Molecular characterization of dissolved organic matter associated with the Greenland ice sheet

Authors:

Maya P. Bhatia[1], Sarah B. Das[2], Krista Longnecker[3], Matthew A. Charette[3], and Elizabeth B. Kujawinski[3](*)

Contact information:

1 – MIT/WHOI Joint Program in Oceanography/Applied Ocean Sciences and Engineering; Department of Geology and Geophysics; Woods Hole Oceanographic Institution; Woods Hole, MA 02543

2 – Department of Geology and Geophysics; Woods Hole Oceanographic Institution; Woods Hole MA 02543

3 – Department of Marine Chemistry and Geochemistry; Woods Hole Oceanographic Institution; Woods Hole, MA 02543

(*) – Corresponding author – Department of Marine Chemistry and Geochemistry; Woods Hole Oceanographic Institution; 360 Woods Hole Rd. MS#4; Woods Hole MA 02543; 508-289-3493; ekujawinski@whoi.edu

Index Terms: glacier, ultrahigh resolution mass spectrometry, FT-ICR, organic carbon, DOM

Running Head: DOM associated with the Greenland ice sheet

ABSTRACT

Subsurface microbial oxidation of overridden soils and vegetation beneath glaciers and ice sheets may affect global carbon budgets on glacial-interglacial timescales. The likelihood and magnitude of this process depends on the chemical nature and reactivity of the subglacial organic carbon stores. We examined the composition of carbon pools associated with different regions of the Greenland ice sheet (subglacial, supraglacial, proglacial) in order to elucidate the type of dissolved organic matter (DOM) present in the subglacial discharge over a melt season. Electrospray ionization (ESI) Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry coupled to multivariate statistics permitted unprecedented molecular level characterization of this material and revealed that carbon pools associated with discrete glacial regions are comprised of different compound classes. Specifically, a larger proportion of protein-like compounds were observed in the supraglacial samples and in the early melt season (spring) subglacial discharge. In contrast, the late melt season (summer) subglacial discharge contained a greater fraction of lignin-like and other material presumably derived from underlying vegetation and soil. These results suggest (1) that the majority of supraglacial DOM originates from autochthonous microbial processes on the ice sheet surface, (2) that the subglacial DOM contains allochthonous carbon derived from overridden soils and vegetation as well as autochthonous carbon derived from *in situ* microbial metabolism, and (3) that the relative contribution of allochthonous and autochthonous material in subglacial discharge varies during the melt season. These conclusions are consistent with the hypothesis that, given sufficient time (e.g., overwinter storage), resident subglacial microbial communities may oxidize terrestrial material beneath the Greenland ice sheet.

1. INTRODUCTION

Anticipating how carbon flux patterns might respond to climate change is a principal motivation for understanding the different sources and reservoirs contributing to the global carbon cycle. In aquatic systems, carbon flux patterns result from complex metabolic interactions of diverse biota with a pool of organic matter (Azam, 1998). Previously it was believed that glacial environments were devoid of life and thus, that carbon dynamics in these systems should be dominated by abiotic processes (Chillrud et al., 1994; Raiswell, 1984). However, the recent discovery of large, active microbial communities beneath glaciers and ice sheets has enlightened our understanding of biogeochemical reactions and organic carbon cycling in glaciated regions, namely that subglacial microbial communities may play an active role in the carbon cycle through oxidation of organic carbon stores beneath ice masses (Lanoil et al., 2009; Sharp et al., 1999; Tranter et al., 2002). On glacial-interglacial timescales, microbial activity might provide an important source of acidity to fuel chemical weathering of silicate rocks, a long-term control on atmospheric CO_2 levels (Berner et al., 1983; Brown, 2002). In addition, microbes may respire or ferment soil organic carbon (to CO_2 or to CH_4, respectively), previously considered inert until deglaciation (Sharp et al., 1999). Wadham et al. (2008) estimated that between 418 to 610 Pg of organic carbon was present beneath ice sheets during the last glacial period, of which 63 Pg C was available for conversion to methane over a glacial cycle. Additionally, Skidmore et al. (2000) calculated that aerobic respiration of subglacial organic carbon could convert 8.1 Pg C to carbon dioxide over a glacial cycle. These calculations, however, are constrained by a lack of knowledge concerning the availability of the subglacial organic carbon stores to microbial degradation. This is a potentially large limitation, given the range in biological reactivity within all other organic carbon stores (Eglinton and Repeta, 2003; Hedges et al., 2000). In order to examine the impact of microbial oxidation on subglacial organic carbon stores, it is critical to assess the composition and reactivity of this material.

Carbon is derived from two distinct regions of the glacial environment: (1) on the glacier surface (i.e., the supraglacial environment) from inorganic and organic carbon in snow and ice; and (2) at the glacier base (i.e., the subglacial environment) where carbon is derived from the underlying bedrock,

72 sediments, and ice. These two regions are linked by a hydrological network that becomes activated during
73 the summer melt season when accumulated surface meltwaters drain through crevasses, moulins, and
74 englacial channels to the bed (e.g. Das et al., 2008; Nienow et al., 1998). Once at the bed, the supraglacial
75 meltwaters become connected to a broad subglacial hydrological drainage network, in contact with the
76 underlying till and bedrock (Nienow et al., 1998). Generally, dissolved organic carbon (DOC)
77 concentrations in supraglacial snow and meltwater are very low (~10-40 μM) (Lafreniere and Sharp,
78 2004; Lyons et al., 2007). In contrast, available organic carbon sources in subglacial environments have
79 variable DOC concentrations ranging from 60 to 700 μM as reflected in subglacial outflow waters
80 (Lafreniere and Sharp, 2004; Skidmore et al., 2005) and concentrations up to ~ 4 mM (Dry Valleys,
81 Antarctica) and ~ 20 mM (Ellesmere Island, Canada) in basal ice samples (Barker et al., 2006; Bhatia et
82 al., 2006). Although measurements are limited, this variability observed among subglacial DOC
83 concentrations is likely a function of sampling time and/or of different physical characteristics (e.g.
84 lithologies, sediment content, proximity to land) between and within specific field sites.

85 While bulk DOC abundance studies are useful as first-order investigations, they offer little
86 information regarding the provenance, reactivity and bioavailability of the glacial organic carbon pools.
87 In an effort to address these issues, Lafreniere and Sharp (2004) and Barker et al. (2006) used
88 spectrofluorometric techniques to distinguish subglacial fulvic acids (the portion of humic material which
89 is water-soluble at any pH) derived from terrestrial precursor material from those of microbial origin.
90 Terrestrially derived dissolved organic matter (DOM) would contain fulvic acids from plant and soil
91 organic matter, which are typically more aromatic, due to the presence of compounds such as lignins
92 (McKnight et al., 2001). Alternatively, microbially-derived DOM would contain fulvic acids from
93 microbial cell components and metabolism, and are typically less aromatic (McKnight et al., 1994;
94 McKnight et al., 2001). Both Lafreniere and Sharp (2004) and Barker et al. (2006) found that supraglacial
95 samples contained microbially-derived fulvic acids, which they attributed to primary productivity of algae
96 and bacteria in the snow, ice, and meltwater on the glacial surface. However, results from the subglacial
97 runoff were more variable, with both studies finding sources of fulvic acids with both microbial and

98 terrestrial provenance. These findings were attributed to changing subglacial flow-routing regimes
99 throughout the melt season that access different carbon pools as well as to *in situ* subglacial microbial
100 metabolisms that alter the subglacial carbon pools.

101 Though an important first step in compositional assessment of glacial organic carbon pools,
102 fluorescence spectroscopy studies are limited because (1) they can only assess one fraction of DOM
103 (fulvic acids), and (2) they do not directly identify the presence of specific compounds within the DOM
104 pool, thus permitting only broad distinctions between 'microbial' and 'terrestrial' components. In
105 contrast, electrospray ionization (ESI) coupled to Fourier transform ion cyclotron resonance mass
106 spectrometers (FT-ICR MS) provides an opportunity to study a larger portion of the DOC pool (intact
107 polar molecules), and to characterize the reactivity of specific molecules in biogeochemical processes.
108 ESI is a 'soft' (low-fragmentation) ionization technique that detects polar molecules with acidic and basic
109 functional groups. When coupled to a mass spectrometer, such as FT-ICR MS which is capable of
110 ultrahigh mass resolution (>100,000) and mass accuracy (<1 ppm), tens of molecules can be accurately
111 resolved at each nominal mass (Kujawinski, 2002; Marshall and Rodgers, 2008). The mass accuracy
112 achievable is the key to this technique as it enables the assignment of elemental formulae solely from the
113 mass measurement (Kim et al., 2006; Kujawinski and Behn, 2006). Therefore, ESI FT-ICR MS can be
114 used to identify compositional differences among pools of DOM, as well as to determine the elemental
115 compositions of specific molecules within DOM. Recently, ESI FT-ICR MS has been utilized to
116 characterize DOM in a range of diverse environments, including freshwater systems (Sleighter and
117 Hatcher, 2008), marine systems (Koch et al., 2005), and ice cores (Grannas et al., 2006).

118 The goal of this study was to investigate the compositional nature of carbon pools associated with
119 different regions of the Greenland ice sheet in order to elucidate the type of dissolved organic matter
120 present in the subglacial discharge over a melt season. The carbon pools explored were (1) the
121 supraglacial environment: snow and meltwater on the ice surface, (2) the subglacial environment: water
122 exiting the base of a land-terminating outlet glacier, and (3) the pro-glacial tundra environment: non-
123 glacially derived pond water. From a hydrological perspective, these environments are serially connected

to each other as the majority of the supraglacial meltwater on a glacier surface penetrates to the subglacial environment and eventually exits into the proglacial environment. Thus, the compositional characteristics of the contributing carbon pools as well as physical and microbial processes en route ultimately dictate the composition of the DOM in the subglacial discharge. We employed ESI FT-ICR MS to detect compositional differences among the different carbon pools sampled, and to gain insight into the molecular-level impact of microbial metabolism on subglacial organic carbon. By establishing baseline values of the type of organic carbon present beneath glaciated areas, this study serves as the foundation for broader investigations into the impact of increased meltwater runoff from the Greenland ice sheet to surrounding marine environments, and into the extent of subglacial microbial oxidation of overridden soils and vegetation.

2. METHODS

2.1. Field Sites

This study was conducted at two locations along the western margin of the Greenland ice sheet in 2007 and 2008. In July 2007 two snow samples and one supraglacial meltwater sample were collected from the ablation zone on the ice sheet surface, at 980-m elevation approximately 40 km inland from the edge of the ice sheet (Figure 1). By July most of the seasonal snow deposited the previous winter had already melted, thus our samples were collected from isolated pockets of heavily metamorphosed and colored snow from drifts along the banks of relict stream channels. Of the two snow samples analyzed for this study, one exhibited a yellow and green hue (Yellow Snow) and the other a red and black hue (Red Snow). The supraglacial meltwater sample (Supraglacial Inland) was collected from the edge of a large meltwater lake (~1 km in diameter.) Given the scarcity of seasonal snow on the ice sheet surface during our sampling period, and the high annual ablation rates we measured at this site (~ 2-m ice melt yr^{-1}), this meltwater sample is assumed to be derived almost entirely from glacial ice melt rather than from seasonal snow melt or rainfall.

149 In May and July 2008, samples were collected in the vicinity of a small land-terminating outlet
150 glacier (named glacier 'N' here), approximately 70 km south of the 2007 site (Figure 1). In May, one
151 sample was collected from a small supraglacial meltwater pond (~20 m in diameter) within 1 km of the
152 ice sheet margin (Supraglacial Margin). The water here consisted primarily of snow and ice melt. A
153 second sample was collected from the subglacial stream exiting at the base of glacier 'N' (Subglacial
154 May). A third sample was collected at a proglacial pond (Tarn). In July, two additional samples were
155 collected from the subglacial stream exiting the base of glacier 'N' (Subglacial July-1 and Subglacial
156 July-2, referred to collectively as Subglacial July). A synopsis of the samples collected in this study and
157 the filtration and extraction procedures (details below) is presented in Table 1. Electrical conductivity
158 (EC) measurements were made on-site using a Russell RL060C meter (Thermo Electron) for the
159 Subglacial May and July and Supraglacial Margin samples, and are also presented in Table 1.
160
161 **2.2. Sample collection and filtration**
162 The snow samples were collected aseptically using sterile plastic bags (WhirlPak; Nasco
163 Products), and melted onsite in a warm water bath; conditions in the field precluded melting the samples
164 at a controlled 4°C. The water samples were collected in either combusted glass or acid-cleaned Teflon
165 bottles. All samples were filtered on-site through 0.2-μm filters prior to extraction, except for the Red
166 Snow sample, which was processed back in the laboratory. Most samples (Yellow Snow, Supraglacial
167 Inland, Supraglacial Margin, Subglacial May, Tarn) were filtered using 0.2-μm Sterivex cartridges
168 (Millipore), that had been pre-cleaned by soaking in a 10% HCl bath for at least one day, followed by
169 rinsing with 20 L of Milli-Q water. The background DOC concentration of the pre-cleaned units was
170 approximately 9 μM. Due to limited availability of pre-cleaned Sterivex units in the field, the remaining
171 samples (Red Snow, Subglacial July-1, Subglacial July-2) were filtered through a combusted GFF
172 (Whatman) pre-filter and a combusted 0.2-μm Anodisc membrane (Whatman). All solvents were
173 purchased from Thermo Fisher Scientific (Waltham, MA) and were Optima grade or better. Concentrated

174 HCl was Trace-Metal grade. The final volumes of 0.2-μm filtrate (Table 1) differed to accommodate a
175 range of anticipated DOC contents as well as the difficulties encountered with filtering some samples (for
176 example, Subglacial May contained a significant amount of rock flour that quickly clogged the filters).
177 An aliquot of the 0.2-μm filtrate was acidified and stored in a combusted vial for DOC analysis.
178
179 **2.3. Solvent extraction**
180 Immediately following 0.2-μm filtration, all samples were acidified to pH 3 with 12M HCl and
181 dissolved organic matter (DOM) was extracted with either C_{18} cartridges (Mega Bond Elut, UTC) or C_{18}
182 extraction discs (3M) (Table 1). All of the solvent extractions except for the Subglacial July and Red
183 Snow samples were done on-site. The Subglacial July and Red Snow samples were kept as cold as
184 possible, and extracted approximately two months later. The solvent extraction protocol employed was
185 modified from Kim et al. (2003b). Briefly, the cartridges or discs were pre-cleaned according to
186 manufacturer's instructions. The acidified sample was then passed through the cleaned cartridge/disc and
187 the cartridge/disc was left to dry for 15 minutes prior to solvent extraction with methanol (MeOH) (Table
188 1). Extracts were evaporated to dryness under vacuum at 30°C. For Red Snow, the 70% and 100% MeOH
189 aliquots were combined prior to vacuum evaporation. A procedural blank (MeOH) was also evaporated to
190 dryness under vacuum. The samples and solvent blank were stored dry at -20°C until further analysis. We
191 estimated our DOM extraction efficiency by drying an aliquot of the solvent extract on a pre-weighed
192 combusted GFF, and measuring the carbon by dynamic flash combustion on a ThermoQuest EA1112
193 Carbon/Nitrogen Analyzer. The extraction efficiency for each sample was calculated as the percent of
194 carbon recovered from the solvent extract relative to the total amount of carbon in the sample (as
195 determined by TOC analysis). The extraction efficiencies (Table 1) ranged from 10% to 94%, with a
196 mean of 44% and a median of 28%. Although we obtained a low extraction efficiency (10%) for the
197 Subglacial July-2 sample, we do not anticipate being limited in our conclusions since this sample is
198 duplicated by Subglacial July-1 and the mass spectral characteristics of the two samples are nearly

identical (see Section 3.2 and Figure 3). The Tarn sample was the one most similar to previously described freshwater samples and the extraction efficiency of this sample (60%) is well within the range documented to other freshwater studies (Dittmar et al., 2008; Kim et al., 2003b).

2.4. DOC Concentrations

Total and dissolved organic carbon (TOC, DOC) concentrations were quantified as non-purgeable organic carbon (NPOC) by high temperature combustion (680°C) with a Shimadzu TOC-V_{CSH} analyzer equipped with a high sensitivity platinum catalyst (Shimadzu Scientific Instruments). Samples were quantified using a 5-point standard curve made with potassium hydrogen phthalate (KHP). Blanks and reference standards were analyzed routinely within each sample run. Reference standards for low carbon water and deep-sea water were obtained from the Consensus Reference Materials Project, Hansell Laboratory, University of Miami. DOC was not quantified for the 2007 samples due to post-acquisition contamination in Greenland.

2.5. FT-MS data acquisition

All samples and the solvent blank were analyzed on a 7-T ESI FT-ICR mass spectrometer (LTQ-FT-MS, Thermo Fisher Scientific, Waltham MA). For positive ion mode analyses, sample aliquots were reconstituted in 80% MeOH with 0.1% acetic acid (final concentration). Acetic acid promoted positive ion formation. For negative ion mode analyses, reconstituted sample aliquots were reconstituted in 70% MeOH. The solvents used to dilute the samples were also analyzed as instrument blanks (100% MeOH in positive ion mode and 70% MeOH in negative ion mode).

For both positive and negative ion modes, samples were infused into the ESI interface at 4 μL min^{-1}, and instrument parameters were optimized for each sample. Samples were diluted to optimize spray conditions; dilutions ranged from 1:5 to 1:40. The capillary temperature was set at 250°C, and the spray voltage varied between 4.40-4.60 kV. About 200 scans were collected for each sample, a sufficient number of scans for peak reproducibility in our samples (Kido Soule, Longnecker, Giovannoni, and

225 Kujawinski, unpublished data). The mass ranges for full-scan collection were 200 < *m/z* < 1200 and 200
226 < *m/*z < 1000 in positive and negative ion modes, respectively. Weekly mass calibrations were performed
227 with an external standard (Thermo Calibration Mix), and resulted in mass accuracy errors < 1 ppm. The
228 target average resolving power was 400,000 at *m/z* 400 (where resolving power is defined as $m/\Delta m_{50\%}$
229 where $\Delta m_{50\%}$ is the width at half-height of peak m). Good quality data could not be collected for the
230 Subglacial July-2 sample in positive ion mode, nor for the Red Snow sample in negative ion mode. This
231 was due to unacceptable spray stabilities in the former and fluctuating ion currents in the latter.
232
233 **2.6. FT-MS data analysis**
234 *2.6.1. Peak Detection and Blank Correction*
235 We collected individual transients as well as a combined raw file using *xCalibur 2.0*. Transients
236 were co-added and processed with custom-written MATLAB code provided by Southam et al. (2007).
237 This code was used as provided with the following parameters. Within each sample, only those transients
238 whose total ion current (TIC) was greater than 20% of the maximal TIC were co-added and then
239 processed with Hanning apodisation, and zero-filled once prior to fast Fourier transformation. We
240 retained all *m/z* values with a signal-to-noise ratio above 5 (as calculated in Southam et al. (2007)). The
241 individual sample and solvent blank peak lists were then aligned using MATLAB code provided by
242 Mantini et al. (2007). Positive and negative ion mode data were aligned separately in MATLAB with an
243 error tolerance of 1 ppm. Following alignment, all peaks found in each mode's solvent blanks were
244 removed from the appropriate master list. These blank-corrected master peak lists in each sample were
245 used in all downstream statistical analyses and elemental formula assignments.
246
247 *2.6.2. Calibration*
248 Positive and negative ion mode spectra were internally re-calibrated using a short list of *m/z*
249 values present in a majority of samples. This list of calibrants was chosen according to the following
250 criteria: (1) presence in the majority of samples; (2) elemental formulae could be assigned with C, H, O

and N; (3) similar mass errors for all; and (4) distribution along the *m/z* range of each spectrum. The resulting calibrants and their elemental formulae are provided in EA Tables 1a and 1b. After internal re-calibration, the root mean square (RMS) errors for the calibrants ranged from 0.09 to 0.12 in positive ion mode and 0.04 to 0.69 in negative ion mode.

2.6.3. Elemental Formula Assignments

Elemental formulae were assigned to the aligned blank-corrected peaks (*m/z* values) using the Compound Identification Algorithm (CIA), described by Kujawinski and Behn (2006) and modified in Kujawinski et al. (2009). In the CIA, we set the following parameters: (a) formula error was 1 ppm, (b) the relationship error was 20 ppm, and (c) the mass limit above which elemental formulae were only assigned by functional group relationships was 500 Da. For this study, elemental formulae were determined for *m/z* values below 500 Da by comparison to an in-house database of mathematically and chemically legitimate formulae within the 1 ppm error window. Elemental formula assignments were constrained to ^{12}C, ^{13}C, ^{1}H, ^{16}O, ^{14}N, ^{34}S, and ^{31}P. Error testing for formula assignments containing these elements was done using synthetic datasets and is documented in Kujawinski and Behn (2006). Accuracy of formula assignments ranges from 78% to 100%, depending on included elements (Kujawinski and Behn, 2006). These elemental formulae were extended to *m/z* values above 500 Da through identification of functional group relationships. The functional group relationships used by CIA are common to refractory dissolved organic matter (e.g. humic acids); CIA does not presently include many functional group relationships resulting from metabolic (biological) reactions (Kujawinski and Behn, 2006). Isotopomers with a ^{13}C atom are identified in the last step of CIA and elemental formulae are corrected to reflect ^{13}C content. In order to identify terrestrially-derived components of our samples, we compared the elemental formulae for our Greenland samples with those assigned to Suwannee River Fulvic Acid Standard I (Suwannee River - International Humic Substances Society, Stock #1S101F), previously analyzed in our laboratory with negative ion mode ESI FT-ICR MS. Magnitude-averaged elemental ratios and double bond equivalencies were calculated (Sleighter and Hatcher, 2008).

277

278 *2.6.4. Assessment of Potential Contamination*

279 Analysis of the negative and positive ion mode mass spectra revealed potential contamination
280 likely originating from plasticizers or the C_{18} extraction cartridges/discs. In negative ion mode, potential
281 contamination was most prevalent in the Yellow Snow sample. We assigned elemental formulae to the
282 contaminated *m/z* values (18 peaks) and identified peaks belonging to this series in other negative ion
283 mode spectra. Contaminated peaks did not occupy any particular region of the van Krevelen diagram (EA
284 Figure 1). We realize that any contamination may skew the overall composition of the DOM through ion
285 suppression; nonetheless, we believe we attained an adequate representation of DOM composition within
286 our samples because the maximum percentage of peaks represented by the suspected plasticizer
287 contamination was less than 0.6% in any one sample. In addition, to further minimize the potential impact
288 of this contamination, we based our statistical analyses and subsequent conclusions on the diversity of
289 resolved peaks (presence/absence) rather than on their relative peak heights. In positive ion mode, the
290 potential contamination was more pervasive. Inspection of the raw mass spectra revealed likely
291 contamination in the Yellow Snow, Subglacial May, and Tarn samples. Given this observation, we
292 focused our statistical analyses and interpretations on the negative ion mode dataset.

293

294 *2.6.5. Multivariate Statistics*

295 We assessed differences in our samples in negative ion mode with cluster analysis as described in
296 Kujawinski et al. (2009). In our analysis, we transformed all relative peak heights to presence (peak
297 height = 1) or absence (peak height = 0). We recognize that ESI is not quantitative and that differences in
298 ionization efficiencies among compounds can lead to misrepresentations of ion peak height, relative to the
299 abundance of the parent molecule in neutral solution (Stenson et al., 2003). To circumvent this known
300 problem, we have used presence / absence comparisons rather than those that rely on relative peak height.
301 The presence/absence transformation allows assessment of how samples differ based solely on peak
302 diversity. A distance matrix was calculated between all the samples in each mode using the Bray-Curtis

303 distance measure (MATLAB code written by David Jones, University of Miami, as part of the Fathom
304 toolbox); a distance measure of 0 indicates samples are identical with regards to peak diversity, whereas a
305 distance measure of 1 indicates that samples share none of their peaks. Ward's linkage method was used
306 to group the samples followed by presentation of the results as a dendrogram.
307
308 *2.6.6. Indicator Species Analysis*
309 We identified specific *m/z* values characteristic of the observed negative ion mode cluster
310 groupings with Indicator Species Analysis (ISA - as implemented in Kujawinski et al., 2009). ISA
311 combines the relative abundance and relative frequency of a peak within a pre-defined group of samples
312 to assign an indicator value (IV) to each peak (McCune and Grace, 2002). A perfect IV (equal to 100) of a
313 particular group would constitute an *m/z* value that was present exclusively in the samples comprising that
314 group (McCune and Grace, 2002). Statistical significance of IVs is calculated by comparison with Monte-
315 Carlo simulations of randomized data. ISA requires a priori assignment of samples to groups; this was
316 achieved using the protocol and criteria described in McCune and Grace (2002). The best number of
317 groups occurred when we used four groups of samples: Group 1 = Yellow Snow; Group 2 = Supraglacial
318 Inland; Group 3 = 'N' glacier May samples (Subglacial May and Supraglacial Margin); and Group 4 =
319 'N' glacier July and Tarn samples (Subglacial July-1,2 and Tarn). This group assignment was used to find
320 indicator *m/z* values for Groups 3 and 4; use of ISA is restricted to those groups with more than one
321 sample, thus no 'indicator peaks' were identified for Groups 1 and 2. The final list of indicator *m/z* values
322 for each group was manually curated using the criteria outlined in Kujawinski et al. (2009).
323
324 **3. RESULTS AND DISCUSSION**
325 **3.1. Sample overview**
326 The eight samples analyzed in this study represent carbon pools associated with different regions
327 of a glacier system. The supraglacial pools are represented by snow (Yellow Snow, Red Snow) and
328 meltwater (Supraglacial Inland) samples from the inland ice surface as well as the meltwater sample

329 collected on the surface of 'N' glacier (Supraglacial Margin). The subglacial pool at the glacier base is
330 represented by samples collected from the subglacial stream exiting at the base of 'N' glacier (Subglacial
331 May, Subglacial July-1,2). Since surface ice melting is minimal in May, the Subglacial May water sample
332 most likely represents early/spring discharge waters that have been stored at the bed overwinter. These
333 waters likely drain a more distributed subglacial hydrological system with relatively slower flow rates,
334 but they may access a greater areal extent of the subglacial bed (Nienow et al., 1998; Sharp et al., 1999).
335 Conversely, the July subglacial water samples represent late/summer discharge waters fed primarily by
336 supraglacial inflow. These waters likely drain through a channelized hydrological system characterized by
337 relatively much higher flow rates, but they may access a more limited part of the bed (Bingham et al.,
338 2005; Nienow et al., 1998). The electrical conductivity (EC) measurements (Table 1) support this
339 interpretation. The Subglacial May sample has a greater content of dissolved solutes compared to the
340 Subglacial July samples. Finally, a proglacial tarn (Tarn) represents a terrestrial carbon end-member,
341 comprised of non-glacial water, situated in the deglaciated arctic tundra and likely containing a large
342 terrestrial contribution from the surrounding vegetation.
343
344 **3.2. Comparison of Ultra-high Resolution Mass Spectra**
345 All of the samples contained highly complex DOM with numerous peaks per nominal mass in
346 both positive and negative ion modes. The total numbers of peaks resolved in each sample in negative ion
347 mode following blank correction are presented in Table 2. Qualitative differences among the raw mass
348 spectra illustrate that samples representing different regions of the Greenland ice sheet have distinct DOM
349 compositions (Figure 2). Although ultra-high resolution mass spectrometry has not been used to date to
350 compare DOM from different glacial sub-environments, this result is not surprising since both bulk DOC
351 concentrations and *in situ* microbial communities can differ vastly among glacial sub-environments
352 (Bhatia et al., 2006).
353 Cluster analysis based on the presence/absence of resolved peaks in negative ion mode (Figure 3)
354 revealed that the samples collected on the inland ice sheet (Yellow Snow, Supraglacial Inland) were

355 distinct from each other as well as from those collected at the ice sheet margin (Subglacial May,
356 Supraglacial Margin, Tarn, Subglacial July-1,2). Indeed, the Yellow Snow and Supraglacial Inland
357 samples share very few peaks (< 20%) with any of the samples collected at the ice margin (Table 3). The
358 cluster analysis for positive ion mode data (not shown) confirmed that the three samples from the inland
359 ice sheet surface (Yellow Snow, Red Snow, and Supraglacial Inland) were distinct from the ice margin
360 samples (Subglacial May, Supraglacial Margin, Tarn, Subglacial July-1). Differentiation between these
361 sample groups is expected since the Yellow Snow and Red Snow should represent very different,
362 localized regions on the ice sheet surface with unique algal and microbial communities. The lack of
363 similarity between the supraglacial meltwater samples (Supraglacial Inland and Supraglacial Margin, only
364 sharing 13% and 10% of their peaks respectively, Table 3) could be attributed to geographical, seasonal
365 and water source differences. For example, the Supraglacial Inland sample was collected from a large
366 supraglacial lake composed almost entirely of inland ice melt. In contrast, the Supraglacial Margin sample
367 was collected from a small meltwater pool closer to the ice edge and much earlier in the melt season, and
368 thus is comprised of a mixture of marginal snow and ice melt.
369 Among the margin-site samples, results from the cluster analyses for positive and negative ion
370 modes indicate that the DOM composition in the subglacial runoff changes during the melt season.
371 Specifically, the negative ion mode cluster analysis illustrates that the 'N' glacier May samples
372 (Supraglacial Margin and Subglacial May) were grouped (sharing 42% and 56% of their peaks
373 respectively, Table 3) as were the Subglacial July-1,2 and Tarn samples (Subglacial July samples sharing
374 73-79% of their peaks with the Tarn sample, Table 3). Interestingly, the Subglacial July samples are quite
375 distinct from the Subglacial May sample even though the two samples were collected from the same
376 location. In addition, there is significant peak overlap between Suwannee River and the Tarn sample
377 (70%) and between the Subglacial July samples (57-62%), but much less between Suwannee River and
378 the Subglacial May sample (30%). Thus, although our samples are temporally limited (May and July), we
379 infer that the type of DOM in subglacial discharge changed during the 2008 melt season.
380

3.3. Elemental Formula Assignments and Indicator Species Analysis

We were able to assign formulae to over 90% of the resolved peaks in the Suwannee River and the Tarn, Subglacial July-1,2, and Subglacial May samples. We achieved slightly lower percentages of formulae assigned to the Yellow Snow (86%) and Supraglacial Margin (85%) samples, with the lowest percentage of formulae found for the Supraglacial Inland sample (63%). In an effort to increase the percentage of formula assignments in this sample, we made two temporary modifications to CIA. First, we included halogens (F, Cl, Br, and I) in our formula assignments; and second, we attempted to account for multiply-charged molecules. Inclusion of halogens did not increase our formula assignment rate appreciably. In contrast, corrections for doubly- and triply-charged molecules produced a marked increase in the Supraglacial Inland formula assignment percentage (up to 98%), suggesting that a good portion of our *m/z* values represented multiply-charged molecules with multiple de-protonation sites. We discarded these improvements, however, because the modified CIA lowered the formula assignment accuracy when tested with Suwannee River formulae and because multiply-charged isotopomers were rarely available for reliable charge-state determination. Thus, we were forced to retain the original lower formula assignment percentages made to the Supraglacial Inland sample.

Elemental formulae containing only C, H, and O dominated the formula assignments for the Tarn and subglacial samples (Subglacial May and Subglacial July-1,2) (Table 2). Conversely, the supraglacial samples were dominated by formulae containing C, H, O, and N (Supraglacial Margin), or C, H, O, N, and S/P (Yellow Snow, Supraglacial Inland) (Table 2). We should note that this result differs from analysis of other supraglacial organic material in ice cores collected from Russia where formulae containing C, H, and O were the most abundant (Grannas et al., 2006). However, the snow and meltwater samples analyzed in this study (i.e., collected from marginal areas where there is snow melt and water in the residual snowpack) are quite different from bulk ice core material (i.e., collected from inland areas where ice is formed in the dry snow zone), so it is not surprising that we resolved different compounds.

For comparison with other DOM compositional studies, we calculated the magnitude-averaged bulk elemental ratios and double-bond equivalency (DBE) for all samples (Table 2) (Koch et al., 2008;

Sleighter and Hatcher, 2008). Molecular H:C and O:C ratios have been reported previously to range broadly from 0.3-1.8 and 0-0.8, respectively (Koch et al., 2008; Sleighter and Hatcher, 2008; Stenson et al., 2003). The elemental ratios of all of our samples fall within this range (Table 2), with Suwannee River being the most aromatic (H:C = 1.05), and the Subglacial May and Supraglacial Margin samples being the most aliphatic (H:C = 1.68 and 1.56, respectively). The low DBE of the Supraglacial Margin and Subglacial May samples also imply that DOM in these samples is relatively aliphatic. The DBE was the highest in the Supraglacial Inland sample. This fact, combined with the relatively lower H:C ratio (1.16) and relatively higher N:C ratio (0.33) of this sample (Table 2), suggest that molecules within this sample may contain condensed nitrogen functionalities (i.e., aromatic nitrogen or nitro groups). Finally, the supraglacial samples (Yellow Snow, Supraglacial Inland, Supraglacial Margin) generally had relatively high N:C ratios (0.30, 0.33, 0.27 respectively, Table 2), suggesting that nitrogen-containing molecules could be major contributors to DOM in these samples (Reemtsma, 2008).

Van Krevelen diagrams were generated for all Greenland samples and Suwannee River in order to compare DOM composition across our samples (representative sample plots in Figure 4). Van Krevelen diagrams illustrate the O:C molar ratio and the H:C molar ratio of each elemental formula on the x- and y-axes, respectively. Generally, major biogeochemical compound classes (such as condensed hydrocarbons, lipids, proteins, lignins, and carbohydrates) have characteristic H:C and/or O:C molar ratios, and thus should occupy specific regions of the plot (Kim et al., 2003a; Kujawinski and Behn, 2006; Wu et al., 2004). The percentages of negative ion mode formula assignments located in the different regions of the van Krevelen diagram are presented in Table 4. However, we should note that van Krevelen diagrams should be interpreted with caution as inconsistent definitions of particular compound classes across the literature (e.g., lipid), and variable O:C or H:C ratios within particular compound classes (e.g., proteins) may lead to exclusion of elemental formulae from the prescribed compound class regions (Kujawinski and Behn, 2006). Nonetheless, at present, they remain the best way to graphically depict elemental formula assignments for mass spectra comprised of thousands of peaks.

The van Krevelen plot of the negative ion mode Suwannee River sample (not shown) is consistent with previous work (Stenson et al., 2003). Over 99% of formulae were assigned and most occur in the region associated with lignin-derived materials (Stenson et al., 2003). Very few formulae are present in the regions associated with proteins and lipids (Table 4). Because of these results and the fact that Suwannee River is well-cited as a terrestrial DOM end-member (e.g. McKnight et al., 2001; Stenson et al., 2003), we label the region encompassing the majority of its elemental formula assignments as "terrestrial" (shown in Figures 4 and 5), and use this information to aid our analyses of our negative ion mode spectra.

The van Krevelen diagrams may explain the observed cluster groupings in Figure 3. In negative ion mode, the separation between the samples collected on the ice sheet surface and those collected at the margin may be the result of the Yellow Snow and Supraglacial Inland samples having a greater representation in the condensed hydrocarbon region and a lower proportion in the lignin region (Table 4). The grouping of the Subglacial May and Supraglacial Margin samples may be due to greater proportions of protein-like and lipid-like material in these samples compared to the remainder of the dataset (Table 4). The grouping of the Tarn and Subglacial July samples results from a commonality in every region of the van Krevelen plot, particularly in the terrestrial Suwannee River and lignin regions (Table 4).

Apart from these general trends, each sample also has some noteworthy features on the van Krevelen diagram. In addition to a large protein-like component, the Supraglacial Margin sample also contains more formulae in the lipid and the condensed hydrocarbon regions than the Subglacial May sample (Table 4). Even though both the Supraglacial Margin and Subglacial May samples contain lignin-like molecules, the Subglacial May sample has a larger proportion of formulae in the "terrestrial" Suwannee River region (Table 4). The Tarn and Subglacial July samples all contain a larger proportion of formulae in the condensed hydrocarbon and protein regions than the Suwannee River sample (Table 4). The results of our analyses of the van Krevelen plots for each of the samples are summarized in Figure 5.

Indicator species analysis revealed that a higher content of biologically-derived elemental formulae is responsible for the differentiation of the Subglacial May and Supraglacial Margin samples

458 (Group 3) from the Tarn and Subglacial July samples (Group 4). Indicator *m/z* values for the Group 3
459 samples are dominated by high H:C compounds occupying the protein region of the van Krevelen
460 diagram (Figure 6a). Conversely, the indicator *m/z* values for the Group 4 samples are dominated by low
461 H:C compounds found in the terrestrial Suwannee River region. There is a significant terrestrial
462 component within all the ice margin samples, as evidenced by the presence of indicator *m/z* values
463 common to groups 3 and 4 (yellow dots, Figure 6b) in this region. This component is absent in the
464 samples collected on the inland ice sheet surface (Yellow Snow and Supraglacial Inland, Groups 1 and 2).
465

466 **3.4. Potential Sources of Observed Peaks**

467 *3.4.1. Microbially-derived material (lipid-like and protein-like signatures)*

468 Similar to previous fluorescence studies (Barker et al., 2006; Lafreniere and Sharp, 2004), the
469 distinct microbial character of the Supraglacial Margin sample (reflected by its high proportion of protein-
470 like formulae) is likely derived from photosynthetic algae and bacteria communities widely observed to
471 be present in supraglacial environments (Carpenter et al., 2000; Foreman et al., 2007; Grannas et al.,
472 2004). The presence of lipid-like material in the Supraglacial Margin sample also correlates well with
473 previous work identifying biologically-derived lipids in organic matter from snow collected at Summit
474 atop the Greenland ice sheet (Grannas et al., 2006; Grannas et al., 2004).

475 Early season (spring) subglacial waters have also been observed to have a microbial fluorescence
476 signature (Barker et al., 2006; Lafreniere and Sharp, 2004), despite the fact that terrestrial carbon from
477 overridden soils and vegetation is also present at the glacier base (Sharp et al., 1999). The larger
478 proportion of protein-like formulae in the early season subglacial waters (Subglacial May) may reflect *in*
479 *situ* subglacial microbial metabolism of some component of the subglacial organic carbon stores during
480 over winter storage (Tranter et al., 2005). The May subglacial water likely drains a broad distributed
481 hydrological network along the ice-bed interface, and consequently experiences prolonged storage at the
482 bed where active subglacial microbial communities are thought to be present (Tranter et al., 2005).
483 Although no study has documented the presence of subglacial communities beneath the Greenland ice

484 sheet specifically, a mounting body of literature indicates that large, active microbial communities are
485 present beneath glaciers in diverse regions on varying lithologies (the Swiss Alps, southern New Zealand
486 Alps, Alaska, Svalbard, Antarctica, and the Canadian high Arctic) (Lanoil et al., 2009; Mikucki et al.,
487 2009; Sharp et al., 1999; Skidmore et al., 2000). Furthermore, studies show that the abundances of
488 subglacial communities (as high as 1.8×10^9 cells g^{-1}) are similar to the highest microbial abundances in
489 permafrost (10^7-10^9 cells g^{-1}) (Sharp et al., 1999). Documented subglacial communities include
490 heterotrophic bacteria (e.g., aerobic respirers, nitrate- and sulfate-reducers) as well as autotrophic bacteria
491 (e.g., methanogens) (Cheng and Foght, 2007; Foght et al., 2004; Skidmore et al., 2000). The existence of
492 numerically-abundant, enduring biological communities implies that any microbially-mediated
493 biogeochemical activities occur on a continuous temporal basis. The diverse DOM composition in the
494 Subglacial May sample is consistent with the idea of high subglacial microbial activity due in particular to
495 its significant protein and terrestrial components (Table 4).
496
497 *3.4.2. Terrestrial-derived material (lignins and Suwannee River-like components)*
498 Lignins and formulae located in the "terrestrial" region of our van Krevelen plots are likely
499 derived from previously overridden soils and vegetation (subglacial samples) or surrounding terrestrial
500 soils and vegetation (Tarn). The large component of terrestrially derived DOM in the Tarn sample
501 (overlap between Suwannee River and the Tarn sample is 70%), is likely derived from its location in the
502 developed soils and vegetation at our study site. In contrast, the subglacial samples contain terrestrially-
503 derived DOM, present in both May and July, that is most likely derived from previously overridden soils
504 and vegetation during glacial advance. The lack of lignin material in the samples collected on the inland
505 ice sheet surface (Yellow Snow, Supraglacial Inland) suggests that organic matter from these
506 environments is not influenced significantly by non-charred terrestrial inputs. This is in contrast to
507 Grannas et al. (2004) who noted the presence of vascular plant tissue (i.e., lignin) in snow collected from
508 Summit, Greenland.
509

510 *3.4.3. Condensed hydrocarbons*

511 Condensed hydrocarbons are generally compounds with a deficiency in both oxygen and
512 hydrogen and often contain aromatic ring structures. Previous studies have illustrated that these
513 compounds originate from black carbon-like molecules (Kim et al., 2004), and could be derived from
514 atmospheric deposition of soot particles (Slater et al., 2002). Evidence of these compound types is present
515 in all ice sheet surface samples (Yellow Snow, Supraglacial Inland, Supraglacial Margin) and the
516 late/summer discharge samples (Subglacial July-1,2). On the ice sheet surface, this material likely
517 originates from atmospheric deposition of combustion products. We do not anticipate a novel source of
518 condensed hydrocarbons in the subglacial environment. Rather, the presence of condensed hydrocarbons
519 in late season subglacial waters (Subglacial July-1,2) may reflect either (1) the increased contribution of
520 supraglacial meltwater to the subglacial outflow at the peak of the summer melt season, or (2) an
521 increased flux of condensed hydrocarbons from the ice sheet surface after the snow cover has melted.
522 Support for this second hypothesis may be provided by Clarke and Noon (1985), who found that soot may
523 be enriched in Arctic snowmelt compared to the snowpack.

524

525 **3.5. Implications for understanding subglacial flow regimes**

526 The fact that the late season subglacial waters still possess an overwhelming terrestrial signature
527 may reflect the ability of the summer hydrological flow regime to mobilize subglacial organic carbon
528 stores. As the melt season progresses on the Greenland ice sheet, meltwater from seasonal snow and ice
529 collects in streams and lakes on the ice sheet surface. The majority of this surface meltwater is thought to
530 descend to the bed via crevasses and moulins at the peak of the summer melt season (Das et al., 2008;
531 Krawczynski et al., 2009). Thus, the late season subglacial waters are primarily comprised of supraglacial
532 inflow passing rapidly through the subglacial environment. Over the course of a melt season, the ice sheet
533 subglacial drainage system is predicted to evolve from a distributed to a more channelized network
534 facilitating rapid water flow to the glacier front, similar to what has been observed in alpine glacier
535 systems (Nienow et al., 1998). The faster flow rates characteristic of this channelized system do not

536 permit extensive water-sediment interaction, thus minimizing the impact of *in situ* microbial metabolism
537 (Tranter et al., 2005). Additionally, the larger volumes of water passing through the subglacial system
538 may facilitate turbulent incidental contact that allows the meltwaters to mobilize terrestrial sources of
539 DOC at the glacier base (i.e., previously overridden soil and vegetation). Previous work in alpine
540 catchments has illustrated that suspended sediment concentrations increase throughout a melt season as
541 sediment sources are accessed by an extending and integrating subglacial drainage network (Clifford et
542 al., 1995; Richards et al., 1996). This reasoning is also consistent with previous fluorescence spectroscopy
543 work by Barker et al. (2006) at a polythermal Canadian high Arctic glacier, which showed that the late
544 season subglacial meltwaters bear a terrestrially-derived signature. The change in subglacial flow rate
545 may explain why condensed hydrocarbons are not present in the early season subglacial waters. Increased
546 residence times of these waters at the glacier bed throughout the preceding winter would permit non-polar
547 hydrocarbon-like, soot-derived compounds to adsorb quantitatively to organic particles in the subglacial
548 environment (Kramer et al., 2004) and thus to be removed from discharge waters. At the peak of the
549 summer melt season, the higher meltwater flow rates and potentially elevated hydrocarbon concentrations
550 would preclude quantitative removal by adsorption, allowing the subglacial waters to retain these
551 compounds in the late season subglacial runoff.
552
553 **3.6. Implications for understanding glacial organic matter cycling**
554 The microbial signatures of the subglacial discharge samples analyzed in our study support the
555 suggestion that glacial systems supply labile material to downstream marine and terrestrial environments
556 (Barker et al., 2006; Hood et al., 2009; Lafreniere and Sharp, 2004) extending these results to an ice sheet
557 environment for the first time. This hypothesis follows earlier discoveries of abundant, active microbial
558 communities associated with supraglacial, subglacial, and proglacial environments (Anesio et al., 2009;
559 Bhatia et al., 2006; Sharp et al., 1999). It has been substantiated by direct investigations of glacially-
560 derived DOM, including fluorescence spectrometry (Barker et al., 2006; Lafreniere and Sharp, 2004),
561 compound specific analyses (i.e. lignin phenols) (Hood et al., 2009), and bulk organic carbon

562 characterizations (C:N ratios, $\delta^{13}C$ values) (Hood et al., 2009; Hood and Scott, 2008). Most recently,
563 Hood et al. (2009) demonstrated that the bioavailability of glacial organic carbon is indirectly correlated
564 with age, so that DOM from glaciated catchments is labile despite having ancient $\Delta^{14}C$ ages. Thus,
565 meltwater streams and rivers draining glaciated areas may potentially provide a significant, previously
566 overlooked source of labile reduced carbon to downstream ecosystems (Barker et al., 2006; Hood et al.,
567 2009). Our study corroborates these findings through a comprehensive molecular-level description of
568 glacially-derived DOM in meltwater runoff from the Greenland ice sheet and offers a novel line of
569 evidence that glacial DOM has a microbial source.
570
571 **4. Conclusions**
572 Previous studies illustrate that the majority of supraglacial DOM likely originates from
573 autochthonous microbial processes, whereas subglacial DOM contains both allochthonous carbon derived
574 from previously overridden soils and vegetation, and autochthonous carbon derived from *in situ* microbial
575 metabolism. Our findings support these provenances. Generally the supraglacial and early season
576 subglacial discharge had a higher proportion of protein-like and lipid-like elemental formulae, whereas
577 the tarn and late season subglacial water DOM had a higher proportion of lignin and terrestrial Suwannee
578 River-like materials. However, evolving subglacial flow regimes also likely exert a heavy influence on
579 the type of DOM present in the subglacial outflow at different times of the year. In this study, this
580 influence is reflected in a smaller terrestrial component in the early season subglacial waters, and the
581 detection of condensed hydrocarbon-like material in late season subglacial waters. Based on the samples
582 analyzed, the DOM composition of subglacial outflow shifts from a terrestrial to microbial signature over
583 winter storage and then back to a terrestrial signature through a melt season. We propose that this shift is
584 dependent on the degree of subglacial microbial metabolism that has occurred. However, additional
585 samples and measurements constraining the subglacial flow regime and resident microbial communities
586 are required to fully test the validity of this conjecture.

This study represents the first molecular-level analyses of subglacial organic carbon stores, and as such, has illustrated that ultrahigh resolution mass spectrometry can provide unprecedented compositional information regarding the interplay among different glacial carbon pools. In addition to these qualitative results, further work with both bulk and compound-specific measurements will be required to confirm that specific compound classes (e.g., proteins, lipids) are present and to constrain the temporal provenances of these pools. Nevertheless, our results suggest that a much more complex and reactive carbon system is associated with glacial environments than previously thought and merit further investigation, given the extent and frequency of glaciation events through Earth's history.

Acknowledgements. This research was supported by: the National Science Foundation (CAREER-OCE-0529101 (EBK), ARC-0520077 (SBD)), National Atmospheric and Space Administration (SBD), the WHOI Clark Arctic Research Initiative (EBK, SBD, MAC), the WHOI Ocean Ventures Fund (MPB), and the National and Science Engineering Research Council of Canada (MPB). We acknowledge M. Kido Soule for assistance with data collection and the funding sources of the WHOI FT-MS Users' facility (National Science Foundation OCE-0619608 and the Gordon and Betty Moore Foundation). We are grateful to I. Joughin, M. Behn, R. Harris, B. Gready, P. Henderson, A. Criscitiello, and M. Evans for their assistance in the field, to P. Henderson for conducting the carbon/nitrogen analyses, and to G. Wolken for his assistance in constructing maps of our field site. We also thank three anonymous reviewers whose comments improved the manuscript.

606 **Citations**

607 Anesio, A. M., Hodson, A. J., Fritz, A., Psenner, R., and Sattler, B., 2009. High microbial activity on
608 glaciers: importance to the global carbon cycle. *Global Change Biol.* **15**, 955-960.
609 Azam, F., 1998. Microbial control of oceanic carbon flux: The plot thickens. *Science* **280**, 694-696.
610 Bamber, J. L., Layberry, R. L., and Gogineni, S., 2001. A new ice thickness and bed data set for the
611 Greenland ice sheet 1. Measurement, data reduction, and errors. *J. Geophys. Res. D: Atmos.* **106**,
612 33773-33780.
613 Barker, J. D., Sharp, M. J., Fitzsimons, S. J., and Turner, R. J., 2006. Abundance and dynamics of
614 dissolved organic carbon in glacier systems. *Arct. Antarct. Alp. Res.* **38**, 163-172.
615 Berner, R. A., Lasaga, A. C., and Garrels, R. M., 1983. The carbonate-silicate geochemical cycle and its
616 effect on atmospheric carbon-dioxide over the past 100 million years *Am. J. Sci.* **283**, 641-683.
617 Bhatia, M., Sharp, M., and Foght, J., 2006. Distinct bacterial communities exist beneath a high arctic
618 polythermal glacier. *Appl. Environ. Microbiol.* **72**, 5838-5845.
619 Bingham, R. G., Nienow, P. W., Sharp, M. J., and Boon, S., 2005. Subglacial drainage processes at a
620 High Arctic polythermal valley glacier. *J. Glaciol.* **51**, 15-24.
621 Brown, G. H., 2002. Glacier meltwater hydrochemistry. *Appl. Geochem.* **17**, 855-883.
622 Carpenter, E., Lin, S., and Capone, D., 2000. Bacterial activity in South Pole snow *Appl. Environ.*
623 *Microbiol.* **66**, 4514-4517.
624 Cheng, S. M. and Foght, J. M., 2007. Cultivation-independent and -dependent characterization of Bacteria
625 resident beneath John Evans Glacier. *FEMS Microbiol. Ecol.* **59**, 318-330.
626 Chillrud, S. N., Pedrozo, F. L., Temporetti, P. F., Planas, H. F., and Froelich, P. N., 1994. Chemical
627 weathering of phosphate and germanium in glacial meltwaters: Effects of subglacial pyrite
628 oxidation *Limnol. Oceanogr.* **39**, 1130-1140.
629 Clarke, A. D. and Noone, K. J., 1985. Soot in the Arctic snowpack: A cause for pertubations in radiative
630 transfer. *Atmos. Environ.* **19**, 2045-2053.
631 Clifford, N. J., Richards, K. S., Brown, R. A., and Lane, S. N., 1995. Scales of variation of suspended
632 sediment concentration and turbidity in a glacial meltwater stream *Geografiska Annaler Series a-*
633 *Physical Geography* **77A**, 45-65.
634 Das, S. B., Joughin, I., Behn, M. D., Howat, I. M., King, M. A., Lizarralde, D., and Bhatia, M. P., 2008.
635 Fracture propagation to the base of the Greenland Ice Sheet during supraglacial lake drainage.
636 *Science* **320**, 778-781.
637 Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G., 2008. A simple and efficient method for the solid-
638 phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology and*
639 *Oceanography-Methods* **6**, 230-235.
640 Eglinton, T. I. and Repeta, D. J., 2003. *Treatise on Geochemistry: Marine Organic Geochemistry*.
641 Foght, J., Aislabie, J., Turner, S., Brown, C. E., Ryburn, J., Saul, D. J., and Lawson, W., 2004. Culturable
642 bacteria in subglacial sediments and ice from two Southern Hemisphere glaciers. *Microb. Ecol.*
643 **47**, 329-340.
644 Foreman, C. M., Sattler, B., Mikucki, J. A., Porazinska, D. L., and Priscu, J. C., 2007. Metabolic activity
645 and diversity of cryoconites in the Taylor Valley, Antarctica. *J. Geophys. Res.-Biogeosci.* **112**,
646 11.
647 Grannas, A. M., Hockaday, W. C., Hatcher, P. G., Thompson, L. G., and Mosley-Thompson, E., 2006.
648 New revelations on the nature of organic matter in ice cores. *J. Geophys. Res. D: Atmos.* **111**.
649 Grannas, A. M., Shepson, P. B., and Filley, T. R., 2004. Photochemistry and nature of organic matter in
650 Arctic and Antarctic snow. *Global Biogeochem. Cycles* **18**.
651 Hedges, J. I., 1990. Compositional indicators of organic acid sources and reactions in natural
652 environments In: Perdue, E. M. and Gjessing, E. T. Eds.), *Organic Acids in Aquatic Ecosystems*.
653 John Wiley & Sons, Ltd. .
654 Hedges, J. I., Eglinton, G., Hatcher, P. G., Kirchman, D. L., Arnosti, C., Derenne, S., Evershed, R. P.,
655 Kogel-Knabner, I., de Leeuw, J. W., Littke, R., Michaelis, W., and Rullkotter, J., 2000. The

molecularly-uncharacterized component of nonliving organic matter in natural environments. *Org. Geochem.* **31**, 945-958.
Hood, E., Fellman, J., Spencer, R. G. M., Hernes, P. J., Edwards, R., D'Amore, D., and Scott, D., 2009. Glaciers as a source of ancient and labile organic matter to the marine environment. *Nature* **462**, 1044-U100.
Hood, E. and Scott, D., 2008. Riverine organic matter and nutrients in southeast Alaska affected by glacial coverage. *Nat. Geosci.* **1**, 583-587.
Kim, S., Kaplan, L. A., Benner, R., and Hatcher, P. G., 2004. Hydrogen-deficient molecules in natural riverine water samples--evidence for the existence of black carbon in DOM. *Mar. Chem.* **92**, 225-234.
Kim, S., Kramer, R. W., and Hatcher, P. G., 2003a. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **75**, 5336-5344.
Kim, S., Rodgers, R. P., and Marshall, A. G., 2006. Truly "exact" mass: Elemental composition can be determined uniquely from molecular mass measurement at similar to 0.1 mDa accuracy for molecules up to similar to 500 Da. *Int. J. Mass Spectrom.* **251**, 260-265.
Kim, S., Simpson, A. J., Kujawinski, E. B., Freitas, M. A., and Hatcher, P. G., 2003b. High resolution electrospray ionization mass spectrometry and 2D solution NMR for the analysis of DOM extracted by C-18 solid phase disk. *Org. Geochem.* **34**, 1325-1335.
Koch, B. P., Ludwichowski, K. U., Kattner, G., Dittmar, T., and Witt, M., 2008. Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. *Mar. Chem.* **111**, 233-241.
Koch, B. P., Witt, M. R., Engbrodt, R., Dittmar, T., and Kattner, G., 2005. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Acta* **69**, 3299-3308.
Kramer, R. W., Kujawinski, E. B., and Hatcher, P. G., 2004. Identification of black carbon derived structures in a volcanic ash soil humic acid by Fourier transform ion cyclotron resonance mass spectrometry. *Environ. Sci. Technol.* **38**, 3387-3395.
Krawczynski, M. J., Behn, M. D., Das, S. B., and Joughin, I., 2009. Constraints on the lake volume required for hydro-fracture through ice sheets. *Geophys. Res. Lett.* **36**.
Kujawinski, E. B., 2002. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex environmental mixtures. *Environ. Forensics* **3**, 207-216.
Kujawinski, E. B. and Behn, M. D., 2006. Automated analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra of natural organic matter. *Anal. Chem.* **78**, 4363-4373.
Kujawinski, E. B., Longnecker, K., Blough, N. V., Vecchio, R. D., Finlay, L., Kitner, J. B., and Giovannoni, S. J., 2009. Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochim. Cosmochim. Acta* **73**, 4384-4399.
Lafreniere, M. J. and Sharp, M. J., 2004. The concentration and fluorescence of dissolved organic carbon (DOC) in glacial and nonglacial catchments: Interpreting hydrological flow routing and DOC sources. *Arct. Antarct. Alp. Res.* **36**, 156-165.
Lanoil, B., Skidmore, M., Priscu, J. C., Han, S., Foo, W., Vogel, S. W., Tulaczyk, S., and Engelhardt, H., 2009. Bacteria beneath the West Antarctic Ice Sheet. *Environ. Microbiol.* **11**, 609-615.
Layberry, R. L. and Bamber, J. L., 2001. A new ice thickness and bed data set for the Greenland ice sheet 2. Relationship between dynamics and basal topography. *J. Geophys. Res. D: Atmos.* **106**, 33781-33788.
Lyons, W. B., Welch, K. A., and Doggett, J. K., 2007. Organic carbon in Antarctic snow. *Geophys. Res. Lett.* **34**.

Mantini, D., Petrucci, F., Pieragostino, D., Del Boccio, P., Di Nicola, M., Di Ilio, C., Federici, G., Sacchetta, P., Comani, S., and Urbani, A., 2007. LIMPIC: a computational method for the separation of protein MALDI-TOF-MS signals from noise. *BMC Bioinf.* **8**.

Marshall, A. G. and Rodgers, R. P., 2008. Petroleomics: Chemistry of the underworld. *Proc. Nat. Acad. Sci. U.S.A.* **105**, 18090-18095.

McCune, B. and Grace, J., 2002. *Analysis of Ecological Communities*. MjM Software Design, Gleneden Beach, Oregon.

McKnight, D. M., Andrews, E. D., Spaulding, S. A., and Aiken, G. R., 1994. Aquatic Fulvic Acids in Algal Rich Antarctic Ponds. *Limnol. Oceanogr.* **39**, 1972-1979.

McKnight, D. M., Boyer, E. W., Westerhoff, P. K., Doran, P. T., Kulbe, T., and Andersen, D. T., 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnol. Oceanogr.* **46**, 38-48.

Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A. D., Priscu, J. C., and Lee, P. A., 2009. A Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science* **324**, 397-400.

Nienow, P., Sharp, M., and Willis, I. C., 1998. Seasonal changes in the morphology of the subglacial drainage system, Haut Glacier d'Arolla, Switzerland. *Earth Surf. Process. Landf.* **23**, 825-843.

Raiswell, R., 1984. Chemical models of solute acquisition in glacial meltwaters *J. Glaciol.* **30**, 49-57.

Reemtsma, T., These, A., Linscheid, M., Leenheer, J., Spitzy, A. , 2008. Molecular and structural characterization of dissolved organic matter from the deep ocean by FTICR-MS, including hydrophilic nitrogenous organic molecules. *Environ. Sci. Technol.* **42**, 1430-1437.

Richards, K., Sharp, M., Arnold, N., Gurnell, A., Clark, M., Tranter, M., Nienow, P., Brown, G., Willis, I., and Lawson, W., 1996. An integrated approach to modelling hydrology and water quality in glacierized catchments. *Hydrol. Processes* **10**, 479-508.

Sharp, M., Parkes, J., Cragg, B., Fairchild, I. J., Lamb, H., and Tranter, M., 1999. Widespread bacterial populations at glacier beds and their relationship to rock weathering and carbon cycling. *Geology* **27**, 107-110.

Skidmore, M., Anderson, S. P., Sharp, M., Foght, J., and Lanoil, B. D., 2005. Comparison of microbial community compositions of two subglacial environments reveals a possible role for microbes in chemical weathering processes. *Appl. Environ. Microbiol.* **71**, 6986-6997.

Skidmore, M. L., Foght, J. M., and Sharp, M. J., 2000. Microbial life beneath a high Arctic glacier. *Appl. Environ. Microbiol.* **66**, 3214-3220.

Slater, J. F., Currie, L. A., Dibb, J. E., and Benner, B. A., 2002. Distinguishing the relative contribution of fossil fuel and biomass combustion aerosols deposited at Summit, Greenland through isotopic and molecular characterization of insoluble carbon. *Atmos. Environ.* **36**, 4463-4477.

Sleighter, R. L. and Hatcher, P. G., 2008. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar. Chem.* **110**, 140-152.

Southam, A. D., Payne, T. G., Cooper, H. J., Arvanitis, T. N., and Viant, M. R., 2007. Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Anal. Chem.* **79**, 4595-4602.

Stenson, A. C., Marshall, A. G., and Cooper, W. T., 2003. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **75**, 1275-1284.

Tranter, M., Sharp, M. J., Lamb, H. R., Brown, G. H., Hubbard, B. P., and Willis, I. C., 2002. Geochemical weathering at the bed of Haut Glacier d'Arolla, Switzerland - a new model. *Hydrol. Processes* **16**, 959-993.

Tranter, M., Skidmore, M., and Wadham, J., 2005. Hydrological controls on microbial communities in subglacial environments. *Hydrol. Processes* **19**, 995-998.

Wadham, J. L., Tranter, M., Tulaczyk, S., and Sharp, M., 2008. Subglacial methanogenesis: A potential climatic amplifier? *Global Biogeochem. Cycles* **22**.

Wu, Z., Rodgers, R. P., and Marshall, A. G., 2004. Two- and Three-Dimensional van Krevelen Diagrams: A Graphical Analysis Complementary to the Kendrick Mass Plot for Sorting Elemental Compositions of Complex Organic Mixtures Based on Ultrahigh-Resolution Broadband Fourier Transform Ion Cyclotron Resonance Mass Measurements. *Anal. Chem.* **76**, 2511-2516.

Table 1. Synopsis of the samples collected in this study in preparation for DOM extraction and mass spectrometry analysis. (*) - The [DOC] reported for Subglacial July-1 is from a sample collected 6 hours prior to the sample analyzed for DOM composition in this study. N/A = data not available.

Region	Sample	Collection Date	Location	Volume filtered	C_{18} DOM extraction	Solvent Extract	DOC concentration (μM)	Electrical Conductivity (μS/cm^3)	Extraction Efficiencies
Snow	Yellow Snow	July 17 2007	68°33'N 49°23'W	2 L	Cartridge	40 mL 100% MeOH	N/A	N/A	N/A
Snow	Red Snow	July 17 2007	68°34'N 49°22'W	87 mL	Discs	5 ml 70%, 5 mL 100% MeOH	N/A	N/A	N/A
Supraglacial	Supraglacial Inland	July 14 2007	68°34'N 49°21'W	15 L	Cartridge	40 mL 100% MeOH	N/A	N/A	N/A
Supraglacial	Supraglacial Margin	May 31 2008	68°02'N 50°15'W	4 L	Cartridge	15 mL 100% MeOH	16 ± 0.7	0.2	28%
Subglacial	Subglacial May	May 31 2008	68°02'N 50°16'W	500 mL	Cartridge	15 mL 100% MeOH	28 ± 0.2	17	94%
Proglacial	Tarn	May 29 2008	68°02'N 50°17'W	1 L	Cartridge	15 mL 100% MeOH	406 ± 3	N/A	57%
Subglacial	Subglacial July-1	July 12 2008	68°02'N 50°16'W	4.5 L	Cartridge	15 mL 100% MeOH	*15 ± 0.4	3.2	28%
Subglacial	Subglacial July-2	July 16 2008	68°02'N 50°16'W	3.45 L	Cartridge	15 mL 100% MeOH	51 ± 0.3	2.3	10%

Table 2. Synopsis of general parameters regarding negative ion mode formula assignments. Elemental ratios were calculated as magnitude-averaged values (Sleighter and Hatcher, 2008) for *m/z* values with assigned elemental formulae.

Sample	Total Number of Peaks	Number of Formulas Assigned	% Formulas Assigned	$H:C_w$	$O:C_w$	$N:C_w$	$S:C_w$	$P:C_w$	DBE_w	% Formulae With CHO	% Formulae With CHON	% Formulae With CHONP, CHONS, CHONSP
Yellow Snow	5113	4380	85.7	1.22	0.41	0.30	0.04	0.05	9.79	17.4	23.4	42.5
Supraglacial Inland	1865	1169	62.7	1.16	0.40	0.33	0.03	0.05	12.15	1.7	32.0	50.1
Supraglacial Margin	2331	1980	84.9	1.68	0.27	0.27	0.01	0.03	6.21	23.3	34.7	25.0
Subglacial May	1737	1662	95.7	1.56	0.38	0.17	0.00	0.01	6.66	55.6	26.1	11.3
Subglacial July-1	3330	3249	97.6	1.26	0.38	0.16	0.00	0.02	9.62	69.2	8.1	18.9
Subglacial July-2	3048	2800	91.9	1.24	0.38	0.21	0.00	0.02	10.08	58.9	10.8	26.2
Tarn	5958	5826	97.8	1.27	0.43	0.12	0.00	0.01	10.28	65.7	12.3	17.5
Suwannee River	2092	2079	99.4	1.05	0.55	0.03	0.00	0.01	10.85	91.3	2.0	4.5

774 **Table 3.** Percentage of negative ion mode peaks shared between the different samples analyzed in this study and Suwannee River.
775

Sample	Yellow Snow	Supraglacial Inland	Supraglacial Margin	Subglacial May	Subglacial July-1	Subglacial July-2	Tarn	Suwannee River
% Yellow Snow shared with	100	17	16	13	15	15	18	9
% Supraglacial Inland shared with	46	100	13	7	11	14	9	4
% Supraglacial Margin shared with	35	11	100	42	39	35	35	11
% Subglacial May shared with	39	7	56	100	61	56	60	36
% Subglacial July-1 shared with	23	6	27	32	100	73	79	39
% Subglacial July-2 shared with	24	9	27	32	79	100	73	39
% Tarn shared with	16	3	14	18	44	37	100	25
% Suwannee River shared with	22	4	12	29	62	57	70	100

776
777

Table 4. Percentage of negative ion mode formula assignments located in different regions of the van Krevelen diagram. Group numbers refer to groups determined by Indicator Species Analysis (see text for details).

Sample	Condensed Hydrocarbons	Lipids	Lignin	Protein	Carbohydrate	Terrestrial
Yellow Snow (Group 1)	12.6	1.1	3.0	12.7	0.6	29.2
Supraglacial Inland (Group 2)	16.0	0.5	2.4	9.2	0.6	23.5
Supraglacial Margin (Group 3)	6.8	4.7	3.6	27.0	0.3	14.3
Subglacial May (Group 3)	1.1	1.5	5.5	25.5	0.1	39.2
Subglacial July-1 (Group 4)	6.9	0.9	10.2	10.4	0.1	55.6
Subglacial July-2 (Group 4)	7.8	1.1	8.5	8.4	0.0	55.8
Tarn (Group 4)	9.8	0.1	7.5	13.3	0.3	59.3
Suwannee River	2.0	0.0	4.5	1.9	0.0	85.6

Figure Captions

Figure 1. Locations of the 2007 and 2008 sample sites. Panel A is a map of Greenland, with the black circle representing the 2007 field site and the red circle representing the 2008 field site. The green contour lines represent the surface elevation (5-km DEM from Bamber et al., 2001; Layberry and Bamber, 2001). Panel B is an expanded image of the two field sites. The 2007 ice surface field site is ~ 40 km inland from the ice sheet edge, and approximately 70 km north of the 2008 field site, located at the glacier margin. Panel C is a Landsat image of the 2008 ice marginal sample location (named 'N' glacier in this study).

Figure 2. Negative ion mode blank-corrected, calibrated mass spectra from the groups identified in indicator species and cluster analysis. Group 1: Yellow Snow (not shown); Group 2: Supraglacial Inland; Group 3: N glacier May (Subglacial May and Supraglacial Margin); and Group 4: terrestrial / N glacier July (Tarn and Subglacial July-1,2). The inset shows the region $375.0 \leq m/z \leq 375.2$ and the indicator m/z values for group 3 (black stars) and group 4 (black ovals).

Figure 3. Cluster diagram of the seven negative ion mode samples, based on Bray-Curtis distance measure and Ward's linkage method.

Figure 4. Van Krevelen diagrams of all formulae assigned (grey dots) to negative ion mode peaks detected within the Supraglacial Inland (A), Subglacial May (B), and Tarn (C) samples. The colored boxes represent elemental compositions for some major compound classes, as approximated from Kim et al. (2003) and Hedges (1990). The grey box represents condensed hydrocarbons, the blue box represents lipids, the green box represents lignin, the yellow box represents proteins, and the pink box represents carbohydrates. The black oval represents elemental formula assignments made for a sample of Suwannee River Fulvic Acid.

Figure 5. Van Krevelen diagram summarizing the formula assignments for the negative ion mode samples. The samples/groups containing a high proportion of peaks in the different compound classes are named.

Figure 6. Van Krevelen negative ion mode diagrams with indicator peaks determined by Indicator Species Analysis. In Panel A, indicator peaks exclusive to either group 3 (N glacier May (Subglacial May and Supraglacial Margin)) or group 4 (terrestrial / N glacier July (Tarn and Subglacial July-1,2)) are shown; in Panel B, peaks from Panel A are shown as well as indicator peaks found in both groups 3 and 4.

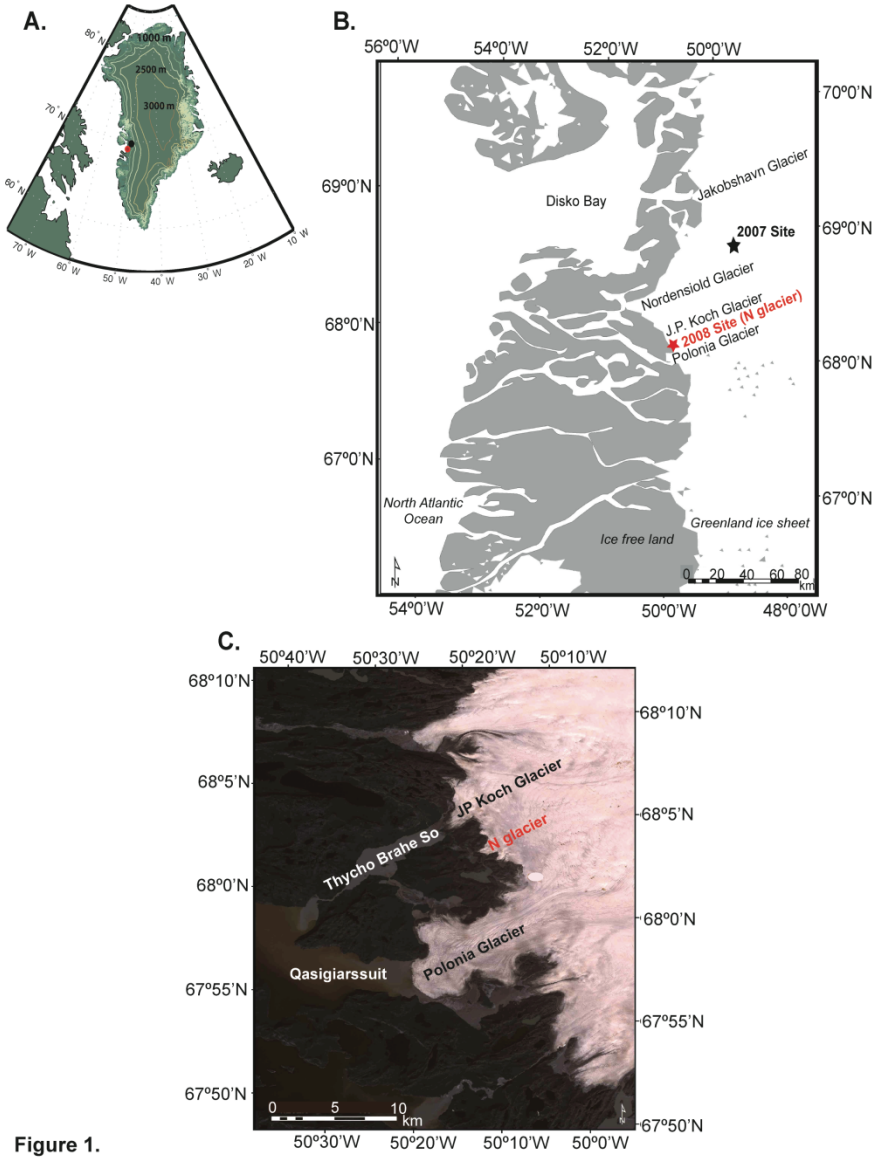

A.
1000 m
2500 m
3000 m
B.
56°0'W
54°0'W
52°0'W
50°0'W
70°0'N
69°0'N
68°0'N
67°0'N
Disko Bay
Jakobshavn Glacier
2007 Site
Nordensiold Glacier
J.P. Koch Glacier
2008 Site (N glacier)
Polonia Glacier
North Atlantic Ocean
Ice free land
Greenland ice sheet
0 20 40 60 80
km
54°0'W
52°0'W
50°0'W
48°0'0W
C.
50°40'W
50°30'W
50°20'W
50°10'W
68°10'N
68°5'N
68°0'N
67°55'N
67°50'N
JP Koch Glacier
N glacier
Thycho Brahe So
Qasigiarssuit
Polonia Glacier
0 5 10
km

Figure 1.

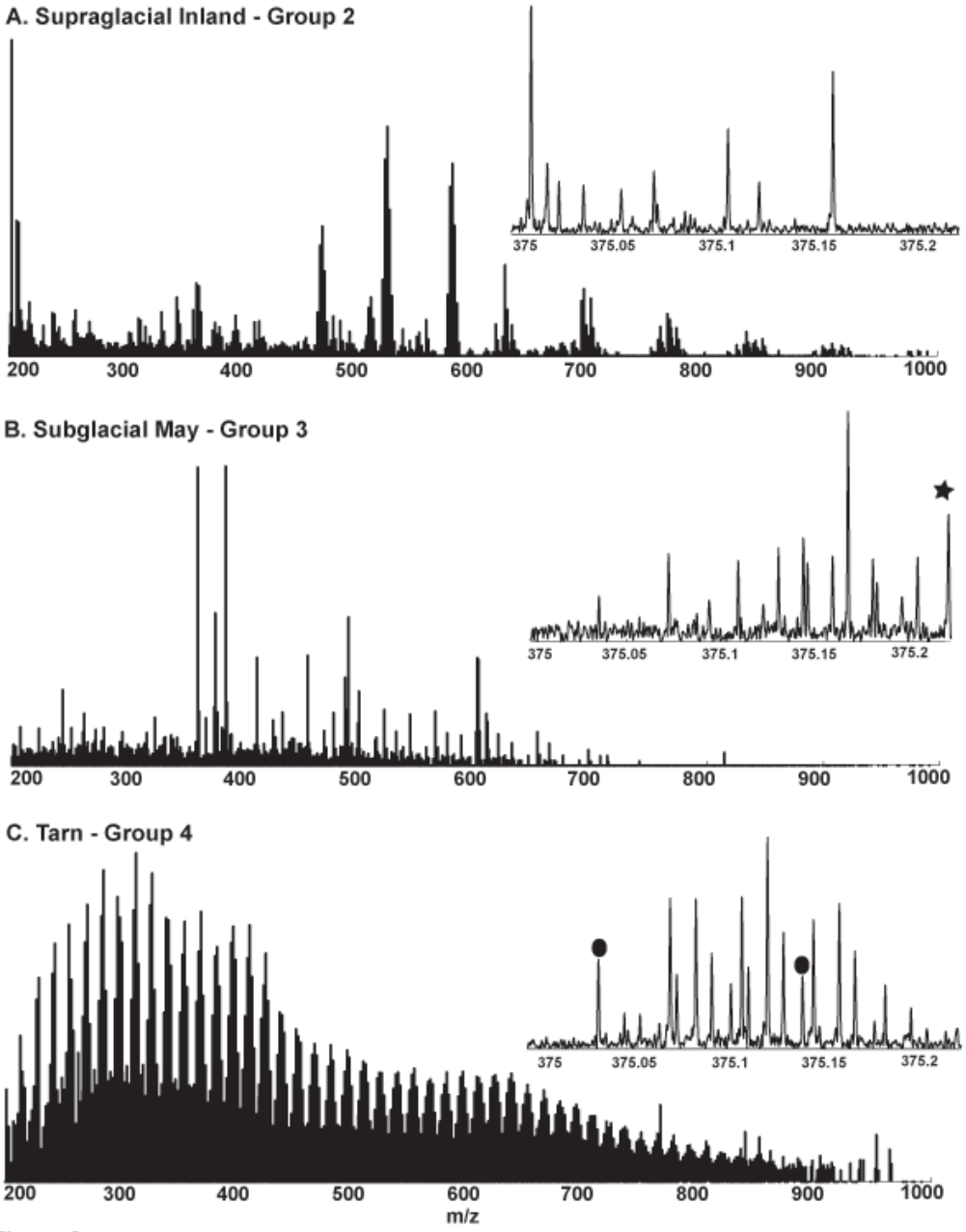


Figure 2.

819
820

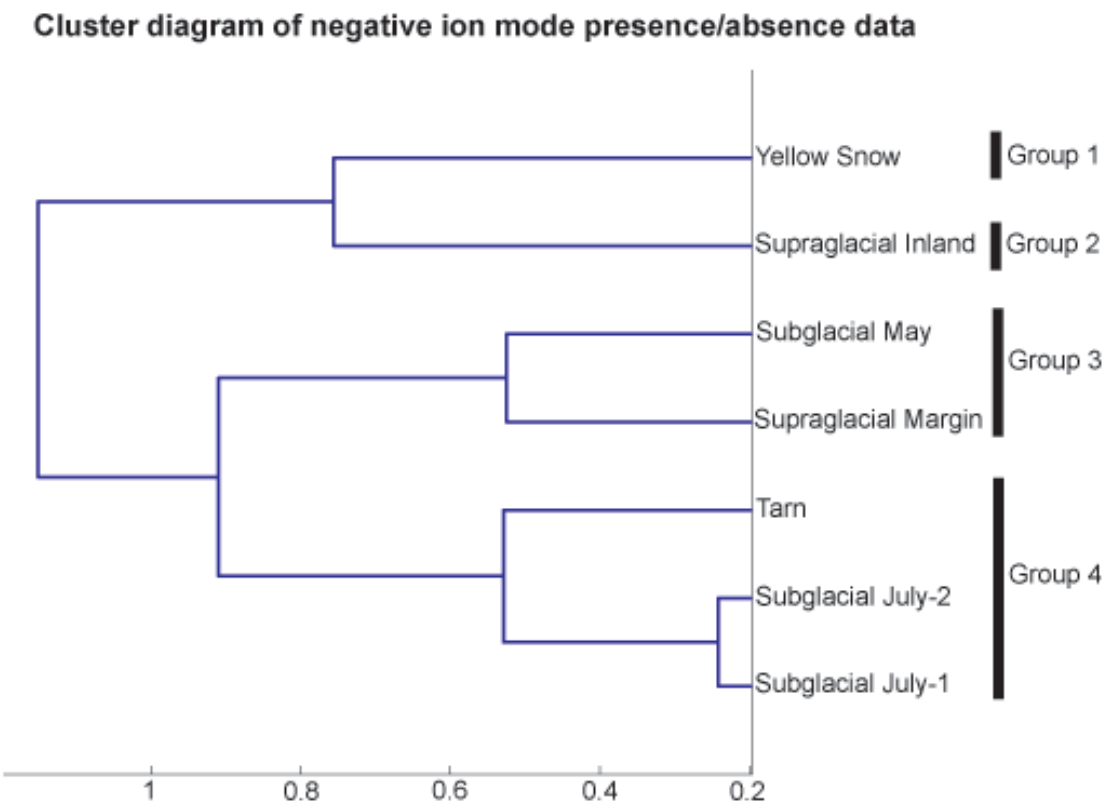
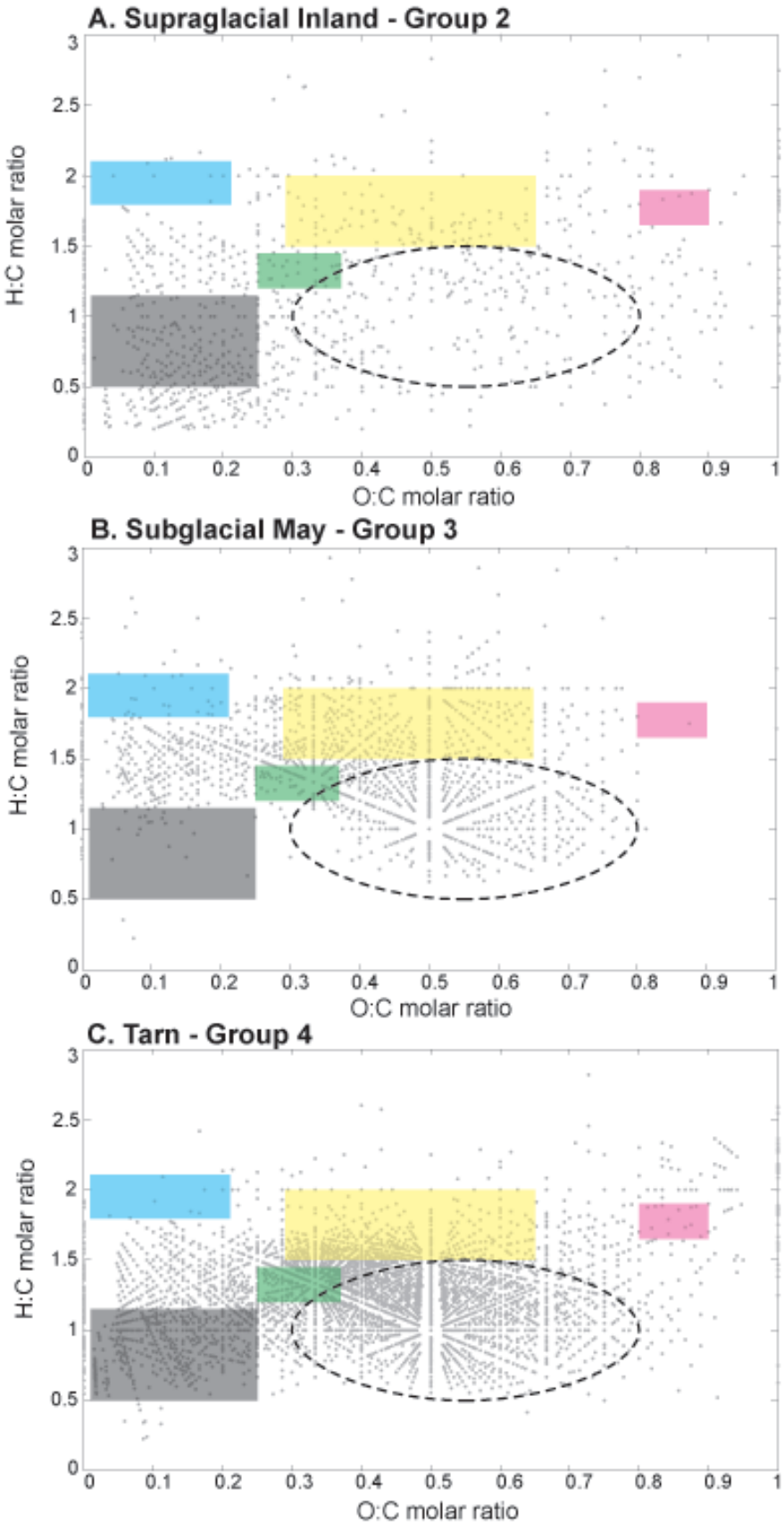


Figure 3.

820
821


A. Supraglacial Inland - Group 2
B. Subglacial May - Group 3
C. Tarn - Group 4
H:C molar ratio
O:C molar ratio
0
0.5
1
1.5
2
2.5
3
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

821 **Figure 4.**

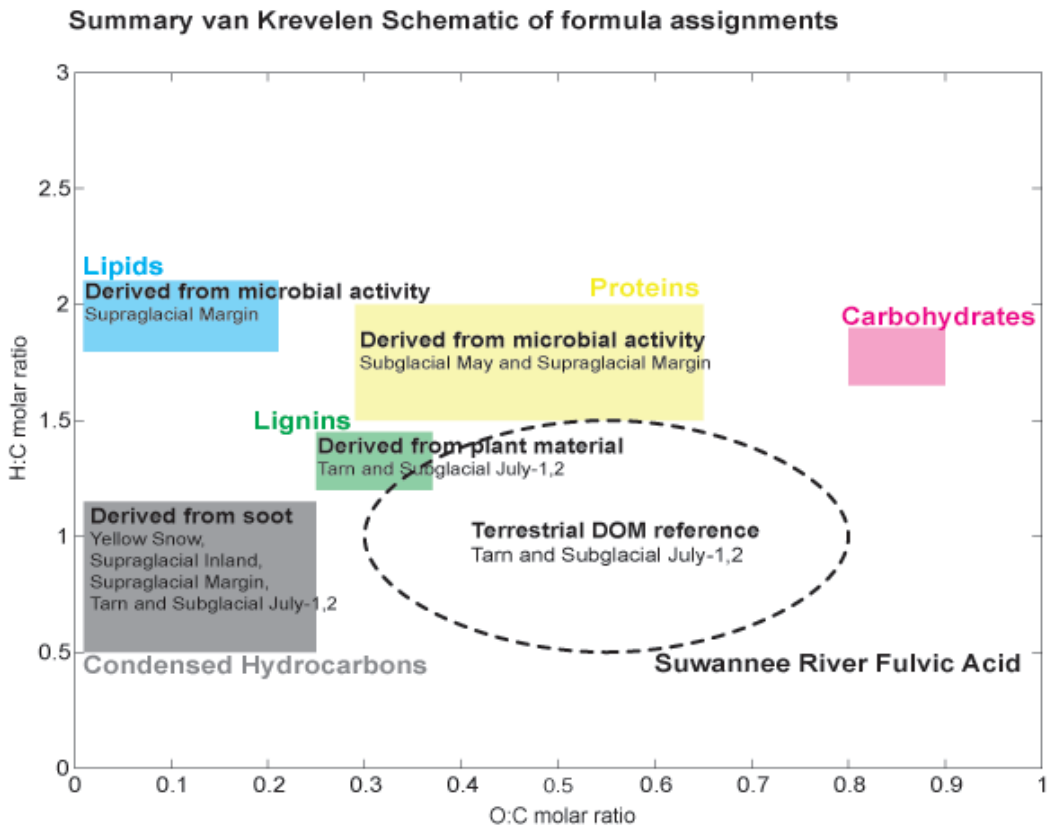
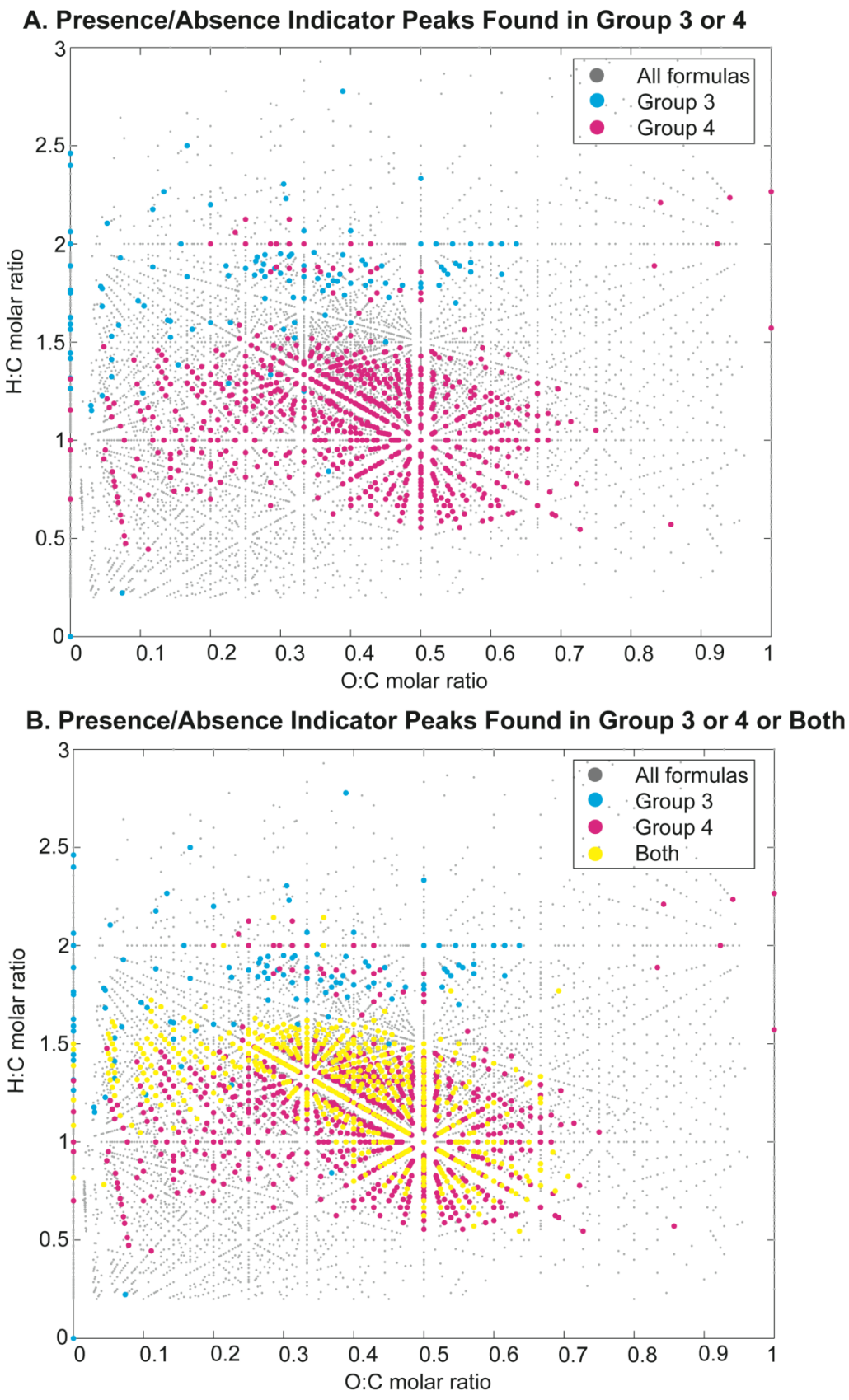


Figure 5.

822
823


A. Presence/Absence Indicator Peaks Found in Group 3 or 4
All formulas
Group 3
Group 4
H:C molar ratio
O:C molar ratio
B. Presence/Absence Indicator Peaks Found in Group 3 or 4 or Both
All formulas
Group 3
Group 4
Both
H:C molar ratio
O:C molar ratio

823 **Figure 6.**

EA Table 1. List of *m/z* values used for internal calibration of (A) positive ion mode data and (B) negative ion mode data. Exact mass refers to the mass calculated from the elemental formula, and charged mass is the exact mass value corrected for positive mode (by adding a Na atom and subtracting an electron) or negative mode (by subtracting a H atom and adding an electron). For the positive ion mode data, we utilized Na adducts. These compounds were chosen because of their frequent occurrence among the different samples analyzed in each mode, and their low error of observed *m/z* values (e.g. the error in mass accuracy ranged from 0.5 to 1.4 for the positive mode calibrants, and 0.3 to 0.8 ppm for the negative ion mode calibrants). In positive mode, calibrants were present in at least six of the seven samples, and in negative mode, calibrants were present in at least half the samples. On occasion, calibrants were added for specific spectra when the original list of calibrants was insufficient to calibrate the desired mass range. In positive mode the internal calibrants span the full range of observed *m/z* values; whereas, in negative mode it was not possible to find calibrants above ~ 600 *m/z* that fit our criteria. However, it is unlikely that the mass error of peaks outside our calibrated range fall outside the 1 ppm error set by the external calibrants because all of the negative mode samples were run within one week.

A. Positive Mode Calibrants (Na Adducts)

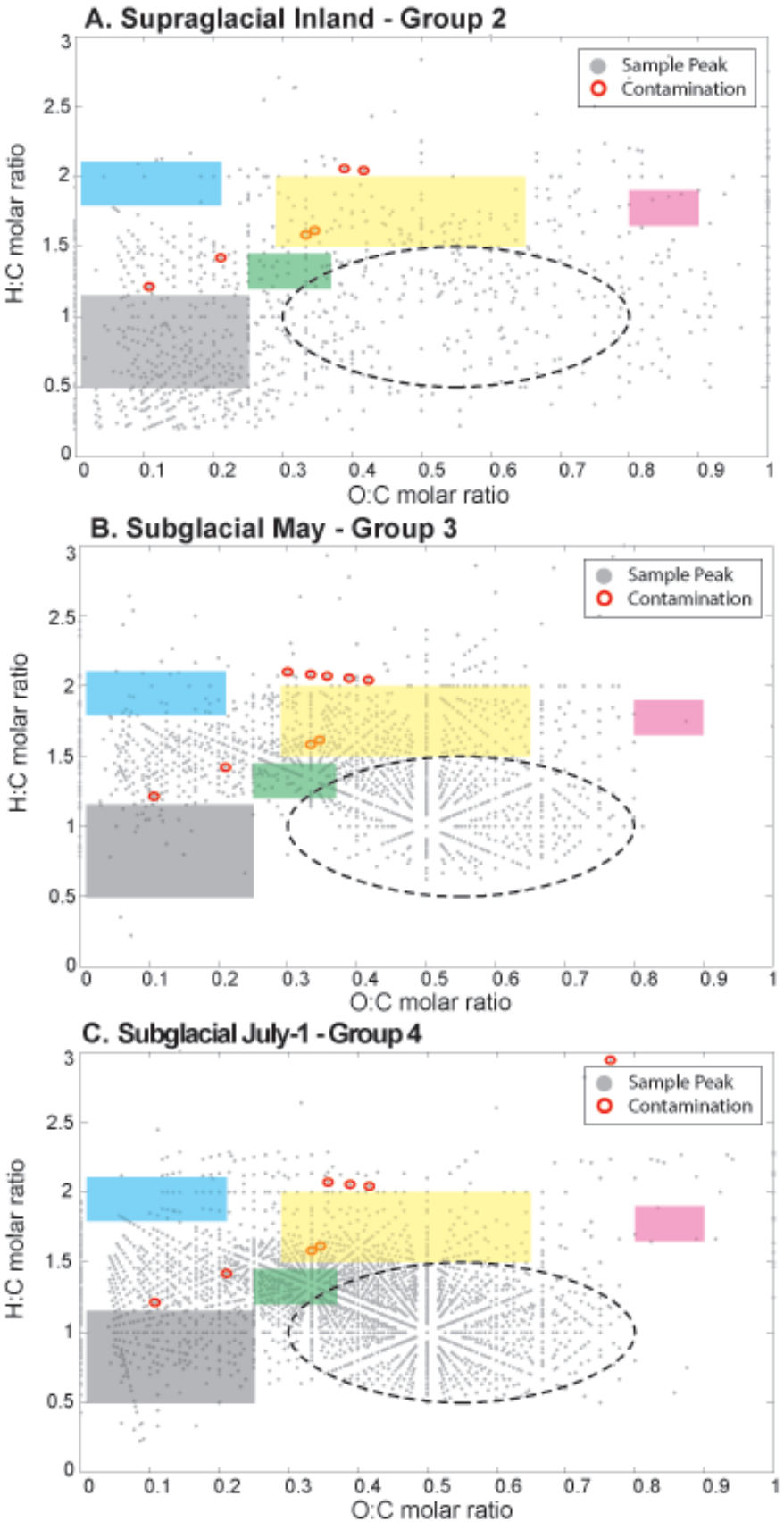
	Elemental Formula	Exact Mass	Charged Mass
1	$C_8H_{18}O_5$	194.115423	217.104642
2	$C_{10}H_{22}O_6$	238.141638	261.130856
3	$C_{12}H_{26}O_7$	282.167853	305.157071
4	$C_{17}H_{36}O_6$	336.251188	359.240407
5	$C_{24}H_{38}O_4$	390.277009	413.266228
6	$C_{18}H_{38}O_{10}$	414.246497	437.235715
7	$C_{20}H_{42}O_{11}$	458.272712	481.261930
8	$C_{22}H_{46}O_{12}$	502.298926	525.288145
9	$C_{24}H_{50}O_{13}$	546.325141	569.314360
10	$C_{26}H_{54}O_{14}$	590.351356	613.340574
11	$C_{28}H_{58}O_{15}$	634.377571	657.366789
12	$C_{30}H_{62}O_{16}$	678.403785	701.393004
13	$C_{37}H_{68}O_{12}$	704.471077	727.460296
14	$C_{35}H_{62}O_{16}$	738.403785	761.393004
15	$C_{42}H_{86}O_{15}$	830.596672	853.585890
16	$C_{45}H_{92}O_{16}$	888.638536	911.627755

B. Negative Mode Calibrants

	Elemental Formula	Exact Mass	Charged Mass
1	$C_{10}H_{16}O_6$	232.094688	231.087411
2	$C_{10}H_{21}O_5N_3$	263.148120	262.140844
3	$C_{13}H_{20}O_6$	272.125988	271.118711
4	$C_{13}H_{10}O_9$	310.032481	309.025205
5	$C_{16}H_{24}O_8$	344.147117	343.139841
6	$C_{24}H_{18}O_3N_2$	382.131742	381.124466
7	$C_{19}H_{24}O_9$	396.142032	395.134755
8	$C_{21}H_{26}O_{10}$	438.152597	437.145320
9	$C_{25}H_{32}O_8$	460.209718	459.202441
10	$C_{21}H_{24}O_{14}$	500.116605	499.109329
11	$C_{27}H_{26}O_{12}$	542.142426	541.135149
12	$C_{28}H_{24}O_{15}$	600.111520	599.104243
13	$C_{26}H_{52}O_{15}$	604.330620	603.323344

Electronic Annex Captions

EA Figure 1. Negative ion mode van Krevelen diagrams illustrating the potential contamination present within the Supraglacial Inland (A), Subglacial May (B), and Subglacial July-1 (C) samples. The contamination was detected in the Yellow Snow mass spectra, likely originating from plasticizers, and consisted of an 18 peak series. Peaks from this potential contamination found in the Supraglacial Inland, Subglacial May, and Subglacial July-1 samples are outlined in red in panels A, B, and C respectively. In Supraglacial Inland the potential contamination represented 6 out of 1865 total sample peaks (0.35%), in Subglacial May the potential contamination represented 9 out of 1737 total sample peaks (0.52%), and in Subglacial July-1 the potential contamination represented 8 out of 3330 total sample peaks (0.24%). The colored boxes represent elemental compositions for some major compound classes, as approximated from Kim et al. (2003) and Hedges (1990). The grey box represents condensed hydrocarbons, the blue box represents lipids, the green box represents lignin, the yellow box represents proteins, and the pink box represents carbohydrates. The black oval represents elemental formula assignments for a sample of Suwannee River Fulvic Acid.



862 **Figure EA1.**