in the chromatograms, and it seems that many of the transformations are similar for similar trended peaks. Interestingly, one of the new peaks is exactly the same for all three adducts, with a significantly altered structure from the original adducts. These results demonstrate for the first time that adduction to a quinone can alter the removal pathways of a peptide in a natural system. This perhaps demonstrates why often the measured concentration of peptides is so low in natural waters, because they have been too severely altered from the original structure (McCarthy et al., 1997).

8.2 Future directions

Much of the research in this dissertation is preliminary in nature and consequently, there are many possible future research options available. The analytical methods I developed in Chapters 2 and 3 can easily be applied to a wide variety of OM samples for FT-ICR-MS characterisation. Employing an extraction technique that reliably removes a substantial, representative fraction of solid OM (e.g., sediment and soil) is invaluable for their analysis by FT-ICR-MS, and I suggest that pyridine extraction is the ideal method to achieve this. Due to the scarcity of studies focused on FT-ICR-MS analysis of N-containing formulae, I necessarily found a graphical method that can display a multifaceted amount of information about the NCOM that can be applied to any study involving MS. It is my intention in proposing the N/H versus O/C diagram that it will become a well-employed diagram, much like the van Krevelen diagram for O-containing formulae, to analyse the N-containing compounds of OM samples.

The identification of a new class of compounds, alkyl amides and diamides that contribute to the recalcitrant amide fraction of sediment, pore water and lake water OM
of MLB could be a major finding with major implications for other systems. In addition, since the route of production of these compounds by amidation is a simple, commonly known reaction that readily occurs within a few weeks (in sunflower oil); it is highly plausible that this could be a major pathway for N incorporation into OM. Future directions for this research would include confirming that amidation is the definitive pathway of formation for these compounds, by using isotope labelling and specifically examining the isotopic signature of alkyl amides and linking it to the ammonia present in the pore water. In order to achieve this, it will probably be necessary to develop a strategy to separate these compounds from the rest of the OM matrix, possibly by using HPLC or chemical treatments. Upon confirming this pathway, future research could be directed to indentifying these compounds in other systems such as in ocean water, estuarine systems, low organic matter lacustrine systems, soils and atmospheric aerosols. Since ammonia and glyceride esters as well as other types of humic-like esters, are observed in all of these environments it is possible that alkyl amides are actively produced on a global scale.

My research on peptide preservation can be expanded to many new research projects based on my initial results. Firstly, I suggest increasing the number of synthesised peptides in future studies in terms of amino acids included, and number of residues per peptide (e.g., ranging from 4 – 10 amino acids) and measuring their removal rates. This would provide further evidence for the differences in length I observed for removal rate and confirm that the amino acid chemistry has little influence regarding its removal rate. This would greatly increase the applicability of the results to natural conditions. In addition, I would incubate synthesised peptides in a number of different waters collected from a variety of different systems. This would include more waters
heavily polluted by anthropogenic sources like the ER water in my research and compare the removal rate of peptides incubated in water from pristine locations without such influences like MLB. Through this research I could confirm whether the observed differences in removal rate are due to the presence of anthropogenically produced compounds. Concurrently, I would analyse the different collected waters for their inherent chemistries, employing such measurements as I have in my research such as measuring salinity, pH, temperature and analysing the OM molecular composition. In addition, I would also study the microbial community in more detail, to identify the specific differences in the type and number of bacteria and phytoplankton in each water system. In doing so, I should be able to provide more specific reasons for differences in removal rates of peptides in these waters.

My results regarding Michael adduction of peptides to quinones, has many future research options. Since this is the first time, to my knowledge, that Michael adduction has been shown to alter the peptide degradation pathway in natural waters there is much possibility for future studies. Initially I suggest including more peptides, of varying amino acid types and residue lengths to adduct to Nap to ascertain whether the common peak that I observed in this study is produced by all peptides. This will involve synthesising more adducts that do not contain the first amino acid as alanine bonded to Nap. This will prove more difficult, due to their inevitably larger size and will likely involve substantial research in order to develop a successful synthetic pathway. Since I was unable to unambiguously identify the structures of the peaks formed during the incubations of Nap-AVFA, Nap-AFRV and Nap-AREG, future research could be directed to uncovering their structures by re-incubating them in water freshly collected from ER (since it is an easier location to collect at from ODU) but at a higher quantity
and analysing them by additional methods such as NMR.

An additional future direction of Michael adduction research could involve investigating the exact production pathway of the observed extra peaks in this study. I suggest this could be accomplished by several different methods, this includes incubating peptide adducts of varying lengths (e.g., incubate Nap-AVFA, Nap-AVF, Nap-AV and Nap-A) to assess whether they produce similar peaks and also different sequence orders of the same peptide (e.g., Nap-AVFA, Nap-AAVF, and Nap-AFVA). These two research projects would provide additional detail regarding the pathway by which additional peaks are produced.

Additional future research could include similar incubation studies to investigate other hypothesised methods by which peptides are proposed to be preserved through including formation of Melanoidins by reacting with carbohydrates, reaction with mineral matter, and further investigation of Michael reactions with other oxidised polycyclic aromatic hydrocarbons type structures e.g., other quinones, such as 1, 2-naphthoquinone, benzoquinone, and anthraquinone (Kirchman et al., 1989; Maillard, 1916; Rillig et al., 2007).

It is clear that the research in my dissertation gives further information regarding the molecular composition of amide N, and their proposed formation and preservation pathways. The use of advanced analytical techniques to their investigation has proved invaluable to shed new light on this area of research. It is clear that more research is required before some of the postulated hypotheses in my dissertation, particularly regarding the presence of alkyl amides, are demonstrated widely applicable but overall this research brings the organic geochemical community one step closer to understanding the intricate complexities organic N have upon on the global N cycle.
REFERENCES


Harvey H. R. and Mannino A. (2001) The chemical composition and cycling of particulate and macromolecular dissolved organic matter in temperate estuaries as


Knicker H., Lüdemann H.-D. and Haider K. (1997) Incorporation studies of NH$_4^+$ during


Leahy J. G., Tracy K. D. and Eley M. H. (2003) Degradation of mixtures of aromatic and


Maie N., Parish K. J., Watanabe A., Knicker H., Benner R., Abe T., Kaiser K. and Jaffé


Stevenson F. J. and Goh K. M. (1971) Infrared spectra of humic acids and related


COPYRIGHT PERMISSION

Displayed below is the copyright permission for reproducing Chapter 4 in this dissertation:

This is a License Agreement between Georgina A McKee ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier

Elsevier Limited
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB, UK

Registered Company Number: 1982084

Customer name: Georgina A McKee

Customer address: 1347 Battleford Drive, Virginia Beach, VA 23464

License number: 2692121072710

License date: Jun 18, 2011

Licensed content publisher: Elsevier

Licensed content publication: Geochimica et Cosmochimica Acta

Licensed content title: Alkyl amides in two organic-rich anoxic sediments: A possible new abiotic route for N sequestration

Licensed content author: Georgina A. McKee, Patrick G. Hatcher
APPENDIX B

EVIDENCE FOR SUCCESSFUL PEPTIDE AND ADDUCT SYNTHESSES

The following figures display the resulting spectra from nuclear magnetic resonance (NMR) spectroscopy and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) analyses of synthesised peptides and 1, 4-naphthoquinone peptide adducts utilised in Chapters 6 and 7.
Fig. B1. Results of analysis of the synthesised AAFRV by (a) nuclear magnetic resonance spectroscopy (NMR) and (b) Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) in positive ion mode.
Fig. B2. Results of analysis of the synthesised AFRV by (a) NMR and (b) FT-ICR-MS in positive ion mode.
Fig. B3. Results of analysis of the synthesised AREG by a. NMR and b. FT-ICR-MS in positive ion mode.
Fig. B4. Results of analysis of the synthesised AVFA by (a) NMR and (b) FT-ICR-MS in positive ion mode.
Fig. B5. Results of analysis of the synthesised 1, 4-naphthoquinon-2-yl-AFRV (Nap-AFRV) by (a) NMR and (b) FT-ICR-MS in positive ion mode.
Fig. B6. Results of analysis of the synthesised Nap-AREG by a. NMR and b. FT-ICR-MS in positive ion mode.
Fig. B7. Results of analysis of the synthesised Nap-AVFA by (a) NMR and (b) FT-ICR-MS in positive ion mode.
APPENDIX C

ADDITIONAL FIGURES FOR A FUTURE PUBLICATION

The following figures and tables are intended to be formulated into a publication, as an extension of work in Chapter 5.

![Image of various measurements]

Fig. C1. Depth profiles of various measurements of the PW and sediment samples collected from MLB in November 2009. Measurements displayed include, TOC and IC concentration (mM) in the PW (top left), OC % in the sediment (bottom left), TN concentration (mM) in the PW (middle top), N and H % in the sediment (middle bottom), NH$_4^+$ concentration (mM) in PW (top right) and pH of the PW (bottom right). When applicable, error bars are displayed that were measured using standard deviation.
Fig. C2. Pyridine extract mass spectra of the mass range m/z 397 of 9 different depths of sediment collected from MLB in 2009, analysed by FT-ICR-MS negative ion mode. Peaks are scaled to the same relative magnitude and as a consequence some peaks are displayed off scale. Most peaks are assigned elemental formulae, and are numbered 1 – 16; the formulae are listed in Table C1. Peaks labelled as *, could not be assigned an elemental formula.
Table C1

Data pertaining to the pyridine extracted sediments analysed by FT-ICR-MS negative ion mode of the spectra of the 397 Da mass range of a depth profile from MLB. Peaks are labelled in Fig. C2, and their formulae are displayed below.

<table>
<thead>
<tr>
<th>Peak Label</th>
<th>Formula</th>
<th>Peak Label</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>([\text{C}<em>{15}\text{H}</em>{17}\text{N}<em>{4}\text{O}</em>{5}\text{S}_2])⁻</td>
<td>9</td>
<td>([\text{C}<em>{17}\text{H}</em>{13}\text{O}_8\text{S}])⁻</td>
</tr>
<tr>
<td>2</td>
<td>([\text{C}<em>{14}\text{H}</em>{21}\text{O}_{11}S])⁻</td>
<td>10</td>
<td>([\text{C}<em>{14}\text{H}</em>{28}\text{N}_2\text{O}_5])⁻</td>
</tr>
<tr>
<td>3</td>
<td>([\text{C}<em>{13}\text{H}</em>{22}\text{N}_2\text{O}_5\text{S}_2])⁻</td>
<td>11</td>
<td>([\text{C}<em>{13}\text{H}</em>{19}\text{O}_5\text{S}])⁻</td>
</tr>
<tr>
<td>4</td>
<td>([\text{C}<em>{14}\text{H}</em>{22}\text{N}_2\text{O}_5\text{S}_2])⁻</td>
<td>12</td>
<td>([\text{C}<em>{25}\text{H}</em>{33}\text{O}_4})⁻</td>
</tr>
<tr>
<td>5</td>
<td>([\text{C}<em>{15}\text{H}</em>{21}\text{N}_2\text{O}_6})⁻</td>
<td>13</td>
<td>([\text{C}<em>{25}\text{H}</em>{37}\text{N}_2\text{O}_2])⁻</td>
</tr>
<tr>
<td>6</td>
<td>([\text{C}<em>{14}\text{H}</em>{20}\text{N}_2\text{O}_5\text{S}_2])⁻</td>
<td>14</td>
<td>([\text{C}<em>{25}\text{H}</em>{35}\text{O}_4})⁻</td>
</tr>
<tr>
<td>7</td>
<td>([\text{C}<em>{20}\text{H}</em>{29}\text{O}_3\text{S}])⁻</td>
<td>15</td>
<td>([\text{C}<em>{25}\text{H}</em>{49}\text{O}_1})⁻</td>
</tr>
<tr>
<td>8</td>
<td>([\text{C}<em>{20}\text{H}</em>{29}\text{O}_3\text{P}])⁻</td>
<td>16</td>
<td>([\text{C}<em>{25}\text{H}</em>{55}\text{O}_2})⁻</td>
</tr>
</tbody>
</table>
Fig. C3. PW DOM extract mass spectra (397 Da) of a depth profile from MLB sediment collected in November 2009 and analysed by FT-ICR-MS negative ion mode. Peaks with assigned formulae are numbered and their formulae are listed in Table C2. * indicates that it was not possible to assign a formulae.
Table C2

Information regarding the spectra of the 397 Da mass range of PPL extracted pore water DOM analysed by FT-ICR-MS negative ion mode. Peaks are labelled in Fig C3, and their corresponding formulae are listed below.

<table>
<thead>
<tr>
<th>Peak Label</th>
<th>Formula</th>
<th>Peak Label</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>([\text{C}<em>{1}\text{H}</em>{1}\text{O}<em>{2}\text{S}</em>{2}])</td>
<td>15</td>
<td>([\text{C}<em>{25}\text{H}</em>{29}\text{O}_{8}\text{S}])</td>
</tr>
<tr>
<td>2</td>
<td>([\text{C}<em>{1}\text{H}</em>{1}\text{O}_{2}\text{S}])</td>
<td>16</td>
<td>([\text{C}<em>{24}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>3</td>
<td>([\text{C}<em>{14}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>17</td>
<td>([\text{C}<em>{26}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>4</td>
<td>([\text{C}<em>{19}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>18</td>
<td>([\text{C}<em>{18}\text{H}</em>{29}\text{N}<em>{4}\text{O}</em>{7}\text{S}])</td>
</tr>
<tr>
<td>5</td>
<td>([\text{C}<em>{24}\text{H}</em>{29}\text{O}_{6}\text{S}])</td>
<td>19</td>
<td>([\text{C}<em>{25}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>6</td>
<td>([\text{C}<em>{18}\text{H}</em>{29}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>20</td>
<td>([\text{C}<em>{24}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>7</td>
<td>([\text{C}<em>{18}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>21</td>
<td>([\text{C}<em>{23}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>8</td>
<td>([\text{C}<em>{23}\text{H}</em>{29}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>22</td>
<td>([\text{C}<em>{24}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>9</td>
<td>([\text{C}<em>{23}\text{H}</em>{29}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>23</td>
<td>([\text{C}<em>{25}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>10</td>
<td>([\text{C}<em>{18}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>24</td>
<td>([\text{C}<em>{26}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>11</td>
<td>([\text{C}<em>{18}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>25</td>
<td>([\text{C}<em>{26}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>12</td>
<td>([\text{C}<em>{18}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>26</td>
<td>([\text{C}<em>{27}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>13</td>
<td>([\text{C}<em>{18}\text{H}</em>{25}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>27</td>
<td>([\text{C}<em>{28}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
<tr>
<td>14</td>
<td>([\text{C}<em>{23}\text{H}</em>{29}\text{O}<em>{6}\text{S}</em>{2}])</td>
<td>28</td>
<td>([\text{C}<em>{29}\text{H}</em>{29}\text{O}_{7}\text{S}])</td>
</tr>
</tbody>
</table>
Average m/z, C, H, N, O, and S of assigned formulae by depth as determined by in sediment and PW extracts analysed by negative ion mode FT-ICR-MS.
Fig. C5. Average ratios of bulk and MS data of PW and sediment samples from a core collected in MLB. Ratios are determined by FT-ICR-MS negative ion mode analysis, elemental analysis and TOC/TN analysis.
Fig C6 van Krevelen diagrams of the LW and four different depths of PW and sediment extracts collected from MLB and analysed by FT-ICR-MS. The two sample types (PW and sediments) are plotted on the same diagrams and are separated as shown in the legend. Two regions are drawn on the figures to bring attention to the most prominent changes by depth occurring to the PW and sediment sample types.
Fig. C7. Information regarding the similar formulae observed in both the PW and sediment extracts (duplicates) analysed by FT-ICR-MS negative ion mode of samples from MLB. (a) lists the percentage of formulae that are duplicates in the PW and sediment by depth and (b) displays the duplicate formulae as classified by formula type for each of the depths.

### a

<table>
<thead>
<tr>
<th>Depth (cm)</th>
<th>Percent of duplicate formulae in PW compared to sediment</th>
<th>Percent of duplicate formulae in sediment compared to PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 1</td>
<td>8.52</td>
<td>13.24</td>
</tr>
<tr>
<td>4 - 5</td>
<td>9.24</td>
<td>11.89</td>
</tr>
<tr>
<td>10 - 15</td>
<td>11.45</td>
<td>9.03</td>
</tr>
<tr>
<td>20 - 25</td>
<td>11.65</td>
<td>16.49</td>
</tr>
<tr>
<td>40 - 45</td>
<td>11.44</td>
<td>9.63</td>
</tr>
<tr>
<td>65 - 70</td>
<td>10.91</td>
<td>7.39</td>
</tr>
<tr>
<td>80 - 85</td>
<td>8.66</td>
<td>11.98</td>
</tr>
<tr>
<td>120 - 125</td>
<td>11.44</td>
<td>10.70</td>
</tr>
<tr>
<td>320 - 330</td>
<td>11.33</td>
<td>8.34</td>
</tr>
</tbody>
</table>

### b

![Diagram showing percent of formulae in each classification](chart.png)

- CHO
- CHOS
- CHNO
- CHNOP
- CHNOS
- CHS
- CHOP
VITA

GEORGINA ANNE McKEE

Address: 1347 Battleford Drive
Virginia Beach
VA, 23464

Email: gmckee@odu.edu

Nationality: British

Date of Birth: 3rd September 1983

Education and Qualifications

Department of Chemistry and Biochemistry, Old Dominion University, Norfolk, Virginia, 23529
Ph.D Chemistry

Department of Chemistry, University of Sheffield, Sheffield, S3 7HF, United Kingdom
M.Chem Chemistry, Chemistry with year in USA (class 2:1)

Publications


