Extracellular peptidases in subsurface sediments of the White Oak River estuary, NC, suggest microbial community adaptation to oxidize degraded organic matter

Andrew D. Steen[1,*], Richard T. Kevorkian[2], Jordan T. Bird[2], Nina Dombrowski[3], Brett J. Baker[3], Shane M. Hagen[1,6], Katherine H. Mulligan[1,4,7], Jenna M. Schmidt[1], Austen T. Webber[1,8], Marc J. Alperin[5]

[1]Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, Tennessee, USA

[2]Department of Microbiology, University of Tennessee, Knoxville, Tennessee, USA

[3]Department of Marine Science, University of Texas-Austin, Marine Science Institute, Port Aransas, Texas, USA

[4]Department of Biology, University of North Carolina at Chapel Hill, Knoxville, Tennessee, USA

[5]Department of Marine Sciences, University of North Carolina at Chapel Hill, USA

[6]Present address: College of Medicine, University of Tennessee Health Science Center, Memphis, Tennessee, USA

[7]Present address: Brody School of Medicine, East Carolina University, Greenville, North Carolina, USA

[8]Present address: Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, USA

[*]Corresponding: asteen1@utk.edu

Abstract

Microbial communities inhabiting subsurface sediments contain abundant heterotrophs, which oxidize organic matter to obtain carbon and energy. Subsurface sediments contain very low concentrations of canonically bioavailable compounds, and it is not clear what fraction of sedimentary organic matter the community metabolizes. To gain a more mechanistic understanding of subsurface heterotrophy, we studied both the genetic potential encoded within metagenomes for extracellular peptidase production, and experimentally assayed the potential activities of a wide range of extracellular peptidases in sediments of the White Oak River estuary, NC. Deeply sequenced metagenomes revealed genes coding for at least 15 classes of extracellular peptidases. We observed enzyme-catalyzed hydrolysis of 11 different peptidase substrates in subsurface sediments. Potential activities (V_{max}) of extracellular peptidases decreased downcore, but cell-specific V_{max} was relatively constant and similar to values observed in seawater phytoplankton blooms. Decreases in half-saturation constants and relative increases in activities of D-phenylalanyl aminopeptidase and ornithyl aminopeptidase with depth indicate a community of heterotrophs that is adapted to access degraded organic matter. These results suggest a subsurface heterotrophic community that converts degraded organic matter into a bioavailable form, rather than a surface-adapted community relying on ever-decreasing concentrations of more labile organic matter.

Introduction

Marine sediments are one of the largest microbial environments on earth (Kallmeyer et al. 2012). Many sedimentary microbes appear to be heterotrophs, slowly metabolizing organic matter (Jørgensen and Marshall, 2016; Biddle *et al.*, 2006), but the mechanisms by which these heterotrophs access old, unreactive organic carbon remain poorly characterized.

In surface environments, where photosynthesis fuels carbon fixation, heterotrophic microorganisms gain energy from a combination of small molecules (<600-1000 Da), which can be taken up directly via general uptake porins (Benz and Bauer, 1988) and macromolecules, which must be broken down outside of the cell by extracellular enzymes. Because most freshly-produced organic matter is macromolecular and large molecules tend to be more bioavailable than small ones (Benner and Amon, 2015), the nature and activity of extracellular enzymes present in surface environments is a major control on the rate of microbial carbon oxidation in such environments.

It is not clear whether microbial extracellular enzymes play the same role in subsurface sediments. Extracellular peptidase activity has been identified in sapropels up to 389 cm below seafloor (cmbsf) in the eastern Mediterranean Sea (Coolen and Overmann, 2000; Coolen *et al.*, 2002), in sediment from 600-630 cmbsf in Aarhus Bay sediments (Lloyd *et al.*, 2013b) and in the interior of seafloor basalts at the Loihi seamount (Jacobson Meyers *et al.*, 2014). However, the nature and relative importance of extracellular enzymes in subsurface environments remains poorly constrained because these studies examined few samples, and only limited enzymatic classes and sample numbers were assayed. It is possible that, like heterotrophs in surface environments, heterotrophs in subsurface environments mainly gain access to organic carbon via extracellular enzymes. On the other hand, some of the unique aspects of subsurface sediments suggest that extracellular enzymes might not be an effective strategy to obtain carbon or energy. First, subsurface sediments contain markedly fewer bioavailable compounds such as amino acids and sugars than do surface sediments (Burdige, 2007). Therefore, these compounds may be insufficiently abundant to be viable heterotrophic substrates. Second, in order for the production of extracellular enzymes to be part of a viable metabolic strategy, each enzyme must, over its lifetime, provide the cell with at

82 least as much carbon or energy as was required to synthesize the enzyme (Vetter *et al.*,
83 1998; Allison, 2005; Schimel and Weintraub, 2003). In subsurface sediments, where
84 metabolic rates may be orders of magnitude slower at the surface, enzyme lifetimes
85 would need to be correspondingly longer to become 'profitable'. Since enzyme lifetimes
86 are finite, there must exist a community metabolic rate below which extracellular enzyme
87 lifetimes are too short to become profitable. Unfortunately that limit is difficult to
88 quantify because enzyme lifetimes in any environment are poorly constrained (e.g., Steen
89 and Arnosti, 2011).

90 It is also possible that extracellular enzymes are not the primary mechanism by
91 which heterotrophs access sedimentary organic matter. More exotic mechanisms, such as
92 abiotic liberation by reactive species formed from radioactive decay of naturally-present
93 radioisotopes (e.g. Blair *et al.*, 2007) are conceivable, but have not been demonstrated.
94 The goals of this work are to determine whether subsurface heterotrophic communities
95 make substantial use of extracellular enzymes to access organic matter, and if so, to
96 characterize the set of enzymes used.

97 We investigated genes for extracellular enzymes and the activities of
98 corresponding enzymes in sediments of the White Oak River, NC by estimating the
99 presence of peptidase families in available metagenomic sequencing data and
100 complementing this by measuring the hydrolysis rates of eleven potential peptidase
101 substrates. This site was chosen because the porewater geochemistry and microbiology of
102 these sediments has been well-characterized (Martens and Goldhaber, 1978; Kelley *et al.*,
103 1990; Baker *et al.*, 2015; Lazar *et al.*, 2016; Lloyd *et al.*, 2011) and because they contain
104 abundant Bathyarchaeota and Marine Benthic Group D archaea, which appear to be
105 capable of metabolizing detrital organic matter (Kubo *et al.*, 2012; Lloyd *et al.*, 2013b;
106 Meng *et al.*, 2014). We focused on peptidases because protein degradation appears to be
107 an important metabolism for some subsurface archaea (Lloyd *et al.*, 2013b) and because
108 peptidases were more active than other enzymes in similar environments (Coolen and
109 Overmann, 2000; Jacobson Meyers *et al.*, 2014). Because environmental samples contain
110 a wide range of distinct peptidases at variable activities (Obayashi and Suzuki, 2005;
111 Steen and Arnosti, 2013) we measured the hydrolysis of eleven different substrates that
112 may be hydrolyzed by structurally and genetically diverse extracellular peptidases. By

measuring potential activities (i.e., the capacity of the enzyme to catalyze hydrolysis if substrate concentrations were not limiting) and substrate affinities of microbial extracellular enzymes, we illuminated some of the mechanisms by which subsurface heterotrophic communities access organic carbon.

Materials and Methods

STUDY SITE

Samples were collected from Station H in the White Oak River Estuary, 34° 44.490' N, 77° 07.44' W, first described by Gruebel and Martens (1984). The White Oak River Estuary occupies a drowned river valley in the coastal plain of North Carolina. Station H is characterized by salinity in the range of 10 to 28 and water depth on the order of 2 m. The flux of ΣCO_2 across the sediment-water interface was 0.46 ± 0.02 mmol m^{-2} hr^{-1} (measured in May of 1987), primarily due to organic carbon oxidation via sulfate reduction, and the sediment accumulation rate averages 0.3 cm yr^{-1} (Kelley *et al.*, 1990). Total organic carbon content is approximately 5%. For this study, push cores of 40-85 cm were collected from Station H by swimmers on May 28, 2013, June 14, 2014, and October 22, 2014. In 2013, cores were transported to the nearby Institute of Marine Sciences (University of North Carolina) at Morehead City, where they were sectioned and processed for enzyme activities, porewater geochemistry, and cell counts within 6 hours of sample collection. Porewater sulfate in 2013 was depleted by 43.5 cm, and methane peaked at 79.5 cm (Fig S1). In 2014, cores were transported on the day of sampling to the University of Tennessee, Knoxville, stored at 4 °C, and processed for enzyme activities the following day. Samples for metagenomic analysis were collected separately in October 2010 from three sites (sites 1, 2, and 3, as previously described by Baker et al (2015)), all of which are within 550 m of Station H.

METAGENOMIC ANALYSIS

To resolve the taxonomic distribution of extracellular peptidases we searched a pre-existing White Oak River *de novo* assembled and binned metagenomic dataset (Baker *et al.*, 2015) for genes that were assigned extracellular peptidase functions. These assignments were based on best matches to extracellular peptidases in KEGG, pfam, and

142 NCBI-nr (non-redundant) databases using the IMG annotation pipeline (Markowitz et al.,
143 2014). Genes were additionally screened for signal peptidase motifs using the following
144 programs: PrediSI setting the organism group to gram-negative bacteria (Hiller *et al.*,
145 2004), PRED-Signal trained on archaea (Bagos *et al.*, 2009), the standalone version of
146 PSORT v.3.0 trained against archaea (Yu *et al.*, 2010), and SignalIP 4.1 using gram-
147 negative bacteria as organism group (Petersen *et al.*, 2011). All programs were used with
148 default settings if not stated otherwise. Results are provided in Supplementary Table 1.
149 In total, binned genomes from three different depth zones of White Oak River
150 sediments were examined. The sulfate-rich zone (SRZ) genomes were obtained from sites
151 2 and 3 core sections 8-12 and 8-10 cm, respectively. The sulfate-methane transitions
152 zone (SMTZ) genomes were recovered from site 2 and 3 and depths of 30-32 cm and 24-
153 28 cm. The methane-rich zone (MRZ) was from site 1 and 52-54 cm. Many of these
154 genes were binned to Bacteria (Baker et al. 2015) and Archaea community members
155 (Baker *et al.*, 2016; Lazar *et al.*, 2016; Seitz *et al.*, 2016). Taxonomic assignments of
156 peptidases identified in the community were based on this binning information. However,
157 since not all of the peptidases were binned, we used top matches to NCBI to identify the
158 unbinned genes. The majority of the Archaea present in the shallow (8-12 cm) sulfate-
159 rich zone were confidently binned, thus were used to determine the relative contributions
160 of archaeal extracellular peptidases (Fig 1b). A smaller proportion of the bacterial
161 peptidase genes were (68% of SRZ, 24% of SMTZ, and 27% of MRZ) confidently
162 binned, therefore, classification was based on top BLAST hits to NCBI. These
163 classifications were then refined using the bin assignments.

164 ENZYME ASSAYS

165 Enzyme assays were performed using different protocols in 2013 versus 2014. In
166 2013, enzyme assays were performed according to a protocol similar to the one described
167 in Lloyd et al (2013). Cores were sectioned into 3 cm intervals. The following intervals
168 were selected for enzyme assays: 0-3 cm, 3-6 cm, 27-30 cm, 57-60 cm, and 81-83 cm.
169 Each section was homogenized, and approximately 0.5 ml wet sediment was transferred
170 into separate 5 ml amber glass serum vials, which had been pre-weighed and preloaded
171 with 4 ml anoxic artificial seawater (Sigma Sea Salts, salinity = 15, pH=7.5) Samples
172 were weighed again to determine the precise mass of wet sediment added, and then an

173 appropriate quantity of 20 mM peptidase substrate stock dissolved in DMSO was added,
174 up to 90 μL, for final substrate concentrations of 0, 25, 50, 75, 100, 200, or 300 μM.
175 Triplicate incubations with 400 μM Arg-AMC, Gly-AMC, Leu-AMC and Gly-Gly-Arg-
176 AMC were also created, but these were omitted for Ala-Ala-Phe-AMC and Boc-Phe-Val-
177 Arg-AMC because the latter two substrates are considerably more expensive than the first
178 four substrates. Each serum vial was vortexed, briefly gassed with N_2 to remove oxygen
179 introduced with the sample, and approximately 1.3 ml slurry was immediately removed,
180 transferred to a microcentrifuge tube, and placed on ice to quench the reaction. The
181 precise time of quenching was recorded. This was centrifuged at 10,000 × *g* within
182 approximately 15 minutes. The supernatant was transferred to a methacrylate cuvette and
183 fluorescence was measured with a Turner Biosystems TBS-380 fluorescence detector set
184 to "UV" (λ_{ex}=365-395 nm, λ_{em}=465-485 nm). Samples were then incubated at 16 °C,
185 approximately the in situ temperature, and the sampling procedure was repeated after
186 approximately 3 hours. The rate of fluorescence production was calculated as the increase
187 in fluorescence for each sample divided by the elapsed time between sample quenching.
188 Killed controls were made using homogenized, autoclaved sediments from 35-45 cmbsf.
189 However, we note that autoclaving does not destroy sediment enzymes because sorption
190 to mineral surfaces stabilizes enzyme structure, vastly increasing their ability to maintain
191 a functional conformation at high temperatures (Stursova and Sinsabaugh, 2008; Carter *et*
192 *al.*, 2007; Schmidt, 2016). We therefore used the autoclaved samples as a qualitative
193 control for the null hypothesis that enzymes were responsible for none of the observed
194 substrate hydrolysis, rather than as a quantitative method to distinguish enzymatic
195 substrate hydrolysis from potential abiotic effects. In some sediments, a large fraction of
196 fluorophore can sorb to particles, requiring a correction to observed fluorescence (Coolen
197 *et al.*, 2002; Coolen and Overmann, 2000), but we observed negligible sorption of
198 fluorophore to the White Oak River sediments.
199 In 2014, enzymes were assayed using a protocol based on the approach of Bell et
200 al (2013), which was designed for soil enzyme assays. In this approach, peptidase
201 substrates were mixed with sediment-buffer slurries in 2-mL wells of a deep-well plate.
202 These plates were periodically centrifuged and 250 μL aliquots of supernatant were
203 transferred into a black 96-well microplate. Fluorescence was read using a BioTek

204 Cytation 3 microplate reader (λ_{ex} = 360 nm, λ_{em} = 440 nm). Results from this method
205 proved considerably noisier than the single-cuvette method used in 2013, so kinetic
206 parameters (V_{max} and K_m) were not calculated for these data. Nevertheless, results were
207 qualitatively similar to those from 2013, and we have reported V_{max} from 2014 as v_0
208 measured at 400 µM substrate concentration, which was saturating. In June 2014, the
209 following substrates were assayed: AAF-AMC, Arg-AMC, Boc-VPR-AMC, D-Phe-
210 AMC, Gly-AMC, Leu-AMC, L-Phe-AMC, Orn-AMC, Z-Phe-Arg-AMC, and Z-Phe-Val-
211 Arg-AMC. In October 2014, L-Phe-AMC, D-Phe-AMC, and Orn-AMC were assayed
212 according to the same protocol in 3-cm core sections at 1.5, 4.5, 7.5, 10.5, 19.5, 22.5,
213 25.5, 28.5, 34.5, 37.5, 40.5, 43.5, 49.5, 52.5, 58.5, and 61.5 cmbsf. Substrate details are
214 given in Table S2.

215 GEOCHEMICAL AND MICROBIOLOGICAL MEASUREMENTS

216 Sediment porosity was measured by mass after drying at 80 °C, according to the equation
217

$$\phi = \frac{m_w/\rho_w}{m_w/\rho_w + \frac{m_d - S \times m_w/1000}{\rho_{ds}}}$$

218
219

220 Here, m_w represents mass lost after drying, ρ_w represents the density of pure water, m_d
221 represents the mass of the dry sediment, S represents salinity in g kg^{-1}, and ρ_{ds} represents
222 the density of dry sediment (assumed to be 2.5 g cm^{-3}). Porewater sulfate concentrations
223 were measured using a Dionex Ion Chromatograph (Sunnyvale, CA) in porewater that
224 was separated by centrifugation in 15 ml centrifuge tubes at 5000 × *g* for 5 minutes,
225 filtered at 0.2 µm, and acidified with 10% HCl. Methane was measured using 3 ml
226 sediment subsamples that were collected from a cutoff syringe, entering through the side
227 of a core section, immediately after core extrusion. Subsamples were deposited
228 immediately in a 20 ml serum vial containing 1 ml, 0.1 M KOH. These were immediately
229 stoppered and shaken to mix sediment with KOH. Methane was later measured by
230 injecting 500 µl of bottle headspace into a GC-FID (Agilent, Santa Clara, CA) using a
231 headspace equilibrium method (Lapham *et al.*, 2008).

Cell Enumeration

Cells were enumerated by direct microscopic counts. One mL of sediment was placed in a 2-mL screw-cap tube with 500 µl of 3% paraformaldehyde in phosphate buffered saline (PBS), in which it was incubated overnight before being centrifuged for 5 minutes at 3000 × *g*. The supernatant was removed and replaced with 500 µl of PBS, vortexed briefly and centrifuged again at 3000 × *g*. The supernatant was subsequently removed and replaced with a 1:1 PBS:ethanol solution. Sediments were then sonicated at 20% power for 40 seconds to disaggregates cells from sediments and diluted 40-fold into PBS prior to filtration onto a 0.2 µm polycarbonate filter (Fisher Scientific, Waltham, MA) and mounted onto a slide. Cells were stained with 4',6-diamidino-2-phenylindole (DAPI) and enumerated by direct counts using a Leica Epifluorescence Microscope.

Geochemical modeling

Organic carbon remineralization rates as a function of depth were estimated by applying a multi-component reaction-transport model to depth distributions of sulfate and methane concentration. The model is based on equations described in Boudreau (1996) and includes only sulfate reduction and methane production due to lack of data regarding oxic and suboxic processes. Thus the model is limited to depths greater than 4.5 cm where sulfate reduction and methane production are the dominant processes, and bioirrigation and bioturbation may be assumed to be negligible. The organic matter remineralization rate is parameterized using the multi-G model first proposed by Jørgensen (1978); a two-component model was sufficient to accurately simulate the sulfate and methane data. For solutes, the upper boundary conditions were measured values at 4.5 cm while the lower boundary conditions (200 cm) were set to zero-gradient. The flux of reactive organic carbon to 4.5 cm was calculated from the sulfate flux across the 4.5 cm horizon and an estimate of methane burial below the lower boundary (the methane flux at the upper boundary was observed to be zero), with an assumed oxidation state of reactive carbon of -0.7. The model contains four adjustable parameters that are set to capture the major details of measured sulfate and methane data: first-order rate

261 constants for both fractions of the reactive carbon pool; the partitioning factor for both
262 fractions, and the rate constant for methane oxidation.

263 Results and Discussion

264 A total of 3739 genes encoding extracellular peptidases were identified among
265 metagenomes from the three depth zones examined, including 685 from SRZ, 1994 from
266 SMTZ, and 1060 from MRZ. Of the genes encoding for peptidases, 0-71% (depending on
267 algorithm and sediment depth) contained a signal peptide and are likely secreted by the
268 SEC-dependent transport system. Among the genes with signal peptidases, zinc
269 carboxypeptidases, peptidases of class C25 and genes of the clostripain family are most
270 likely to be secreted (Table S3). We note that alternative secretion pathways exist,
271 including bacterial lysis and Sec-independent secretion systems (Bendtsen *et al.*, 2005),
272 although the importance of these systems has not to our knowledge been assessed. At all
273 three depths, peptidases of class C25, belonging to the gingipain family, were the most
274 abundant extracellular peptidases (Fig 1a). Gingipains are endopeptidases (i.e., enzymes
275 that cleave proteins mid-chain rather than from the N- or C-termini) with strong
276 specificity for the residue arginine on the N-terminal side of the scissile bond (Rawlings
277 and Barrett, 1999). Bathyarchaeota, which are abundant in sediments of the White Oak
278 River, use extracellular gingipain (among other peptidases) to degrade detrital protein-
279 like organic matter (Kubo *et al.*, 2012; Lloyd *et al.*, 2013b). Genes annotated as encoding
280 extracellular methionine aminopeptidases and zinc carboxypeptidases were also
281 abundant. The composition of protein families was generally consistent with depth, but
282 genes for clostripain (another endopeptidase with preference for Arg N-terminal to the
283 scissile bond) and S24 peptidases (a regulatory peptidase involved in the SOS stress
284 response) were slightly enriched in the SMTZ.

285 In SRZ, Bacteria accounted for 61% of genes for extracellular peptidases for
286 which a lineage could be assigned, increasing to 69% in MRZ (Fig 1b). Archaeal
287 peptidases decreased from 39% in SRZ to 30% in MRZ, while Eukaryota accounted for
288 <1.5% at all depths. Consistent with this distribution, bacteria make up ca. 40-60% of
289 cells in White Oak River sediments as determined by CARD-FISH counts, with no
290 pronounced depth trend (Lloyd *et al.*, 2013a). Since the majority of the archaeal

291 peptidases in the SRZ were assigned to genomic bins, we were able to accurately classify
292 them. Interestingly, over half of the archaeal SRZ peptidases belong to marine benthic
293 group D (MBG-D) genomic bins (Fig 1c)(Lazar et al. in review). Additionally, a large
294 portion of the peptidases in the SRZ belong to newly described Archaea belonging to the
295 Asgard superphylum (Zaremba-Niedzwiedzka *et al.*) including Lokiarchaeota (Spang *et*
296 *al.*, 2015) and Thorarchaeota (Seitz *et al.*, 2016). Little is known about the ecological
297 roles of these novel Archaea. It has recently been shown that they contain metabolic
298 pathways for the degradation of proteins and acetogenesis (Seitz *et al.*, 2016). Sources of
299 bacterial peptidases varied less with depth than sources of archaeal peptidases, with
300 Proteobacteria, Bacteroidetes, and Firmicutes among the dominant phyla (Fig 1d). 13-
301 20% of extracellular peptidase genes belonged to bacteria but could not be assigned a
302 phylum, and 19-22 phyla contributed extracellular peptidases at each depth.
303 Deltaproteobacteria and Gammaproteobacteria dominated the SRZ community at the
304 White Oak River (Baker *et al.*, 2015), and accordingly the largest portion of the
305 extracellular peptidases also belong to them. These phylogenetic groups decrease with
306 depth in the SMTZ and MRZ, however, Deltaproteobacteria contribute a large portion of
307 extracellular peptidases downcore. Although these groups are commonly thought to rely
308 on sulfur and nitrogen respiration, these genomes were shown to contain metabolic
309 pathways for the degradation and fermentation of organic carbon (Baker *et al.*, 2015).
310 Additionally, several phyla thought to be involved in fermentation of detrital carbon
311 including *Clostridia*, *Bacteroidetes*, *Planctomycetes*, and the candidate phyla radiation
312 (CPR) constitute larger portions of the peptidases in the SMTZ and MRZ.
313 Bioinformatic tools are a powerful way to investigate the potential of microbial
314 communities to oxidize complex organic molecules, but these tools do not provide
315 information about the expression level of genes or the in situ activities of gene products
316 and annotation algorithms often fail to identify the precise function of hydrolases,
317 particularly in deeply-branching lineages (e.g., Michalska *et al.*, 2015). Thus, we
318 measured the potential activities of a wide range of extracellular peptidases in the WOR
319 sediments. In 2013 (when the assay protocol used was more sensitive), all six peptidase
320 substrates tested were hydrolyzed faster in untreated sediments than in autoclaved

321 controls (Fig 2, Fig S2). Kinetics of substrate hydrolysis were consistent with the
322 Michaelis-Menten rate law,
323
324 $v_0 = (V_{max} + [S])/(K_m + [S])$,
325
326 which is characteristic of purified hydrolases as well as mixtures of isofunctional
327 enzymes in environmental samples (Steen *et al.*, 2015; Sinsabaugh *et al.*, 2014).
328 Together, these lines of evidence show that the observed substrate hydrolysis was due to
329 extracellular peptidases rather than abiotic factors. Combining data from all three
330 sampling dates, unambiguous hydrolysis of eleven different peptidase substrates was
331 observed.
332 The diverse substrates used in this study were apparently cleaved by a wide
333 variety of peptidases, including aminopeptidases (peptidases that cleave an N-terminal
334 amino acid from a protein) and endopeptidases (peptidases that cleave internal peptide
335 bonds). Peptide bonds adjacent to a diverse set of amino acid residues were cleaved,
336 including glycine (the smallest amino acid), phenylalanine (among the largest amino
337 acids), arginine (positively charged at porewater pH) and leucine (uncharged,
338 hydrophobic). Individual extracellular peptidases can often accept a fairly broad range of
339 substrates, and sometimes a substrate may primarily be hydrolyzed by an enzyme that
340 exhibits maximal activity towards a different substrate. For instance, in pelagic samples
341 from Bogue Sound, an estuary in North Carolina, the substrate Leu-AMC, which is
342 putatively a substrate for leucine aminopeptidase, was hydrolyzed more by arginine
343 aminopeptidase than by leucine aminopeptidase (Steen *et al.*, 2015). Nevertheless, the
344 large diversity of substrates that were hydrolyzed in this study suggests a diverse set of
345 peptidases, as has previously been observed in pelagic samples (Obayashi and Suzuki,
346 2008, 2005; Steen and Arnosti, 2013).
347 Bulk enzyme activities (V_{max} values) decreased with increasing depth (Fig 3a).
348 However, no trend was evident in V_{max} values expressed per cell (Fig 3b), and V_{max}
349 expressed relative to bulk organic carbon oxidation rate increased downcore by nearly
350 two orders of magnitude (Figs 3c-d). V_{max} per cell was approximately 100-200 amol $cell^{-1}$
351 hr^{-1} throughout the core, comparable to previous measurements made in a surface

352 sediments (2-100 amol cell^{-1} hr^{-1}) and surface seawater (mostly less than 100 amol cell^{-1}
353 hr^{-1}, but with some measurements up to 10 nmol cell^{-1} hr^{-1}) (Vetter and Deming, 1994).
354 Modeled organic carbon oxidation due to sulfate reduction and methane production
355 decreased from 16.2 µmol C l wet sediment^{-1} hr^{-1} at 4.5 cm (the top of the model
356 domain) to 0.0312 µmol C l wet sediment^{-1} hr^{-1} at 82.5 cm, a decrease of a factor of 519
357 (or a factor of 1650, relative to the extrapolated organic carbon oxidation rate of 51.4
358 µmol C l wet sediment^{-1} hr^{-1} at 1.5 cm, an estimate which should be considered a lower
359 bound for reasons described in the methods section). The sum of V_{max} of all peptidases
360 measured in 2013 decreased from 94.7 µmol g sed^{-1} hr^{-1} at 1.5 cm to 12.8 µmol g sed^{-1} hr^{-}
361 ^{1} at 82.5 cm depth, a decrease of a factor of 7.4. The ratio of summed peptidase V_{max} to
362 organic carbon oxidation rate correspondingly increased from 1.8 to 410. The absolute
363 value of that ratio is sensitive to the precise set of enzyme included in the sum, but the
364 trend is clear: as sediment depth increased, the potential activity of extracellular
365 peptidases increased faster than the actual rate of organic carbon oxidation. V_{max} is a
366 rough proxy for the concentration of enzymes in an environment, so an increase in V_{max}
367 relative to carbon oxidation rate suggests that subsurface microbial communities
368 produced similar quantities of enzyme per cell as surface communities, but those
369 enzymes returned less bioavailable organic matter, presumably due to lower substrate
370 concentrations.
371 In the sense of bulk activities, therefore, subsurface heterotrophic communities in
372 WOR seem to be similar their surface counterparts in terms of reliance on extracellular
373 enzymes to access organic matter, although metabolisms are slower in the subsurface.
374 However, there are indications from enzyme kinetics and activities of specific enzymes
375 that subsurface communities are specialized for their environment. Peptidase K_m values
376 decreased with increasing depth (Fig 4, ANCOVA, $F(5, 22)=4.44$, $p < 0.05$). When
377 substrate concentrations are considerably less than K_m, as they likely are in subsurface
378 sediments, *in situ* substrate hydrolysis rates are controlled more by K_m than by V_{max}
379 (Steen and Ziervogel, 2012; Cornish-Bowden, 2012), so decreasing K_m values may be an
380 adaptive response to low bioavailable substrate concentrations in subsurface sediments
381 (Sinsabaugh *et al.*, 2014).

382 The trends in K_m warrant some discussion of how K_m values are measured and
383 interpreted. As with all environmental enzyme assays, our measurements relied on
384 measuring the rate of reaction of an artificial substrate that was added to the sample. The
385 sample also contained some quantity of natural substrate (i.e., proteins or peptides). This
386 naturally-present substrate can be viewed as a competitive inhibitor of the added artificial
387 substrate (Cornish-Bowden, 2012). As a consequence, the K_m we measure ($K_{m,app}$) is
388 actually the sum of the true enzyme K_m plus the concentration of *in situ* enzyme substrate
389 ($[S_{is}]$); $K_{m,app} = K_m + [S_{is}]$. It is therefore possible that the observed decrease in $K_{m,app}$
390 actually reflects a change in the concentration of natural peptidase substrates with depth,
391 rather than a change in substrate affinity of peptidases. $[S_{is}]$ is extremely difficult to
392 measure (or even to define precisely) because it reflects the sum of the concentrations of
393 all of the individual molecules which can act as substrate for a given peptidase, modified
394 by the degree to which the peptidase is capable of accessing each substrate. $[S_{is}]$ almost
395 certainly decreased downcore, since a decrease in the concentration of enzymatically-
396 hydrolyzable, protein-like organic matter is a diagnostic feature of aged organic matter
397 (Amon *et al.*, 2001; Dauwe *et al.*, 1999). However, it is very unlikely that $[S_{is}]$ could have
398 decreased enough to cause more than a negligible fraction of the observed decreases in
399 K_m, which were ranged from 110 μM (GGR-AMC) to 990 μM (AAF-AMC). Sediments
400 of Aarhus Bay, which are similar to those of the White Oak River in terms of organic
401 matter content and grain size, are characterized by porewater dissolved combined amino
402 acid concentrations in the range of 50-150 μM L^{-1} (Pedersen *et al.*, 2001). However, only
403 a fraction of total dissolved protein-like material can act as substrate for any given
404 peptidase. Only 10-40% of total combined amino acids in sediments are accessible to
405 added peptidases (Dauwe *et al.*, 1999), and a smaller fraction may be available to the
406 specific peptidases measured in this study. Furthermore, K_m values were measured in a
407 9:1 buffer:sample slurry, so $[S_{is}]$ in the slurry would be one-tenth of *in situ* $[S_{is}]$. By
408 multiplying assumed porewater substrate concentrations (from Aarhus Bay) by the
409 fraction that is enzymatically hydrolysable and dividing by 10 to account for dilution in
410 the slurry, we estimate that changes in [Sis] could account for, at most, a 0.5-6 μM
411 decrease in $K_{m,app}$. The remainder of the 110-990 μM decrease must be due to changes in

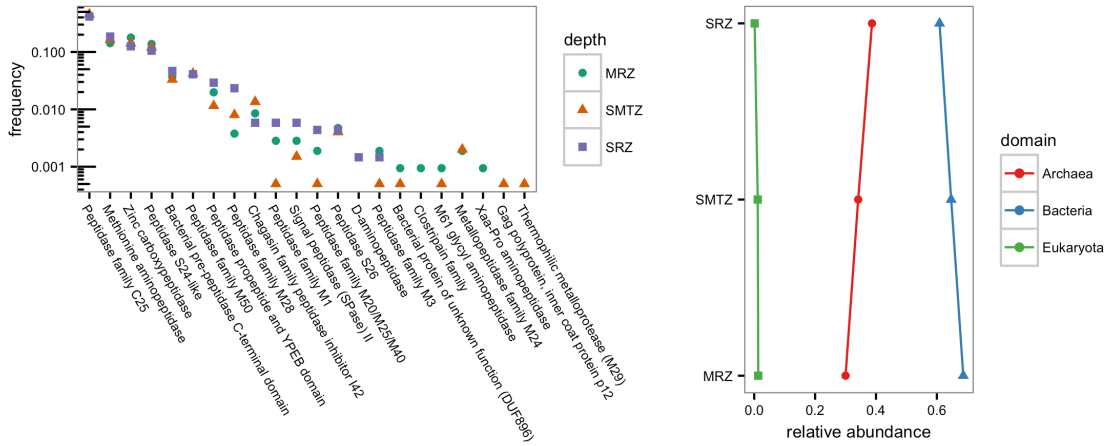
412 the true substrate affinity K_m, indicating that subsurface enzymes hydrolyze low
413 concentrations of protein considerably more efficiently than surface enzymes.
414 The suite of specific peptidases that are active in deeper sediments also seems to
415 reflect adaptation to more degraded organic matter. In order to assess the degree to which
416 sedimentary extracellular peptidases target more recalcitrant organic matter, rather than
417 the extremely small pool of relatively labile organic matter, in 2014 we compared
418 potential activities of D-phenylalanyl aminopeptidase (D-PheAP) and ornithine
419 aminopeptidase (OrnAP) to those of L-phenylalanyl aminopeptidase (L-PheAP). D-
420 phenylalanine and ornithine are both markers for degraded organic matter. Most amino
421 acids are biosynthesized as L-stereoisomers. D-stereoisomers in sedimentary organic
422 matter can be produced via abiotic racemization of biomass (Bada and Schroeder, 1975;
423 Steen *et al.*, 2013) or by bacterial reprocessing of phytoplankton-derived OM (Pedersen
424 *et al.*, 2001; Kaiser and Benner, 2008; Lomstein *et al.*, 2006). Ornithine is non-
425 proteinogenic amino acid, which does not exist in high concentration in fresh biomass,
426 but which can be produced in sediments via deamination of arginine and therefore
427 indicates OM degradation (Hare, 1968; Lee and Cronin, 1984). L-phenylalanine is among
428 the most recalcitrant amino acids (Dauwe and Middelburg, 1998) but is more labile than
429 D-phenylalanine or ornithine, so we take L-PheAP activity as a marker of the
430 community's ability to access relatively fresh OM.
431 D-PheAP:L-PheAP and OrnAP:L-PheAP ratios increased significantly with depth
432 (Fig 5, D-PheAP:L-Phe-AP: $p<0.01$, $r^2=0.35$, n=16; OrnAP:L-PheAP: $p < 0.05$, $r^2=0.26$,
433 n=16), indicating that subsurface communities expressed peptidases that release amino
434 acids from relatively recalcitrant organic matter. These results do not indicate the actual
435 flux of organic matter to communities: V_{max} reflects a potential rate at saturating substrate
436 concentrations, not an *in situ* rate. Nevertheless, these results suggest that deeper
437 heterotrophic communities seek to access the larger pool of more recalcitrant organic
438 matter rather than the smaller pool of more labile organic matter.
439 The absence of a major trend in cell-specific extracellular peptidase activities with
440 depth in the White Oak River indicates that subsurface heterotrophic microbial
441 communities rely on extracellular enzymes to access organic carbon to a similar degree
442 as communities in rapidly-changing environments such as surface seawater. However, the

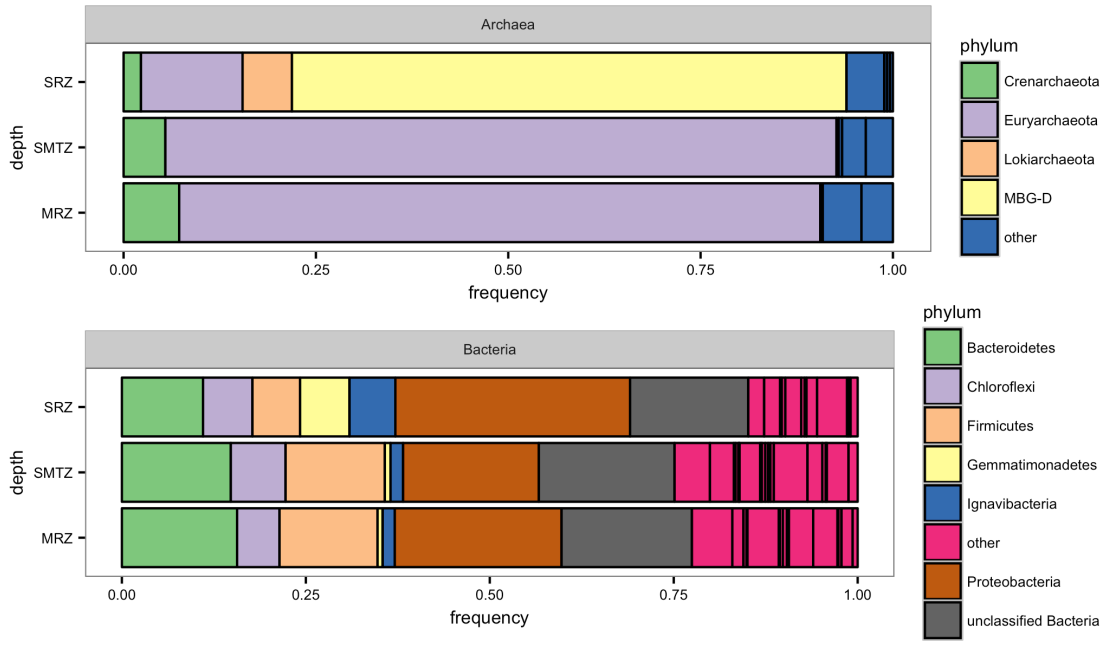
443 nature of the enzymes, as reflected by their substrate affinities, the specific distribution of
444 enzyme activities, and the distribution of peptidase genes, changes with depth. These
445 changes are consistent with community-level adaptation to consuming degraded,
446 relatively recalcitrant organic matter. Further analysis of the mechanisms by which
447 subsurface heterotrophs access organic matter may yield continued insights into how
448 heterotrophic microorganisms live in low-energy environments such as subsurface
449 sediments.
450

Figures

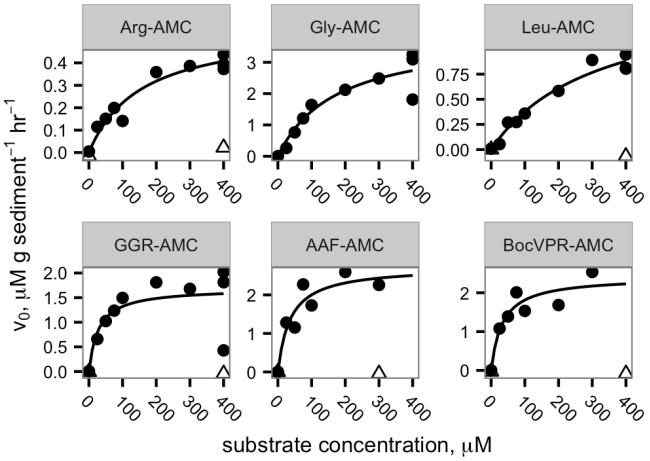
Fig 1:

(a) Frequency of genes for various classes of extracellular peptidases, relative to all genes for extracellular peptidases; (b) sources of extracellular peptidases by domain and depth, and sources of (c) Archaeal and (d) Bacterial peptidases by phylum and depth.





461 Fig 2: Peptidase saturation curves collected in 2013, showing Michealis-Menten kinetics
462 consistent with enzymatic rather than abiotic substrate hydrolysis. Open triangles indicate
463 autoclaved controls.



464
465
466

Fig 3: Peptidase activities measured in 2013 compared to microbial abundance and activity. Panel A: summed V_{max} of the six peptidases measured in 2013, expressed relative to the value at 1.5 cm. Error bars represent standard error of the estimate of the rate of fluorophore production. B: Summed V_{max} relative to cell abundance. Error bars represent error propagated from error of Vmax and standard deviation of cell counts. C: Modeled organic carbon oxidation rate. D: Summed V_{max} relative to organic carbon oxidation rate. Error bars represent standard error of the estimate of the rate of fluorophore production.

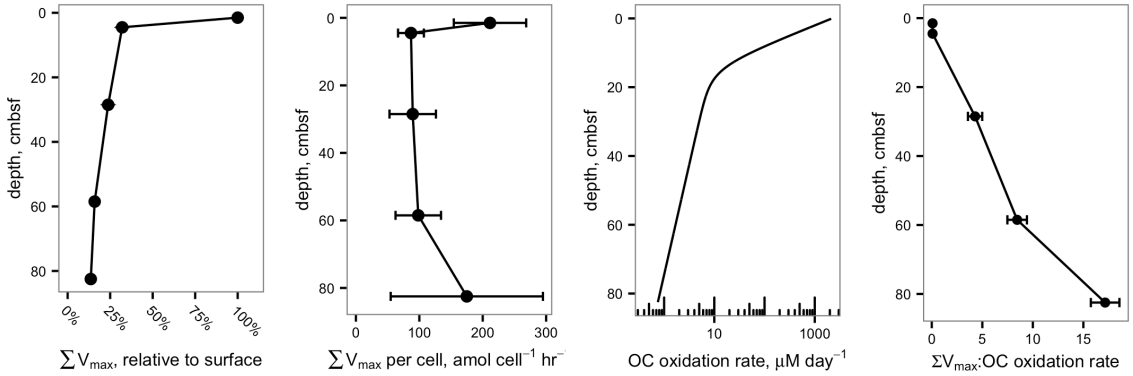


Fig 4: Estimated K_m values as a function of sediment depth. Colored lines indicate linear regressions for individual substrates, while the black line and shaded area represent a regression for all substrates taken together.

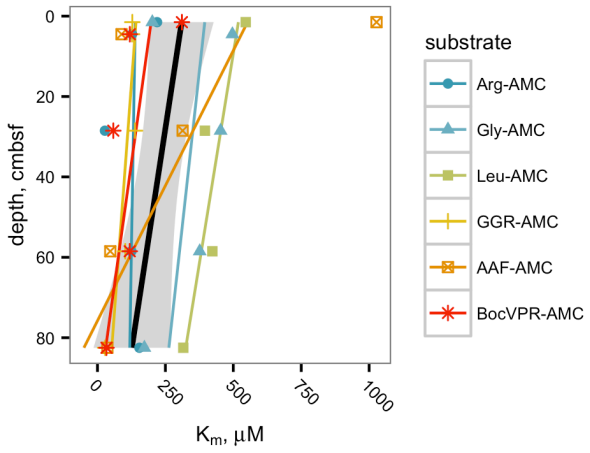
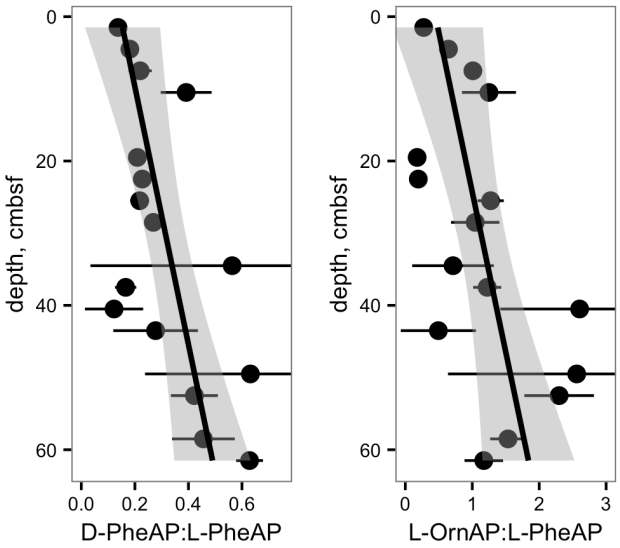


Fig 5: Left panel: Ratio of D-phenylalanyl aminopeptidase activity to L-phenylalanyl aminopeptidase activity versus depth. Right panel: Ratio of ornithyl aminopeptidase activity to L-phenylalanyl aminopeptidase activity. The shaded band indicates the 95% confidence interval of the fitted values.



Conflict of interest

The authors declare no competing financial interests in relation to the work described.

Acknowledgements

We thank Michael Piehler for access to his laboratory facilities at the University of North Carolina Institute of Marine Sciences, the captain of the R/V Capricorn for sampling assistance, and Terry Hazen for use of lab equipment at the University of Tennessee. We thank Oliver Jeffers for reminding ADS that science is cool. Funding for KHM was provided by NSF grant DBI-1156644 to Steven W. Wilhelm. Funding for ADS was provided by NSF grant OCE-1431598 and a C-DEBI subaward. This work is C-DEBI Contribution number <<to be determined.>>

References

Allison SD. (2005). Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments. *Ecol Lett* **8**: 626–635.

Amon RMW, Fitznar H-P, Benner R. (2001). Linkages among the bioreactivity, chemical composition, and diagenetic state of marine dissolved organic matter. *Limnol Oceanogr* **46**: 287–297.

Bada JL, Schroeder R a. (1975). Amino acid racemization reactions and their geochemical implications. *Naturwissenschaften* **62**: 71–79.

Bagos PG, Tsirigos KD, Plessas SK, Liakopoulos TD, Hamodrakas SJ. (2009). Prediction of signal peptides in archaea. *Protein Eng Des Sel* **22**: 27–35.

Baker BJ, Lazar CS, Teske AP, Dick GJ. (2015). Genomic resolution of linkages in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria. *Microbiome* **3**: 14.

Baker BJ, Saw JH, Lind AE, Lazar CS, Hinrichs K-U, Teske AP, *et al.* (2016). Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. *Nat Microbiol* **1**: 16002.

Bell CW, Fricks BE, Rocca JD, Steinweg JM, McMahon SK, Wallenstein MD. (2013). High-throughput fluorometric measurement of potential soil extracellular enzyme activities. *J Vis Exp* e50961.

Bendtsen JD, Kiemer L, Fausbøll A, Brunak S. (2005). Non-classical protein secretion in bacteria. *BMC Microbiol* **5**: 58.

Benner R, Amon RMW. (2015). The size-reactivity continuum of major bioelements in the ocean. *Ann Rev Mar Sci* **7**: 185–205.

Benz R, Bauer K. (1988). Permeation of hydrophilic molecules through the outer membrane of gram-negative bacteria. Review on bacterial porins. *Eur J Biochem* **176**: 1–19.

Biddle JF, Lipp JS, Lever MA, Lloyd KG, Sørensen KB, Anderson R, *et al.* (2006). Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc Natl Acad Sci U S A* **103**: 3846–51.

Blair CC, D'Hondt S, Spivack AJ, Kingsley RH. (2007). Radiolytic hydrogen and microbial respiration in subsurface sediments. *Astrobiology* **7**: 951–70.

Boudreau BP. (1996). A method-of-lines code for carbon and nutrient diagenesis in aquatic sediments. *Comput Geosci* **22**: 479–496.

Burdige DJ. (2007). Preservation of organic matter in marine sediments: controls, mechanisms, and an imbalance in sediment organic carbon budgets? *Chem Rev* **107**: 467–85.

Carter DO, Yellowlees D, Tibbett M. (2007). Autoclaving kills soil microbes yet soil enzymes remain active. *Pedobiologia (Jena)* **51**: 295–299.

Coolen MJL, Cypionka H, Sass AM, Sass H, Overmann J. (2002). Ongoing modification of Mediterranean Pleistocene sapropels mediated by prokaryotes. *Science* **296**: 2407–10.

Coolen MJL, Overmann J. (2000). Functional Exoenzymes as Indicators of Metabolically Active Bacteria in 124,000-Year-Old Sapropel Layers of the Eastern Mediterranean Sea. *Appl Environ Microbiol* **66**: 2589–2598.

Cornish-Bowden A. (2012). Fundamentals of enzyme kinetics. Fourth Edi. Wiley-VCH Verlag & Co. KGaA: Weinheim.

Dauwe B, Middelburg JJ. (1998). Amino acids and hexosamines as indicators of organic matter degradation state in North Sea sediments. *Limnol Oceanogr* **43**: 782–798.

Dauwe B, Middelburg JJ, Van Rijswijk P, Sinke J, Herman PMJ, Heip CHR. (1999). Enzymatically hydrolyzable amino acids in North Sea sediments and their possible implication for sediment nutritional values. *J Mar Res* **57**: 109–134.

Gruebel KA, Martens CS. (1984). Radon-222 tracing of sediment-water chemical transport in an estuarine sediment. *Limnol Oceanogr* **29**: 587–597.

Hare PE. (1968). Geochemistry of Proteins, Peptides, and Amino Acids. In: Eglinton G, Murphy MTJ (eds). *Organic Geochemistry: Methods and Results*. Springer-Verlag: Berlin, Heidelberg, pp 438–462.

Hiller K, Grote A, Scheer M, Münch R, Jahn D. (2004). PrediSi: Prediction of signal peptides and their cleavage positions. *Nucleic Acids Res* **32**. e-pub ahead of print, doi: 10.1093/nar/gkh378.

Jacobson Meyers ME, Sylvan JB, Edwards KJ. (2014). Extracellular enzyme activity and microbial diversity measured on seafloor exposed basalts from Loihi seamount indicate the importance of basalts to global biogeochemical cycling. *Appl Environ Microbiol* **80**: 4854–64.

Jørgensen BB. (1978). A comparison of methods for the quantification of bacterial sulfate reduction in coastal marine sediments: II. Calculation from mathematical models. *Geomicrobiol J* **1**: 29–47.

Jørgensen BB, Marshall IPG. (2016). Slow Microbial Life in the Seabed. *Ann Rev Mar Sci* **8**: 311–32.

Kaiser K, Benner R. (2008). Major bacterial contribution to the ocean reservoir of detrital organic carbon and nitrogen. *Limnol Oceanogr* **53**: 99–112.

Kelley CA, Martens CS, Chanton JP. (1990). Variations in sedimentary carbon remineralization rates in the White Oak River estuary, North Carolina. *Limnol Oceanogr* **35**: 372–383.

Kubo K, Lloyd KG, F Biddle J, Amann R, Teske A, Knittel K. (2012). Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *ISME J* **6**: 1949–65.

Lapham LL, Chanton JP, Martens CS, Sleeper K, Woolsey JR. (2008). Microbial activity in surficial sediments overlying acoustic wipeout zones at a Gulf of Mexico cold seep. *Geochemistry, Geophys Geosystems* **9**: n/a-n/a.

Lazar CS, Baker BJ, Seitz K, Hyde AS, Dick GJ, Hinrichs K-U, *et al.* (2016). Genomic evidence for distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments. *Environ Microbiol* **18**: 1200–11.

Lee C, Cronin C. (1984). Particulate amino acids in the sea: Effects of primary productivity and biological decomposition. *J Mar Res* **42**: 1075–1097.

Lloyd KG, Alperin MJ, Teske A. (2011). Environmental evidence for net methane production and oxidation in putative ANaerobic MEthanotrophic (ANME) archaea. *Environ Microbiol* **13**: 2548–2564.

Lloyd KG, May MK, Kevorkian RT, Steen AD. (2013a). Meta-analysis of quantification methods shows that archaea and bacteria have similar abundances in the subseafloor. *Appl Environ Microbiol* **79**: 7790–9.

Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever M a, Steen AD, *et al.* (2013b). Predominant archaea in marine sediments degrade detrital proteins. *Nature* **496**: 215–8.

Lomstein B, Jorgensen B, Schubert C, Niggemann J. (2006). Amino acid biogeo- and stereochemistry in coastal Chilean sediments. *Geochim Cosmochim Acta* **70**: 2970–2989.

Martens CS, Goldhaber MB. (1978). Early diagenesis in transitional sedimentary environments of the White Oak River Estuary, North Carolina. *Limnol Oceanogr* **23**: 428–441.

Meng J, Xu J, Qin D, He Y, Xiao X, Wang F. (2014). Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. *ISME J* **8**: 650–9.

Michalska K, Steen AD, Chhor G, Endres M, Webber AT, Bird J, *et al.* (2015). New aminopeptidase from 'microbial dark matter' archaeon. *FASEB J* **29**: 4071–4079.

Obayashi Y, Suzuki S. (2008). Occurrence of exo- and endopeptidases in dissolved and particulate fractions of coastal seawater. *Aquat Microb Ecol* **50**: 231–237.

Obayashi Y, Suzuki S. (2005). Proteolytic enzymes in coastal surface seawater: Significant activity of endopeptidases and exopeptidases. *Limnol Oceanogr* **50**: 722–726.

Pedersen A-GU, Thomsen TR, Lomstein BA, Jørgensen NOG. (2001). Bacterial influence on amino acid enantiomerization in a coastal marine sediment. *Limnol Oceanogr* **46**: 1358–1369.

Petersen TN, Brunak S, von Heijne G, Nielsen H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**: 785–6.

Rawlings ND, Barrett AJ. (1999). MEROPS: The peptidase database. *Nucleic Acids Res* **27**: 325–331.

Schimel JP, Weintraub MN. (2003). The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model. *Soil Biol Biochem* **35**: 549–563.

Schmidt J. (2016). Microbial Extracellular Enzymes in Marine Sediments: Methods Development and Potential Activities in the Baltic Sea Deep Biosphere. University of Tennessee - Knoxville.

Seitz KW, Lazar CS, Hinrichs K-U, Teske AP, Baker BJ. (2016). Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. *ISME J* **10**: 1696–1705.

Sinsabaugh RL, Belnap J, Findlay SG, Shah JJF, Hill BH, Kuehn KA, *et al.* (2014). Extracellular enzyme kinetics scale with resource availability. *Biogeochemistry* **121**: 287–304.

633 Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, *et al.*
634 (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*
635 **521**: 173–179.
636 Steen AD, Arnosti C. (2013). Extracellular peptidase and carbohydrate hydrolase
637 activities in an Arctic fjord (Smeerenburgfjord, Svalbard. *Aquat Microb Ecol* **69**: 93–99.
638 Steen AD, Arnosti C. (2011). Long lifetimes of β-glucosidase, leucine aminopeptidase,
639 and phosphatase in Arctic seawater. *Mar Chem* **123**: 127–132.
640 Steen AD, Jørgensen BB, Lomstein BA. (2013). Abiotic Racemization Kinetics of
641 Amino Acids in Marine Sediments. *PLoS One* **8**. e-pub ahead of print, doi:
642 10.1371/journal.pone.0071648.
643 Steen AD, Vazin JP, Hagen SM, Mulligan KH, Wilhelm SW. (2015). Substrate
644 specificity of aquatic extracellular peptidases assessed by competitive inhibition assays
645 using synthetic substrates. *Aquat Microb Ecol* **75**: 271–281.
646 Steen AD, Ziervogel K. (2012). Comment on the review by German et
647 al.(2011)'Optimization of hydrolytic and oxidative enzyme methods for ecosystem
648 studies'[Soil Biology & Biochemistry 43: 1387--1397]. *Soil Biol Biochem* **48**: 196–197.
649 Stursova M, Sinsabaugh RL. (2008). Stabilization of oxidative enzymes in desert soil
650 may limit organic matter accumulation. *Soil Biol Biochem* **40**: 550–553.
651 Vetter YA, Deming JW. (1994). Extracellular enzyme activity in the Arctic Northeast
652 Water polynya. *Mar Ecol Prog Ser* **114**: 23–34.
653 Vetter Y, Deming J, Jumars P, Krieger-Brockett B. (1998). A Predictive Model of
654 Bacterial Foraging by Means of Freely Released Extracellular Enzymes. *Microb Ecol* **36**:
655 75–92.
656 Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, *et al.* (2010). PSORTb 3.0:
657 improved protein subcellular localization prediction with refined localization
658 subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**: 1608–
659 1615.
660 Zaremba-Niedzwiedzka K, Caceres E, Saw J, Backstrom D, Juzokaite L, Vancaester E, *et*
661 *al.* Metagenomic exploration of Asgard arcahea illuminates the origin of eukaryotic
662 cellular complexity. *Nature*.
663

664 ## Supplemental Material

665 Table S1: substrates and enzymes. All amino acids are in the L stereoconformation unless
666 otherwise noted. AMC stands for 7-amido-4-methylcoumarin. N-(carboxylbenzyloxy-)
667 (Z-) and tert-butyl (Boc) are blocking groups which prevent hydrolysis by
668 aminopeptidases due to steric hindrance. Enzymes assayed are listed as "putative"
669 because the enzyme listed shows maximal activity towards the listed substrate, but it is
670 possible (and in some cases very likely) that other enzymes also hydrolyze the listed
671 substrate. None of the peptidases in the E.C. database preferentially hydrolyze N-terminal
672 glycine, ornithine, or phenylalanine, so no E.C. numbers are listed for Gly-AMC, Orn-
673 AMC or L-Phe-AMC, although many peptidases likely exhibit secondary activity
674 towards those residues.

675

Substrate	Abbreviation	Putative enzyme assayed	E.C. number
Ala-Ala-Phe-AMC	AAF-AMC	subtilisin	3.4.21.62
Arg-AMC	Arg-AMC	Aminopeptidase B	3.4.11.6
Gly-AMC	Gly-AMC	Peptidases that preferentially hydrolyze N-terminal Gly	n/a
Gly-Gly-Arg-AMC	Gly-Gly-Arg-AMC	trypsin	3.4.21.4
Leu-AMC	Leu-AMC	leucyl aminopeptidase	3.4.11.1, 3.4.11.10
Ornithine-AMC	Orn-AMC	Peptidses that preferentially hydrolyze N-terminal Orn	n/a
D-Phe-AMC	D-Phe-AMC	D-stereospecific aminopeptidase	3.4.11.19
L-Phe	L-Phe-AMC	Peptidases that preferentially hydrolyze N-terminal Phe	n/a
Z-Phe-Arg-AMC	Z-FR-AMC	gingipain R	3.4.22.37
Z-Phe-Val-Arg-AMC	Z-FVR-AMC	clostripain	3.4.22.8
boc-Val-Pro-Arg-AMC	Boc-VPR-AMC	trypsin	3.4.21.4

676

677 Table S2: Sources for metagenomic data

Site	IMG Genome ID	Genome size	Gene count
White Oak River	3300001751	8247658995	43518175
White Oak River	3300001753	10609117402	52065549
White Oak River	3300001706	7116721444	34096044
White Oak River	3300002053	799007063	857807

678

679 Table S3: called signal peptidases by gene, depth, and prediction algorithm.

Fig. S1: Cell abundance and porewater methane and sulfate concentrations.

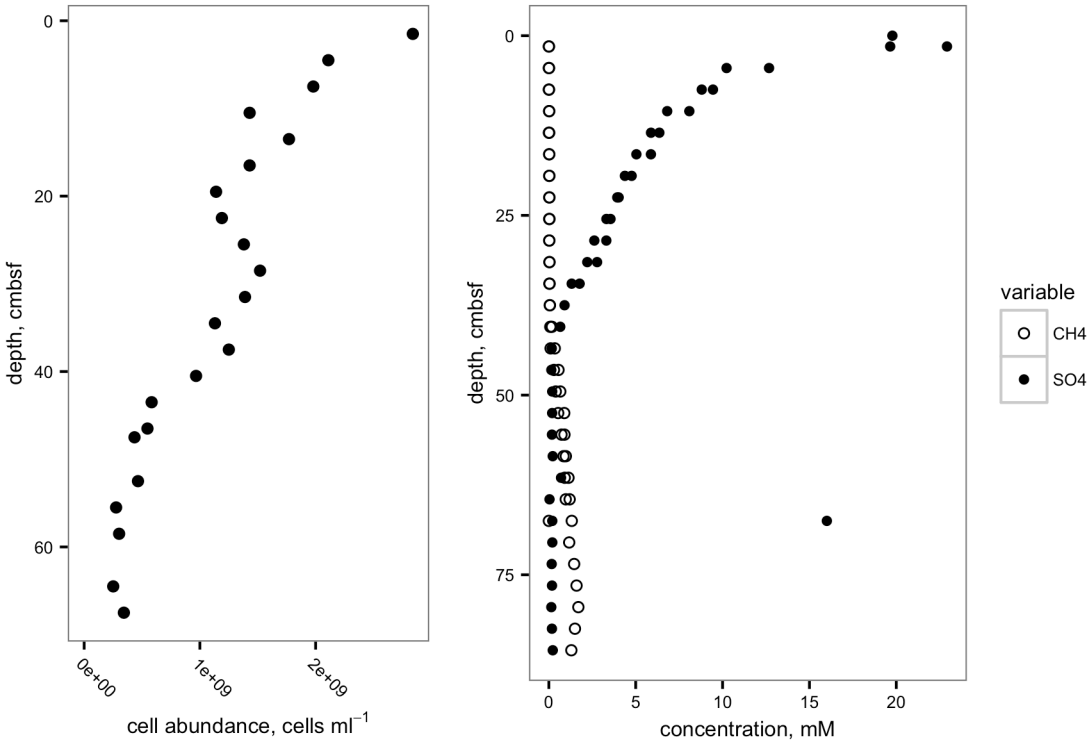


Fig S2a: Full saturation curves from 2013, using the Quantifluor ST fluorescence detector.

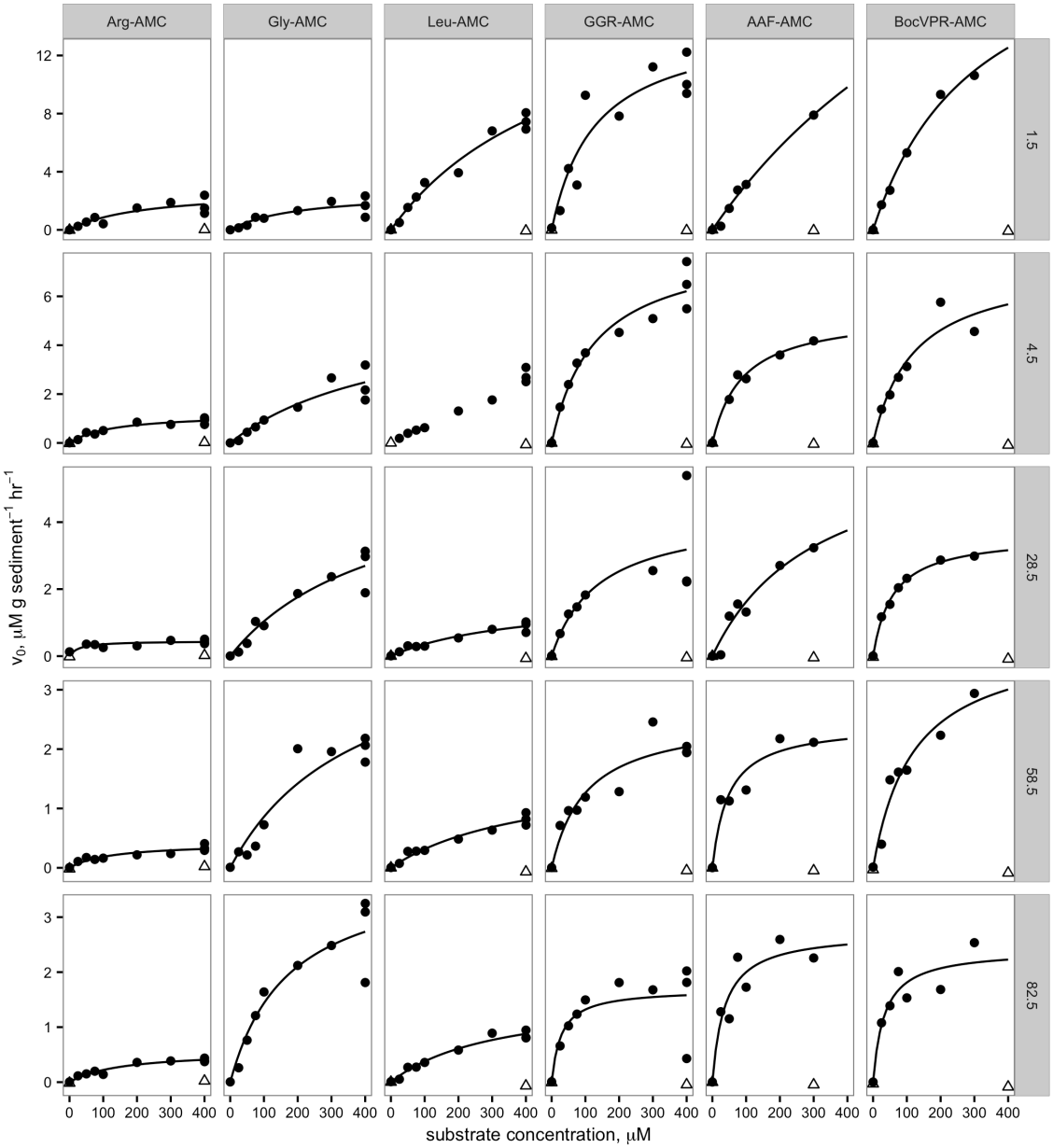


Fig S2b: Saturation curves from 2014, taken using the microplate reader method.

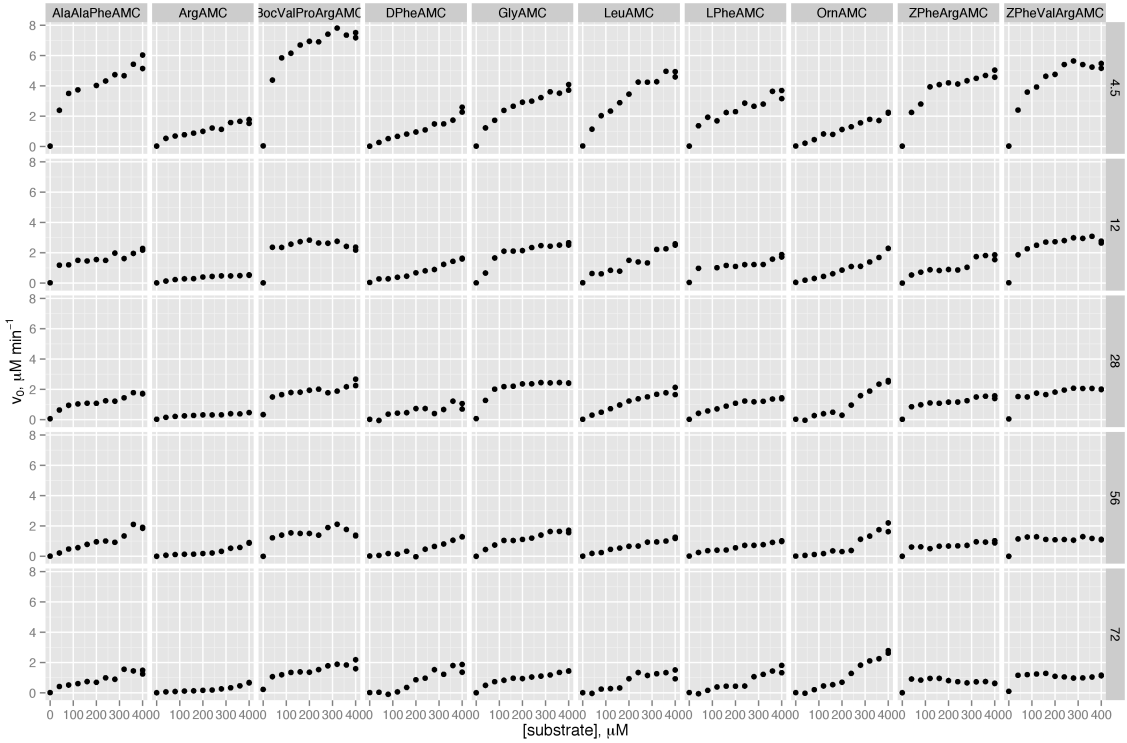
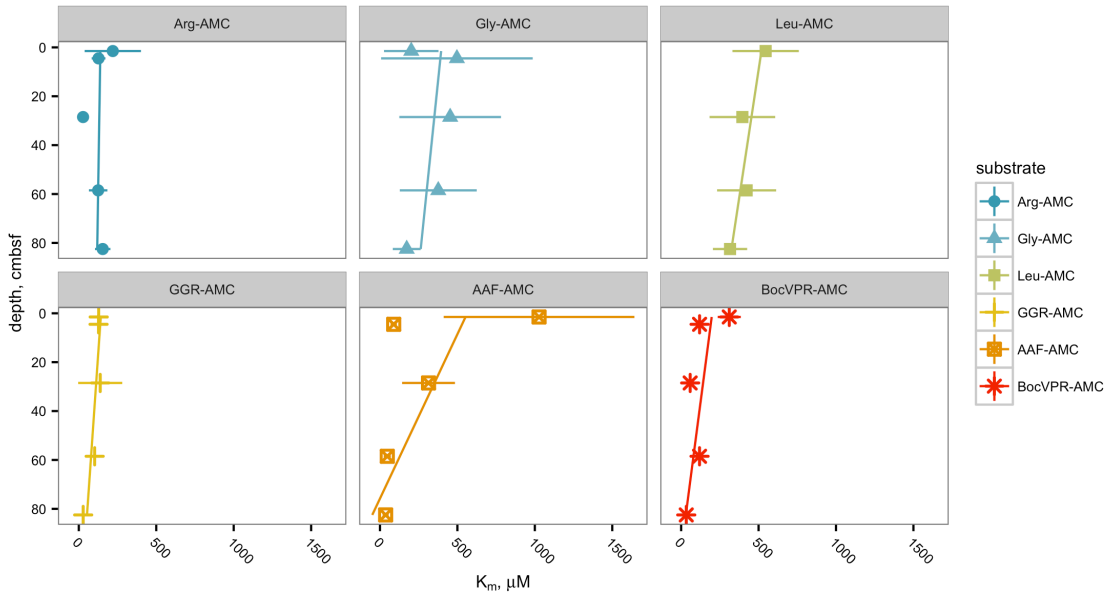
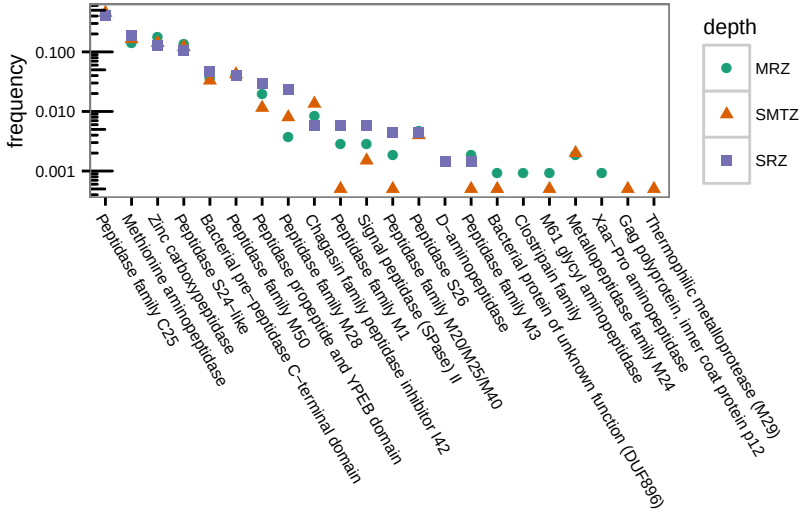
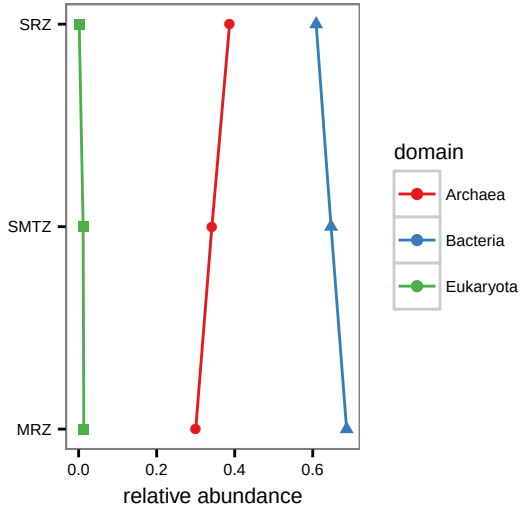


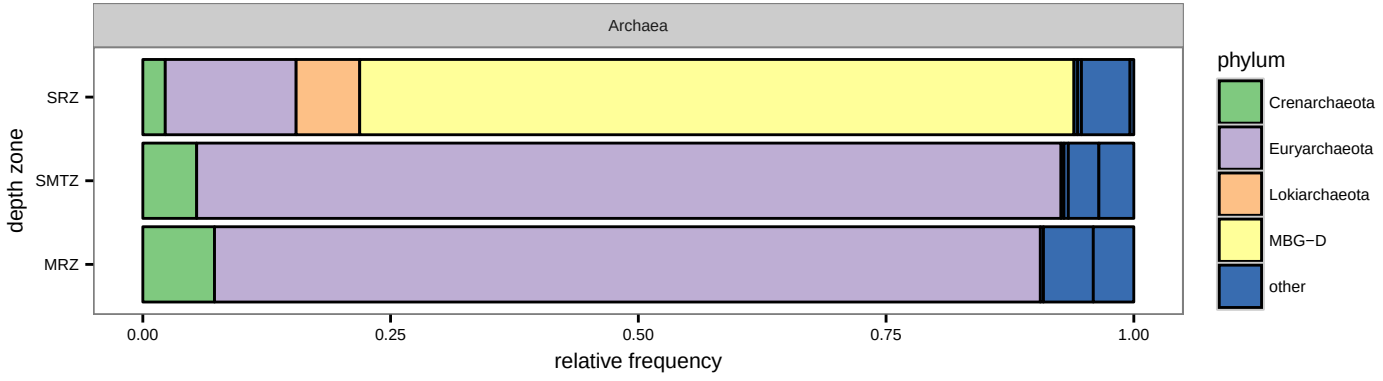
Fig S3: Km values as a function of depth. Error bars represent the standard error of the nonlinear least squares estimate of K_m.

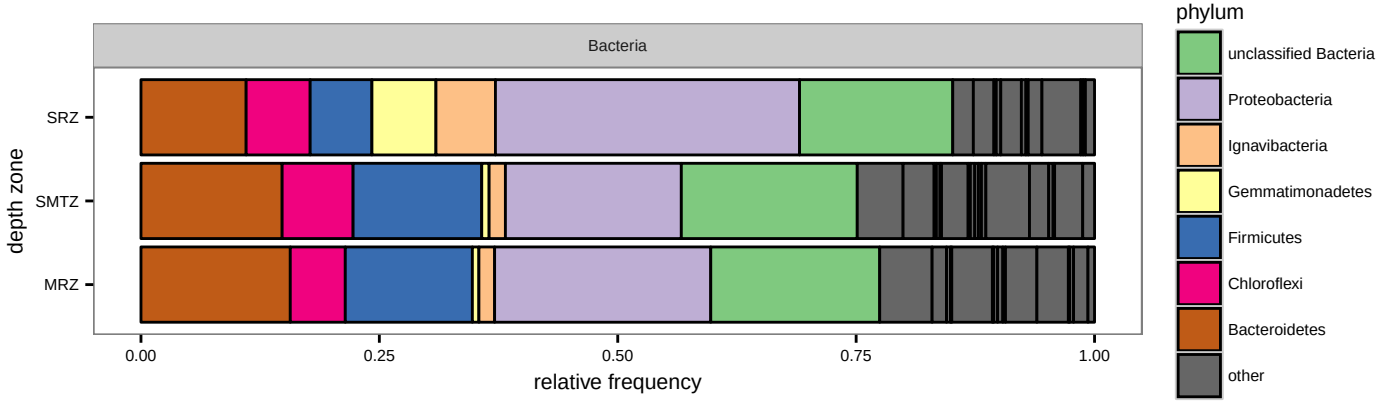


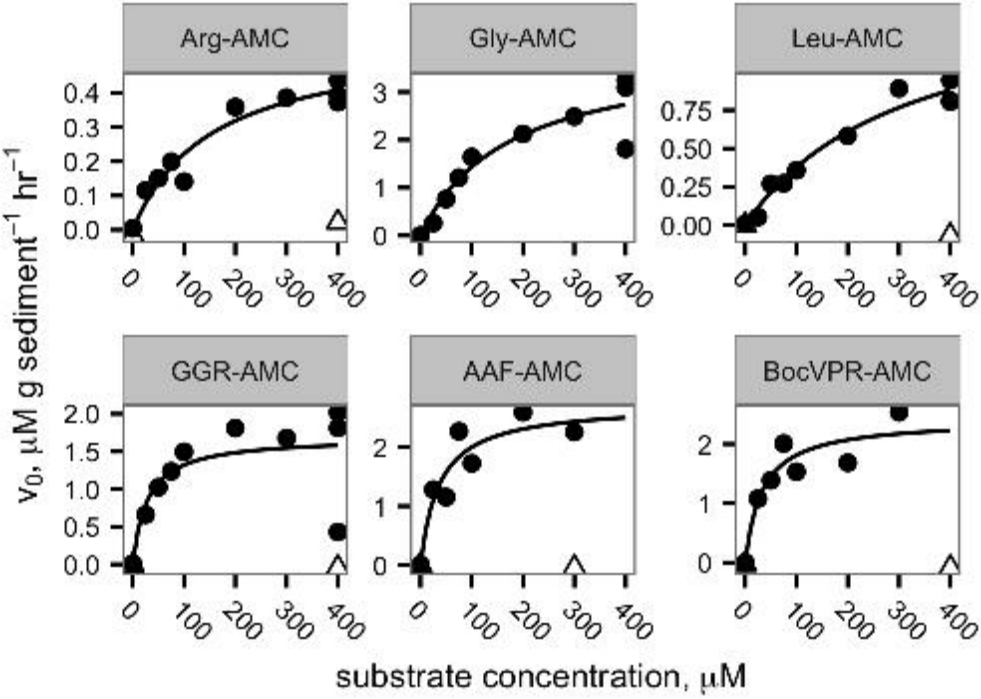
698

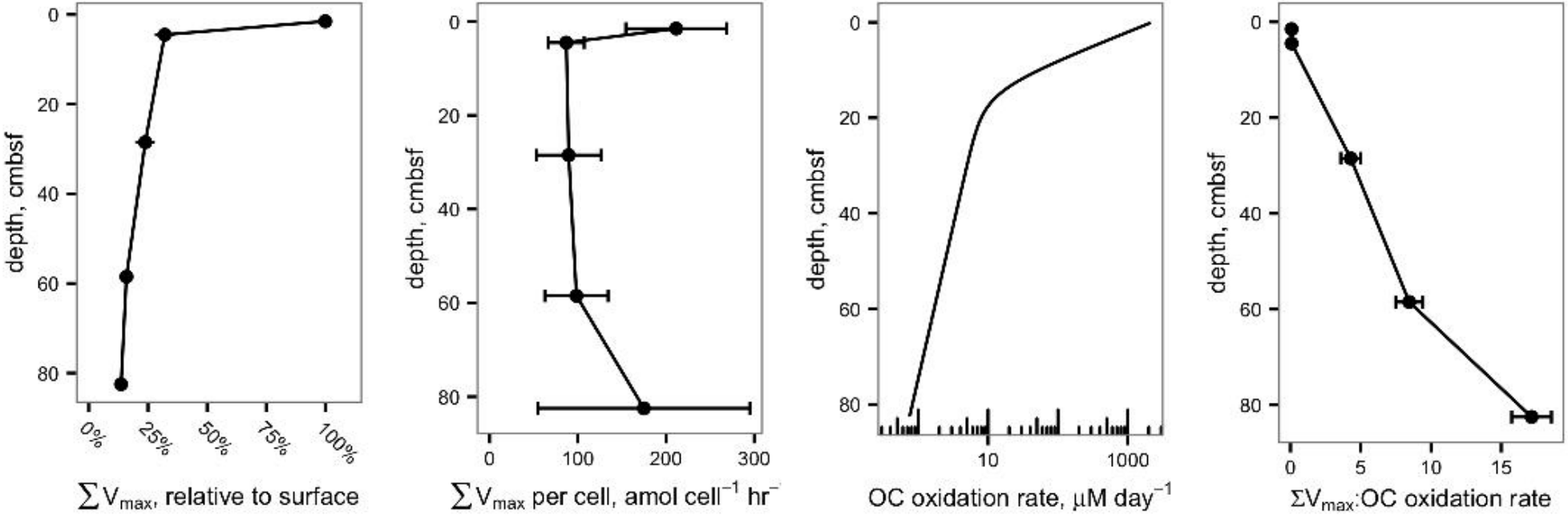

frequency
0.100
0.010
0.001
Peptidase family C25
Methionine aminopeptidase
Zinc carboxypeptidase
Peptidase S24-like
Bacterial pre-peptidase C-terminal domain
Peptidase family M50
Peptidase propeptide and YPEB domain
Chagasin family M28
Peptidase family M1
Signal peptidase (SPase) II
Peptidase family M20/M25/M40
Peptidase S26
D-aminopeptidase
Peptidase family M3
Bacterial protein of unknown function (DUF896)
Clostripain family
M61 glycyl aminopeptidase
Metallopeptidase family M24
Xaa-Pro aminopeptidase
Gag polyprotein, inner coat protein p12
Thermophilic metalloprotease (M29)
depth
MRZ
SMTZ
SRZ

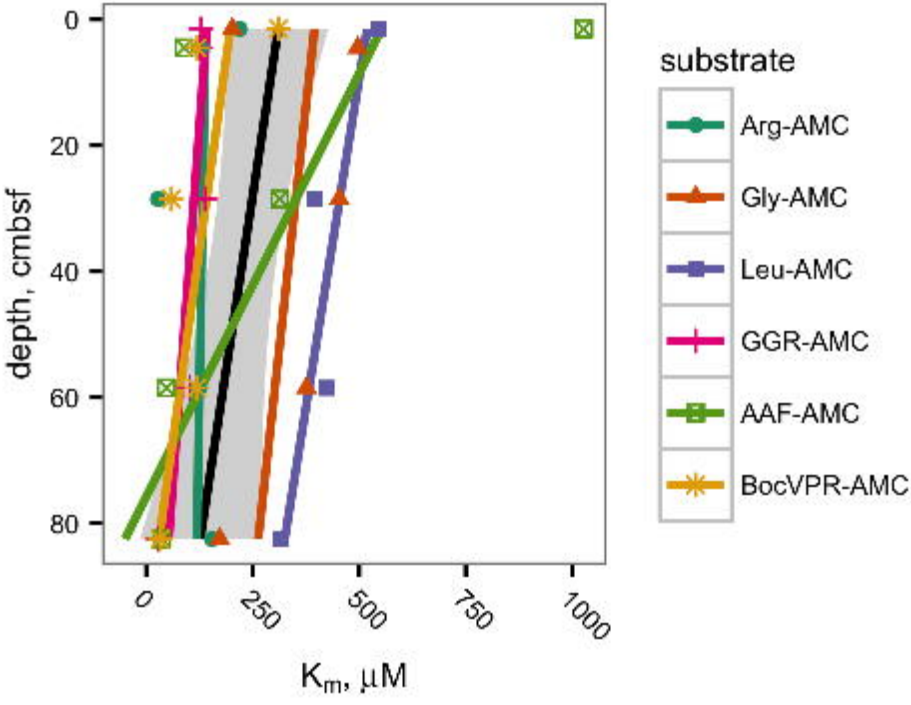

SRZ
SMTZ
MRZ
0.0
0.2
0.4
0.6
relative abundance
domain
Archaea
Bacteria
Eukaryota

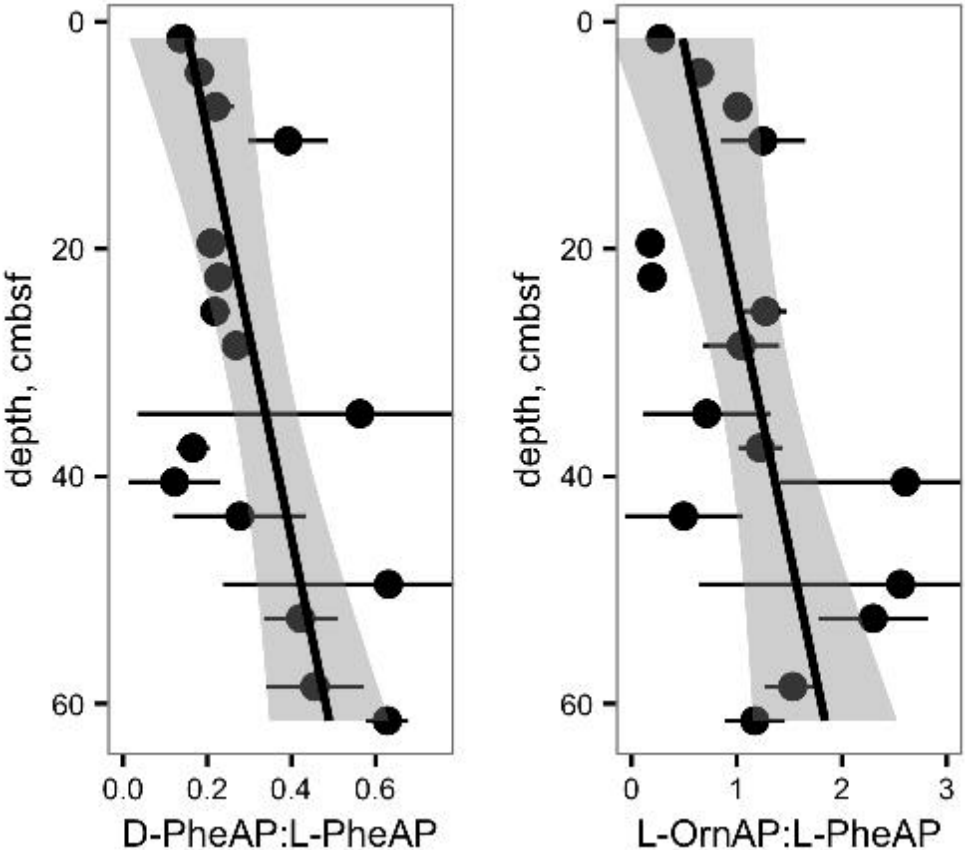

Archaea
depth zone
SRZ
SMTZ
MRZ
0.00
0.25
0.50
0.75
1.00
relative frequency
phylum
Crenarchaeota
Euryarchaeota
Lokiarchaeota
MBG−D
other


Bacteria
depth zone
SRZ
SMTZ
MRZ
0.00
0.25
0.50
0.75
1.00
relative frequency
phylum
unclassified Bacteria
Proteobacteria
Ignavibacteria
Gemmatimonadetes
Firmicutes
Chloroflexi
Bacteroidetes
other


Arg-AMC
Gly-AMC
Leu-AMC
GGR-AMC
AAF-AMC
BocVPR-AMC
v_0, μM g sediment^{-1} hr^{-1}
substrate concentration, μM


depth, cmbsf
0
20
40
60
80
0%
25%
50%
75%
100%
∑Vmax, relative to surface
0
100
200
300
∑Vmax per cell, amol cell−1 hr−
10
1000
OC oxidation rate, µM day−1
0
5
10
15
ΣVmax:OC oxidation rate


depth, cmbsf
0
20
40
60
80
0
250
500
750
1000
K_m, μM
substrate
Arg-AMC
Gly-AMC
Leu-AMC
GGR-AMC
AAF-AMC
BocVPR-AMC


depth, cmbsf
0
20
40
60
0.0
0.2
0.4
0.6
D-PheAP:L-PheAP
depth, cmbsf
0
20
40
60
0
1
2
3
L-OrnAP:L-PheAP