

# **A colloid-based multiplexed method for screening glycoside hydrolase activities using Nanostructure-Initiator Mass Spectrometry**

**Wolfgang Reindl<sup>1,2</sup>, Kai Deng<sup>1</sup>, John M Gladden<sup>1</sup>, Gang Cheng<sup>1</sup>, April Wong<sup>1</sup>, Steven W Singer<sup>1,3</sup>, Seema Singh<sup>1</sup>, Jing-Chyi Lee<sup>4</sup>, Jyun-Syu Yao<sup>4</sup>, Terry C Hazen<sup>1,5</sup>, Anup K Singh<sup>1</sup>, Paul D Adams<sup>1</sup>, Blake A Simmons<sup>1</sup> & Trent R Northen<sup>1,2</sup>**

<sup>1</sup>Joint BioEnergy Institute, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Emeryville, California, USA.

<sup>2</sup>Department of Bioenergy/GTL & Structural Biology, Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.

<sup>3</sup>Department of Geochemistry, Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.

<sup>4</sup>Division of Biotechnology and Pharmaceutical Research, National Health Research Institutes, Zhunan Town, Miaoli County, Taiwan.

<sup>5</sup>Department of Ecology, Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.

Correspondence should be addressed to T.R.N. (trnorthen@lbl.gov)

## **Abstract**

**The enzymatic hydrolysis of long-chain polysaccharides is a crucial step in the conversion of biomass to lignocellulosic biofuels. The identification and characterization of optimal glycoside hydrolases is dependent on enzyme activity assays, however existing methods are limited in terms of compatibility with a broad range of reaction conditions, sample complexity, and especially multiplexity. The method we present is a multiplexed approach based on Nanostructure-Initiator Mass Spectrometry (NIMS) that allowed studying several glycolytic activities in parallel under diverse assay conditions. Although the substrate analogs carried a highly hydrophobic perfluorinated tag, assays could be performed in aqueous solutions due colloid formation of the substrate molecules. We first validated our method by analyzing known  $\beta$ -glucosidase and  $\beta$ -xylosidase activities in single and parallel assay setups, followed by the identification and characterization of yet unknown glycoside hydrolase activities in microbial communities.**

## Introduction

Glycoside hydrolases play important roles in a multitude of biological processes, both in eukaryotes (e.g. processing of glycans in glycoproteins) (Helenius & Aebi, 2001) and prokaryotes (e.g. utilization of sugar polymers as carbon source) (Doi & Kosugi, 2004). Recently, there is growing interest in the application of this enzyme class to the development of biofuels from lignocellulosic biomass (Blanch et al., 2008): long-chain polysaccharides from plant cell walls are enzymatically hydrolyzed and the resulting sugar monomers are fermented into ethanol or advanced biofuels (Steen et al., 2010). The three major components forming plant cell walls are the polysaccharides cellulose and hemicellulose, and the highly phenolic macromolecule lignin. Cellulose is formed of linear chains of  $\beta$ -(1 $\rightarrow$ 4)-linked D-glucose units, while hemicelluloses consist mainly of mixtures of pentoses with D-xylose and D-arabinose being the most abundant (Pauly & Keegstra, 2010). Several enzymatic activities are required for the complete hydrolysis of glycosidic bonds in the polysaccharides. Cellulose is hydrolyzed into glucose through the concerted action of three classes of enzymes: endoglucanases randomly produce free ends from cellulose fibrils that get degraded by exoglucanases releasing cellobiose which in turn is hydrolyzed by  $\beta$ -glucosidases producing glucose (Lynd et al., 2002). Hemicelluloses are degraded by a complex class of multi-domain enzymes known as hemicellulases (Shallom & Shoham, 2003). Microbial communities (e.g. fungi or bacteria) capable of growing on lignocellulose have gained increasing attention as source for these glycoside hydrolases (Lynd et al., 2002).

The identification and characterization of glycoside hydrolases is dependent on enzyme activity assays. Most existing methods for screening samples are based on changes of the spectroscopic properties (e.g. absorbance, fluorescence, etc.) of a substrate analog upon hydrolysis (Sharrock, 1988). However, there are various inherent limitations to these approaches: the overlap of absorption or emission spectra of fluorescent labels greatly complicates the analysis of more than one reaction at a time, the optical density of complex

microbial samples interferes with spectroscopic measurements, and substrate analogues often have only limited thermal or temporal stability. Due to its high sensitivity and specificity mass spectrometry is capable of detecting large numbers of enzymatic reaction products in parallel, however the complexity of crude cell lysates also complicates analysis and often requires time-consuming chromatographic steps that lower sensitivity and throughput. Recently, surface-based mass spectrometry techniques have gained increasing attention as they allow for rapid analysis. For example, self-assembled monolayers for matrix assisted laser desorption ionization time of flight (SAMDI-TOF) mass spectrometry was used for the analysis of glycoside transferase activity on gold surfaces (Ban & Mrksich, 2008), a Nanostructure-Initiator Mass Spectrometry (NIMS) (Northen et al., 2007) based enzymatic activity assay (Northen et al., 2008) was successfully applied for the detection of  $\beta$ -1,4-galactosidase enzymatic activity, or most recently glycan arrays on aluminum oxide-coated glass slides were used to study glycoside hydrolases and determine antibody binding constants (Chang et al., 2010).

Here we report the application of NIMS to a new method based on self-assembled colloids for the multiplexed identification and characterization of glycoside hydrolase activities in a wide range of samples ranging from isolated enzymes to complex crude environmental samples. Due to colloid-formation of the amphiphilic substrate analogs, all reactions could be carried out in solution in tubes or microwell plates. Using this technique, we were able to study isolated  $\beta$ -glucosidase and  $\beta$ -xylosidase activities in separate and multiplexed enzyme assays. Additionally, we simultaneously identified and characterized  $\beta$ -glucosidase, exo-/endoglucanase, and hemicellulase activity directly from crude environmental samples generated by microbial communities. Our approach presents a rapid tool for the identification of yet unknown enzymes involved in the degradation of biomass.

## RESULTS

### NIMS analysis of glycoside hydrolase activity

Glycoside hydrolase assays were performed with amphiphilic substrate analogs and cleavage reactions were subsequently analyzed using NIMS (**Fig. 1**). Used substrates comprise a polar sugar head group and a highly hydrophobic perfluorinated tail (F17); the sugar groups used in this study were based on hexoses, i.e. cellobiose (CB) and cellotetraose (CT), or pentoses, i.e. xylobiose (XB) (**Supplementary Fig. 1**). Due to their amphiphilic character substrates spontaneously formed colloids, most likely micelles (**Figs. 1a** and **2**). This signifies, that despite containing a large hydrophobic moiety, substrates could be dissolve in aqueous solution enabling enzymatic reactions to be performed in standard buffers in conventional tubes or multiwell plates with one or more sugar substrates in parallel (**Fig. 1a**). After enzymatic cleavage samples were spotted onto a NIMS surface without any further pretreatment steps, even if they contained crude environmental extracts (**Fig. 1b**). The substrates selectively bind to the surface via fluororous-phase-interactions, while other sample components (i.e. enzymes, cleaved off sugar units, salts) remained in the aqueous phase. After a short incubation period the sample drop could be removed, leaving only the substrates on the chip (**Fig. 1b**) which has been found to enhance signal intensities (Reindl & Northen, 2010). Next, samples were analyzed by NIMS in a MALDI mass spectrometer, where laser irradiation leads to vaporization of the initiator liquid on the chip surface and subsequent transfer of applied samples into the gas phase (**Fig. 1c**). Acquired spectra show pairs of substrate and product signals (**Fig. 1d**), such that enzymatic activities can always be measured as product-to-substrate ratios, which is completely independent from total signal intensities.

Colloid-formation was analyzed for the cellobiose substrate (CB-F17) by small-angle neutron scattering (SANS) and dynamic light scattering (DLS) (**Fig. 2**). Both techniques showed a consistent radius for the detected main particles of 27.5 Å (polydispersity 0.14%) for SANS (**Fig. 2a**) and 30.0 Å (polydispersity 0.12%) for DLS (**Fig. 2b**). With an estimated molecule size

of ~24 Å for CB-F17, the recorded diameters of ~55-60 Å suggest the formation of micelles. A small portion of particles showed sizes bigger than 1000 Å, probably due to aggregation (**Figs. 2a** and **2b**). Micelles may reflect more natural reaction conditions in comparison to isolated substrate molecules as many glycoside hydrolases utilize insoluble solid substrates (e.g. cellulose fibrils) (Lynd et al., 2002; Shallom & Shoham, 2003).

Two substrates were utilized to validate and illustrate the range of assays that can be prepared using this method. Cellobiose (CB-F17) or xylobiose (XB-F17) were used to analyze  $\beta$ -glucosidase or  $\beta$ -xylosidase activity of commercially available enzymes in separate setups (**Fig. 3**). We first characterized the temperature and pH optima. For the  $\beta$ -glucosidase an optimal reaction temperature of 50-60°C (**Fig. 3a**) and a pH optimum at around pH 4.8 (**Fig. 3b**) were detected, while the  $\beta$ -xylosidase showed highest activity at a temperature of 35°C (**Fig. 3a**) and at around pH 7.5 (**Fig. 3b**). These data match the known reaction optima specified by the manufacturers. Kinetic studies were performed by analyzing aliquots of each reaction sample at various time points (**Fig. 3c**). Product inhibition was studied by analyzing  $\beta$ -glucosidase activity in the presence of untagged cellobiose. As expected there was competition for conversion with increasing amounts of non-perfluorinated cellobiose with an  $IC_{50}$  of about 15-20 mM at the used assay conditions (**Fig. 3d**).

### **Multiplexed glycoside hydrolase assays**

It is difficult to discriminate between several products or simultaneous reactions using conventional spectroscopic methods since these reactions typically yield the same read-outs and there are only a few unique absorption or emission patterns that can be resolved separately. In contrast, mass spectrometry can resolve thousands of ions and therefore is well suited for the parallel detection of large numbers of reaction products as long as the substrates and occurring products differ in mass. Both  $\beta$ -glucosidase and  $\beta$ -xylosidase activities were detected simultaneously utilizing an enzyme cocktail comprised of equal amounts of CB-F17 and XB-F17

(**Fig. 4**). For a non-enzyme control virtually only signals for cellobiose and xylobiose were detectable (**Fig. 4a**). After incubation with  $\beta$ -glucosidase, cellobiose was almost completely converted to glucose, while there was only a slight increase in the xylose signal, probably due to side reactions of the used enzyme (**Fig. 4b**). As expected, for  $\beta$ -xylosidase there was conversion of xylobiose to xylose and no enzymatic cleavage of cellobiose (**Fig. 4c**). Incubation of the two substrates with a  $\beta$ -glucanase/xylanase enzyme cocktail leads to complete cleavage of both (**Fig. 4d**). These data proof the applicability of the presented approach for the analysis of several glycoside hydrolase activities in parallel.

### **Multiplexed detection of glycoside hydrolase activities in environmental samples**

Our multiplexed micelle-based assay system was used for screening for potential glycoside hydrolase activities in environmental samples. The following workflow was applied for obtaining these samples (**Fig. 5a**): soil samples were collected at various sites with different soil characteristics, (**Fig. 5a, left**); a small amount of each sample was used to inoculate minimal growth medium with switchgrass as sole carbon source (**Fig. 5a, middle**); after one week of incubation the remaining soil particles were removed by centrifugation and the resulting supernatants containing secreted enzymes (secretome) of any microbes present were used for further analysis (**Fig. 5a, right**). The secretomes of various environmental samples were incubated with a mixture of equal amounts of CB-F17 and XB-F17 for 1 hour at 50°C and analyzed for conversion of these substrates. This approach proved to be effective, and at least one of the targeted activities was detected in the majority of samples tested, while several samples showed clear substrate conversion for both enzymatic activities (**Fig. 5b**). In total a broad range of activities was exhibited by the samples. Some samples, e.g. chicken manure (CHM) or a soil sample from under an oak tree (OAK), showed evidence of  $\beta$ -glucosidase activity, while in other samples, e.g. a light woodland soil sample (WS) or Zamora (Z), the  $\beta$ -xylosidase activity was dominant (**Fig. 5b**). The Jepson Prairie (JP) sample turned out to be the

most effective in terms of overall liberation of monomeric sugars with  $87.1\% \pm 2.1\%$  cellobiose and  $92.3\% \pm 0.4\%$  xylobiose conversion (**Fig. 5b**).

### Profiling of an environmental sample

The next step was the characterization of enzymatic activities present in an environmental sample. In order to get an initial survey of the enzymes that may be present, a range of purified  $\beta$ -glucosidases (NS50010 and UBG), exoglucanases (two expression versions of Cs\_GH5), and endoglucanases (Pr\_GH5, Cel5A, Cel9A) were tested with cellobiose or cellotetraose alone (**Supplementary Fig. 2**). When incubating these enzymes with cellobiose, the two  $\beta$ -glucosidases hydrolyzed a significant amount of the cellobiose, whereas the exo- and endoglucanases, which are supposed to act on larger polysaccharides (Dashtban et al., 2009), exhibited minimal activity (**Supplementary Fig. 2a**). Surprisingly, the endoglucanase Cel5A was able to convert  $26.2\% \pm 6.6\%$  CB-F17 (**Supplementary Fig. 2a, bar 6**). The incubation with cellotetraose gave a more complex result (**Supplementary Fig. 2b**). Both  $\beta$ -glucosidases were able to stepwise degrade CT-F17 from the tetraose to cellotriose, cellobiose, and glucose. The preferred reaction of  $\beta$ -glucosidases is the conversion of cellobiose to glucose, but the production of sugar monomers from longer cellodextrins is also catalyzed (Dashtban et al., 2009), presumably at a slower reaction speed. The most abundant signal for both  $\beta$ -glucosidases was cellotriose, followed by cellobiose and glucose (**Supplementary Fig. 2b, bars 1/2**). The main reaction for exo- and endoglucanases with cellotetraose is hydrolysis into cellobiose units (Dashtban et al., 2009; Damude et al., 1996). Accordingly, both groups of enzymes showed cellobiose as the most significant signal present (**Supplementary Fig. 2b, bars 3-7**). However, in some cases significant signals for glucose were detectable, showing that these enzymes are also capable of cleaving off cellotriose units directly (**Supplementary Fig. 2b, bars 3/5-7**). This effect was more significant for the endoglucanases. Additionally in accordance with the results from the previous analysis using CB-F17, Cel5A shows the biggest



glucose signal as formed cellobiose can potentially also be cleaved on to glucose (**Supplementary Fig. 2b, bar 6**). Exo- and endoglucanases basically showed no triose signal, indicating that in general these two groups of enzymes cannot cleave off sugar monomers (**Supplementary Fig. 2b**).

The JP secretome was identified as the environmental sample with the  $\beta$ -glucosidase and  $\beta$ -xylosidase activities (**Fig. 5b**). It had been obtained by incubation of the JP environmental sample in minimal medium with switchgrass as sole carbon source (JP SG). The activity profile of JP SG was compared to another JP secretome that was obtained by using microcrystalline cellulose as sole carbon source (JP MC). Both secretomes were incubated with CT-F17 and product formation was monitored over time (**Supplementary Fig. 3**). In the case of JP SG it seems that CT-F17 gets degraded stepwise to cellotriose, cellobiose, and glucose (**Supplementary Fig. 3, bars 1/4/6**), which resembles the degradation of cellotetraose by isolated  $\beta$ -glucosidases (**Supplementary Fig. 2b, bars 1/2**). For the JP MC sample, the cellotriose signal is less prominent and the apparent main reaction is the immediate conversion of cellotetraose to cellobiose or glucose (**Supplementary Fig. 2, bars 2/3/5/7**), which is very similar with the results seen for cleavage of cellotetraose by exo- or endoglucanases (**Supplementary Fig. 2b, bars 3-7**). The data suggests that exposure to different carbon sources can trigger the expression of different sets of glycoside hydrolases.

For a detailed multiplexed characterization of the JP MC secretome, the extract was incubated with a mixture of XB-F17 and CT-F17 for 30 min at 50°C or 80°C at pH 3, 5, 7, and 9, and in the presence of 0%, 10%, 20%, and 30% ionic liquid (IL; ethyl-3-methyl imidazolium (EMIM) acetate) (**Fig. 6**). The ionic liquid reagent is a molten salt used for pre-treatment of biomass and IL-tolerance is highly desirable for enzymes and microbes involved in biomass degradation (Datta et al., 2010). The  $\beta$ -xylosidase activity showed a clear preference for higher temperatures under every reaction condition tested (**Fig. 6a and 6b**). This clearly shows that the  $\beta$ -xylosidase(s) that are present in the JP MC secretome are thermophilic enzymes. Essentially

no  $\beta$ -xylosidase activity was detectable at pH 3, while XB-F17 was converted at pH 5, 7, and 9. At 80°C more than 90% of XB-F17 was degraded at pH 5, 7, and 9. The 50°C data show a pH optimum at around pH 7 (**Fig. 6a**). The IL profile of  $\beta$ -xylosidase activity revealed a surprisingly strong tolerance. Even at 30% IL still  $32.3\% \pm 4.2\%$  of XB-F17 was converted at 80°C (**Fig. 6b**). Regarding cellulolytic activity in the samples without IL, activities were also higher in the 80°C samples (**Fig. 6c**), suggesting that in this sample there are primarily thermophile enzymes involved. Strong cellobiose and glucose signals suggest that enzymatic activity as is mainly based on exo- and endoglucanases, as seen for purified enzymes (**Supplementary Fig. 2b, bars 3-7**), but as no triose formation was observed for these enzymes (**Supplementary Fig. 2b, bars 3/5-7**) and a significant amount of triose was detected for the JP MC secretome (**Fig. 6a**), this shows that  $\beta$ -glucosidase activity is present. The pH optimum for cellulolytic activity was determined to be at pH 5 (**Fig. 6a**). In comparison to the  $\beta$ -xylosidase activity, the present cellulolytic activities were much less tolerant for IL, for the 50°C samples already there was hardly any activity detectable at 30% IL with  $89.6\% \pm 5.1\%$  CT-F17 left (**Fig. 6b**). One surprising observation was that although there was higher enzymatic activity at 80°C when no IL was present, activities were higher at 50°C if the samples contained any IL (**Fig. 6b**). We hypothesize that the particular reaction leading to the inactivation of the affected glucoside hydrolase(s) must have been kinetically favored at higher temperatures.

## DISCUSSION

We report an integrated method utilizing substrate colloids/micelles coupled with NIMS for the analysis of glycoside hydrolase activities. This method was shown to correctly determine the optimal reaction conditions for already known enzymes, and to be suitable for the detection of activities in complex environmental samples. It is compatible with conventional microwell plate formats and has several benefits over standard assays: Multiple enzymatic activities can easily

be assayed in parallel, the only requirement is a difference in mass. This is the case, e.g. when using pentose- and hexose-based substrates at the same time. The hydrolysis of polysaccharides with identical masses (e.g. cellobiose and maltose) can be analyzed in the same run by using different tags: differences in mass can be created by using different chemical structures in the linker or by using perfluorinated tags varying in length. This indicates that potentially dozens of enzymatic conversions can be tested simultaneously using this method. The potential of this rapid parallel analysis scheme of several enzyme reactions in combination with low sample volumes (less than 1  $\mu$ l spotted per sample) makes this approach suitable for high throughput applications in the detection of enzymes that can be used for biofuel and other associated fields of development. Another advantage is that there is no need for pure samples. Even crude environmental samples containing visible soil particles (**Fig. 5a, middle**) did not interfere with the mass spectrometric analysis. Due to the unique fluororous phase interactions substrates selectively bind to the chip surface while all other sample components including soil particles or cell debris can simply be washed away *in situ*. Our assay system comprises the benefits of chromatographic sample purification, while requiring only a minimal fraction of time in comparison to standard chromatography techniques, e.g. liquid chromatography coupled mass spectrometry (LC-MS). Particles present in a sample would strongly interfere with any traditional spectrophotometric analysis due to multiple light scattering effects. Additionally, some of the used environmental samples were strongly colored (**Fig. 5a, right**), and would also potentially cause false readouts in spectrophotometric measurements. This NIMS-based assay is found to be very stable. Once spotted onto the chip surface, samples are stable for weeks, while the used substrates in absorption or fluorescence measurements sometimes are only stable for minutes. Furthermore our method could be used under a broad range of assay conditions. Reactions can be performed at basically any pH, while some substrate analogs used in standards assays are functional over a small pH range. The presence of ionic liquid did not interfere with the assay,

which is particularly important for biofuel development that is using these green solvents for biomass pretreatment.

It is noted that the method presented here has the same types of limitations as any method based on substrate analogues in that it requires chemical synthesis and shares concerns over reaction specificity vs. the actual substrates. Therefore this method is best used for rapid screening purposes and comparing relative performances, and not sample quantification.

Using this method, both crude and purified enzymes were studied for the major classes of enzymatic reactions required to deconstruct (hemi)cellulose:  $\beta$ -glucosidases, exo- and endoglucanases, and  $\beta$ -xylosidases. For a more comprehensive analysis of plant cell wall degradation, it would be interesting to use several more substrates for the analysis of additional enzymatic activities. Longer  $\beta$ -(1 $\rightarrow$ 4)-linked glucose chains would help to study endoglucanases in more detail. Longer xylose chains or also other sugars, e.g. arabinose or mannose, would aid in the identification of other enzymes involved in the degradation of hemicelluloses. While the method presented here is focused on enabling biofuel development, it can potentially be used for a wide range of other glycan assays including medical applications, e.g. screening for viruses based on profiles of neuraminidase activity.

## **METHODS**

### **Substrates.**

The substrates used in this study were cellobiose, cellotetraose, and xylobiose, all attached to a perfluorinated tag (**Supplementary Fig. 1**). Sugar molecules were purchased from the following sources: cellobiose from Sigma-Aldrich (St. Louis, MO), cellotetraose from Toronto Research Chemicals (North York, Canada), and xylobiose from TCI America (Portland, OR). Synthesis of substrates has been described in extensive detail previously (Northen et al., 2008). Briefly, a

(CH<sub>2</sub>)<sub>5</sub>-linker was coupled to the reducing end of each sugar molecule by Schmidt imidate chemistry. Hydrogenation by Pd/C removed the Carbobenzyloxy (Cbz) protection group to give a primary amine. The heptadecafluoro-1,1,2,2-tetrahydrodecyl (F17) tag was attached to a dimethyl-arginine by an amide bond forming reaction. Finally, a peptide coupling reaction linked the sugar moiety with the fluororous tag to yield the desired substrate.

### **Measuring colloid formation.**

*Small-angle neutron scattering (SANS).* SANS experiments were conducted at Oak Ridge National Laboratory (ORNL) on the CG2 (GP-SANS) instrument with a neutron wavelength of  $\lambda=4.8 \text{ \AA}$  ( $\Delta\lambda/\lambda \sim 0.14$ ). Liquid solutions of 0.35% (w/w) and 0.97% (w/w) CB-F17 in D<sub>2</sub>O were put into 1 mm quartz cells. Measurements were carried out at room temperature. Two sample-detector distances were used (4.0 and 14 m with a 40 cm detector offset), which resulted in a range of  $0.004 \text{ \AA}^{-1} < q < 0.3 \text{ \AA}^{-1}$  for the scattering vector  $q (=4\pi \sin\theta/\lambda)$ . The data was corrected for instrumental background as well as detector efficiency and put on absolute scale (cross section  $I(q)$  per unit volume in units of  $\text{cm}^{-1}$ ) by means of precalibrated secondary standard (Wignall & Bates, 1987).

*Dynamic light scattering (DLS).* DLS experiments were performed at 25°C with a 0.35% (w/w) solution of CB-F17 in D<sub>2</sub>O on a DynaPro plate reader (Wyatt Technology; Santa Barbara, CA) with a wavelength of 832.6 nm and a detection angle of 158°. The size distribution was obtained by analyzing the auto correlation function using regularization analysis (Chu, 1991).

### **Enzymes.**

The used  $\beta$ -glucosidase (NS50010) and the  $\beta$ -glucanase/xylanase mixture (NS22002) were part of the 'Biomass Kit' from Novozymes (Davis, CA). 1,4- $\beta$ -D-xylosidase from *Bacillus pumilus* was purchased from Megazyme (Wicklow, Ireland). Two different expression versions of the exoglucanase Cs\_GH5 from *Caldicellulosiruptor saccharolyticus* (Chandrasekaran et al., 2010),

the endoglucanases Pr\_GH5 from *Prevotella ruminicola* (Gardner et al., 1997), Cel5A from *Thermotoga maritima* (Chhabra et al., 2002), and Cel9A from *Alicyclobacillus acidocaldarius* (Eckert et al., 2002) were kindly provided by Joshua Park and Supratim Datta (Joint BioEnergy Institute, Emeryville, CA). An additional undisclosed  $\beta$ -glucosidase (UBG) was also included.

### **Enzymatic activity assays.**

*Separate analysis of  $\beta$ -glucosidase and  $\beta$ -xylosidase activity.* The separate characterization of  $\beta$ -glucosidase and  $\beta$ -xylosidase activity was carried out in reaction volumes of 20 – 50  $\mu$ l, by mixing 0.1 mM cellobiose with 8  $\mu$ g of NS50010 solution in 50 mM sodium acetate buffer (pH 4.8), or 0.1 mM xylobiose with 0.13  $\mu$ g of 1,4- $\beta$ -D-xylosidase in 50 mM potassium phosphate buffer (pH 7.5) with 1 mg/ml BSA. For the competition studies with cellobiose, samples additionally contained 10, 25, or 50 mM of unlabeled cellobiose (Sigma-Aldrich; St.Louis, MO). Other buffers used for the determination of the pH optima were McIlvaine's citrate/phosphate buffer (pH 2.6 and 6.0), 50 mM Tris buffer (pH 9.0) for NS50010, and McIlvaine's citrate/phosphate buffer (pH 6.0), 50 mM Tris buffer (pH 9.0), 50 mM sodium borate buffer (pH 11.0), all with 1 mg/ml BSA for the  $\beta$ -xylosidase. Samples were incubated for 1 h at 35°C ( $\beta$ -xylosidase) or 50°C (NS50010), for determination of temperature optima samples were additionally incubated at the given temperatures. Reactions were quenched by adding one sample volume of ice-cold methanol. For monitoring the enzymatic reactions over time, small aliquots were taken out of a bigger reaction volume at the given time points, quenched by adding one aliquot volume of methanol, and kept on ice until further analysis.

*Multiplexed  $\beta$ -glucosidase/ $\beta$ -xylosidase assay.* For the simultaneous analysis of  $\beta$ -glucosidase and  $\beta$ -xylosidase activity a mixture of 0.1 mM cellobiose and 0.1 mM xylobiose was incubated with 8  $\mu$ g NS50010 solution in 50 mM sodium acetate buffer (pH 4.8), with 2.4 mg NS22002 solution in 50 mM sodium acetate buffer (pH 6.0), or with 0.26  $\mu$ g  $\beta$ -xylosidase in 50 mM potassium phosphate buffer (pH 7.5) with 1 mg/ml BSA in a total reaction volume of 20  $\mu$ l

each. All samples were incubated for 1 h at 40°C. Reactions were quenched by adding one sample volume of ice-cold methanol.

*Characterization of  $\beta$ -glucosidase, exo- and endoglucanase activity.* To assess the spectrum of potentially catalyzed reactions by the three different groups of cellulose degrading enzymes, two  $\beta$ -glucosidases, two exo-glucanases, and three endo-glucanases were incubated for 15 min with either 0.15 mM cellobiose or 0.02 mM cellotetraose in a total reaction volume of 25  $\mu$ l. The used enzymes, buffers and temperatures are shown in **Supplementary Table 1**. All enzymes were used at their optimal reaction temperatures. Reactions were quenched by adding one sample volume of ice-cold methanol.

**Supplementary Table 1**

Enzyme name	Enzyme type	Enzyme amount	Buffer	Temp
NS50010	$\beta$ -glucosidase	1.2 mg	10 mM sodium acetate (pH 4.8)	50°C
UBG	$\beta$ -glucosidase	2 $\mu$ g	Mcllvaine's citrate/phosphate buffer (pH 7.0)	50°C
Cs_GH5-1	exo-glucanase	2.9 $\mu$ g	Mcllvaine's citrate/phosphate buffer (pH 6.0)	80°C
Cs_GH5-2	exo-glucanase	3.2 $\mu$ g	Mcllvaine's citrate/phosphate buffer (pH 6.0)	80°C
Pr_GH5	endo-glucanase	3.8 $\mu$ g	Mcllvaine's citrate/phosphate buffer (pH 6.0)	37°C
Cel5A	endo-glucanase	1.9 $\mu$ g	Mcllvaine's citrate/phosphate buffer (pH 6.0)	80°C
Cel9A	endo-glucanase	12.5 $\mu$ g	Mcllvaine's citrate/phosphate buffer (pH 6.0)	70°C

### Environmental samples.

Two commercially available compost samples and four soil samples collected at various sites in Berkeley, CA or Walnut Creek, CA were used: cow (COM) and chicken (CHM) manure, soil mixed with leafs from under an oak tree (OAK), clay-rich soil (CS), mixed organic compost (MOC), and a sample from light woodland (WS). These samples were manually fragmented with a hammer. Additionally two samples collected from two different green waste compost sites in northern California were included: Zamora (Z) and Jepson Prairie (JP). Zamora compost collection and processing has been described previously (Allgaier et al., 2010). The latter compost sample was collected at the Jepson Prairie Organics facility (Vacaville, CA). This facility

processes municipal green waste in turned and watered windrows. Compost was collected from 7, 30, and 60 day windrows. The 7 day windrow was in the mesophilic stage of composting and slightly warm to the touch. The 30 and 60 day windrows were in the thermophilic composting stage, hot to the touch, and steaming. The top 12 inches of each windrow was removed with a spade and the exposed biomass underneath was packed into 50ml Falcon tubes, stored at room temperature during transport, and then frozen at -80°C.

### **Cultivation and extraction of environmental samples.**

200 mg of small fragments ( $<3\text{ mm}^3$ ) of each of the eight environmental samples were used to inoculate 50 ml liquid cultures containing 1 g of extracted switchgrass (washed exhaustively with water and ethanol in a soxhlet apparatus) in M9 minimal medium with trace elements (M9TE; DeAngelis et al., 2010). Cultures were incubated at 37°C or 60°C with shaking at 200 rpm for one week. Two ml of each sample was then collected: each sample was placed to a 2 ml tube, spun at 21,000x g for 5 minutes, the supernatant containing secreted enzymes (secretome) was aliquoted to a new tube, and stored at 4°C till further use in enzymatic activity assays.

The Jepson Prairie (JP) compost sample was also used to generate switchgrass-adapted thermophilic consortia. Briefly, JP was used to inoculate liquid cultures containing switchgrass as the sole carbon source and grown at 60°C. The compost-derived microbial communities were allowed to adapt to switchgrass in liquid culture for a total of 32 weeks (16 x two week passages). After adaptation, the supernatant (JP SG secretome) of the liquid cultures were collected for enzyme assays as described below. An additional liquid culture was established using 2 ml of passage #15 of the JP switchgrass-adapted culture to inoculate a culture containing 50 ml of M9TE and 0.5 g of microcrystalline cellulose (Sigma-Aldrich; St. Louis, MO). This culture was incubated for two weeks at 60°C with shaking at 200 rpm. Ten ml of the culture was placed into five 2 ml tubes, spun at 21,000 x g for 5 minutes, the supernatant (JP MC



secretome) was collected and filtered through a 0.2  $\mu\text{m}$  filter, and stored at 4°C until used in enzyme assays.

### **Direct analysis of enzyme activity in environmental samples.**

*Screening for  $\beta$ -glucosidase/ $\beta$ -xylosidase activity.* For the detection of glycoside hydrolase activity a mixture of 0.03 mM cellobiose and 0.03 mM xylobiose were mixed with 75% (v/v) per environmental sample secretome in 50 mM sodium acetate buffer (pH 5.0). The total reaction volume was 30  $\mu\text{l}$ . Samples were incubated for 1 h at 50°C and reactions quenched by adding four reaction volumes of ice-cold methanol.

*Influence of carbon source on enzyme expression.* To determine if different carbon sources in the growth medium lead to differential enzyme expression/activities the JP SG and JP MC secretome was tested for glucoside hydrolase activity. To this end 0.01 mM of cellotetraose was mixed with 72% (v/v) per secretome in 40 mM sodium acetate buffer (pH 5.0). For the MC secretome additional samples with 36% (v/v) and 18% (v/v) environmental sample were included. Samples were incubated at 80°C and small aliquots were taken out of a bigger reaction volume at 5min, 15min, and 30min, quenched by adding one aliquot volume of methanol, and kept on ice until further analysis.

*Enzymatic profiling of the JP MC environmental sample.* For analysis of the activity profile for cellulose-degrading glucoside hydrolase and  $\beta$ -xylosidase of the JP MC sample a mixture of 0.005 mM cellotetraose and 0.02 mM xylobiose were mixed with 42% (v/v) of the JP MC secretome. For determination of the pH profile reactions were carried out in McIlvaine's citrate/phosphate buffer (pH 3.0, 5.0, 7.0) or 10 mM Tris buffer (pH 9.0). For determination of the ionic liquid (IL) tolerance 0%, 10%, 20%, or 30% (v/v) ethyl-3-methyl imidazolium (EMIM) acetate (Sigma-Aldrich; St. Louis, MO) were added to the sample in McIlvaine's citrate/phosphate buffer (pH 5.0). All samples were incubated for 30 min at 50°C and 80°C. Reactions were quenched by adding one sample volume of ice-cold methanol.

### **Fabrication of NIMS chips.**

The production of NIMS chips has been described in extensive detail elsewhere (Woo et al., 2008) In brief, a 4" silicon wafer (single-sided polished P/Boron, orientation <1-0-0>, resistivity 0.01-0.02  $\Omega$  cm, thickness  $525 \pm 25$   $\mu$ m) obtained from Silicon Quest International (Santa Clara, CA) was cut into a 70x70 mm square and cleaned thoroughly with methanol, followed by anodic etching with 25% hydrofluoric acid in ethanol in a custom made Teflon etching chamber using extreme caution. A current of 2.4 A was applied for 15 minutes. After etching, chips were coated by adding 400  $\mu$ l of the initiator liquid bis(heptadecafluoro-1,1,2,2-tetrahydrodecyl)tetramethyl-disiloxane (Gelest; Morrisville, PA) for 20 minutes. Excess initiator was blown off with nitrogen.

### **Nanostructure-Initiator Mass Spectrometry (NIMS).**

In each case 0.5 – 1  $\mu$ l per quenched reaction sample was spotted onto the NIMS surface and removed after an incubation of ~30s. A grid drawn manually on the NIMS chip using a diamond-tip scribe helped with spotting and identification of sample spots in the spectrometer. Chips were loaded using a modified standard MALDI plate. NIMS was performed on a 4800 MALDI TOF/TOF mass analyzer from Applied Biosystems (Foster City, CA). In each case signal intensities were identified for the ions of the used substrates and occurring products. Enzyme activities were determined by forming product-to-substrate ratios.

### **ACKNOWLEDGMENTS**

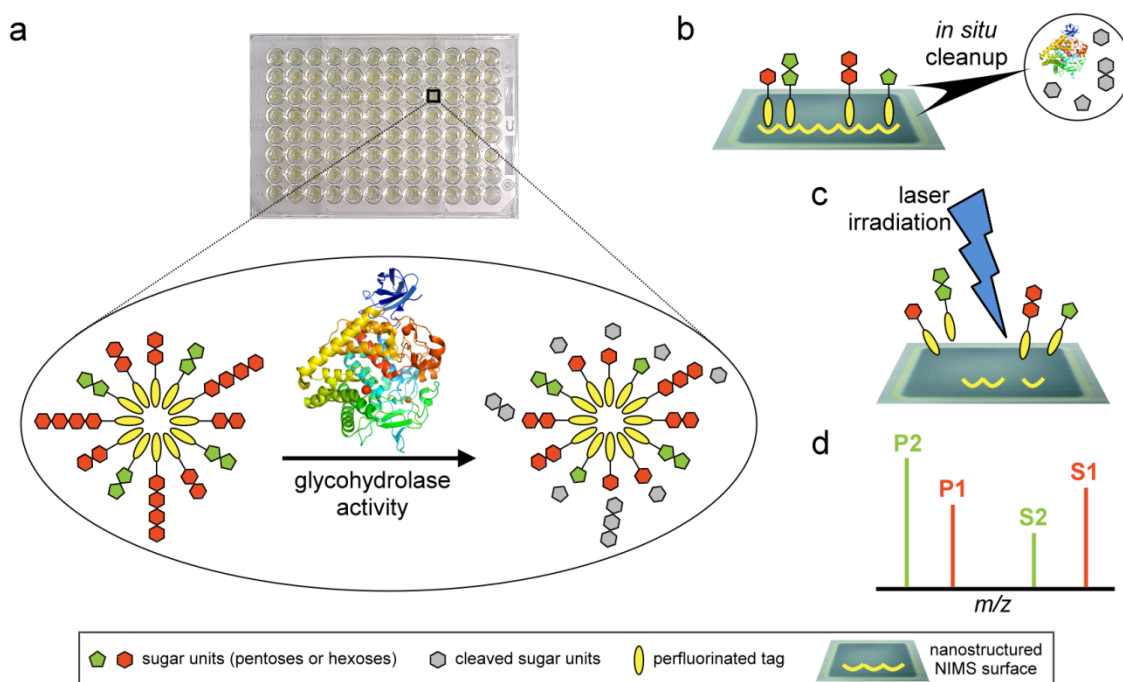
This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy. We thank Joshua Park and Supratim Datta for providing enzymes and helpful discussions. G.C. thanks Dr. Yuri Melnichenko for discretionary neutron beam time

at Oak Ridge National Laboratory (ORNL). Measurements at ORNL's High Flux Isotope Reactor were sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy and the EPSCoR Neutron Travel Fellowship through University of Tennessee to G.C.

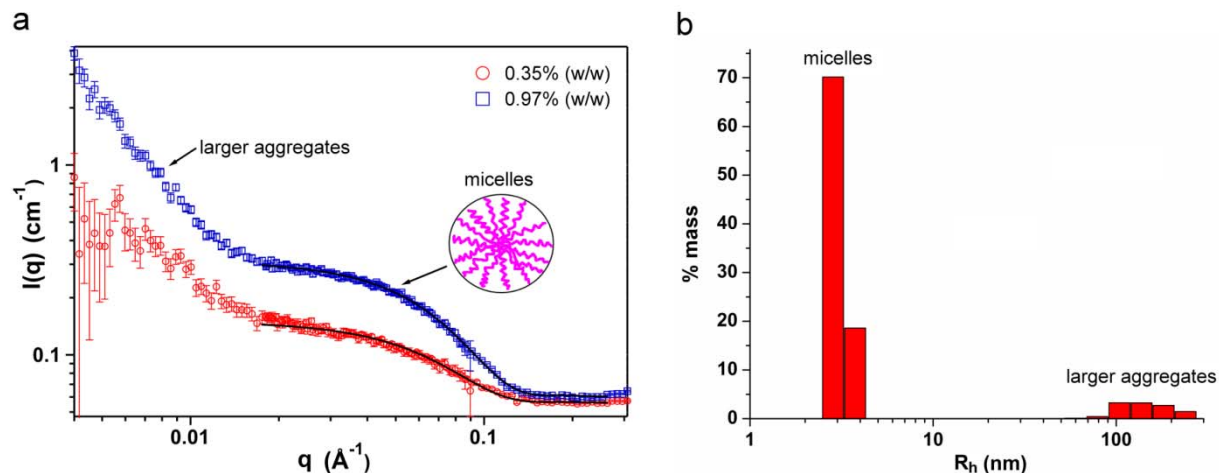
## **AUTHOR CONTRIBUTIONS**

T.R.N. conceived the study. T.R.N., S.W.S, and W.R. designed and W.R., J.M.G., and A.W. performed experiments. K.D., J.-C.L., and J.-S.Y. synthesized the perfluorinated substrates. G.C. and S.S. measured micelle formation. T.R.N. and W.R. wrote the manuscript. T.C.H., A.K.S., P.D.A, and B.A.S. discussed, revised, and checked the manuscript.

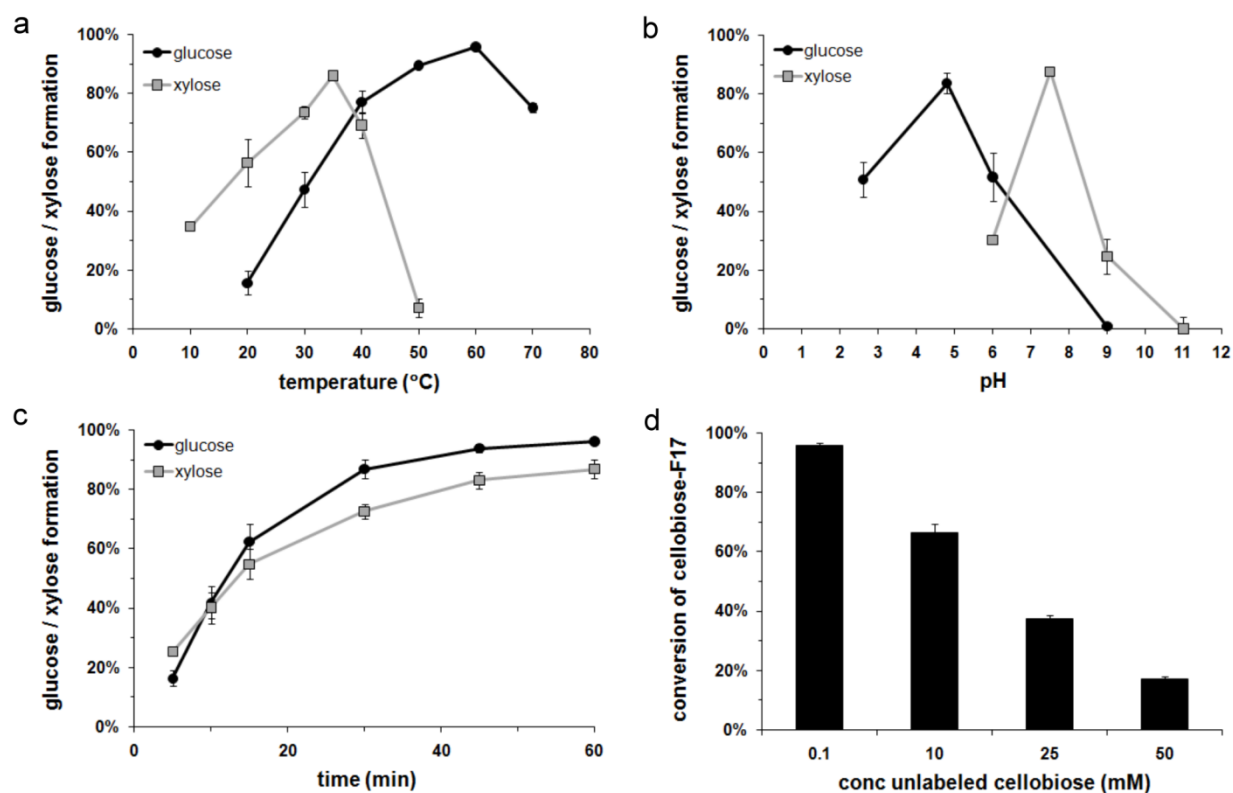
## **Figures**



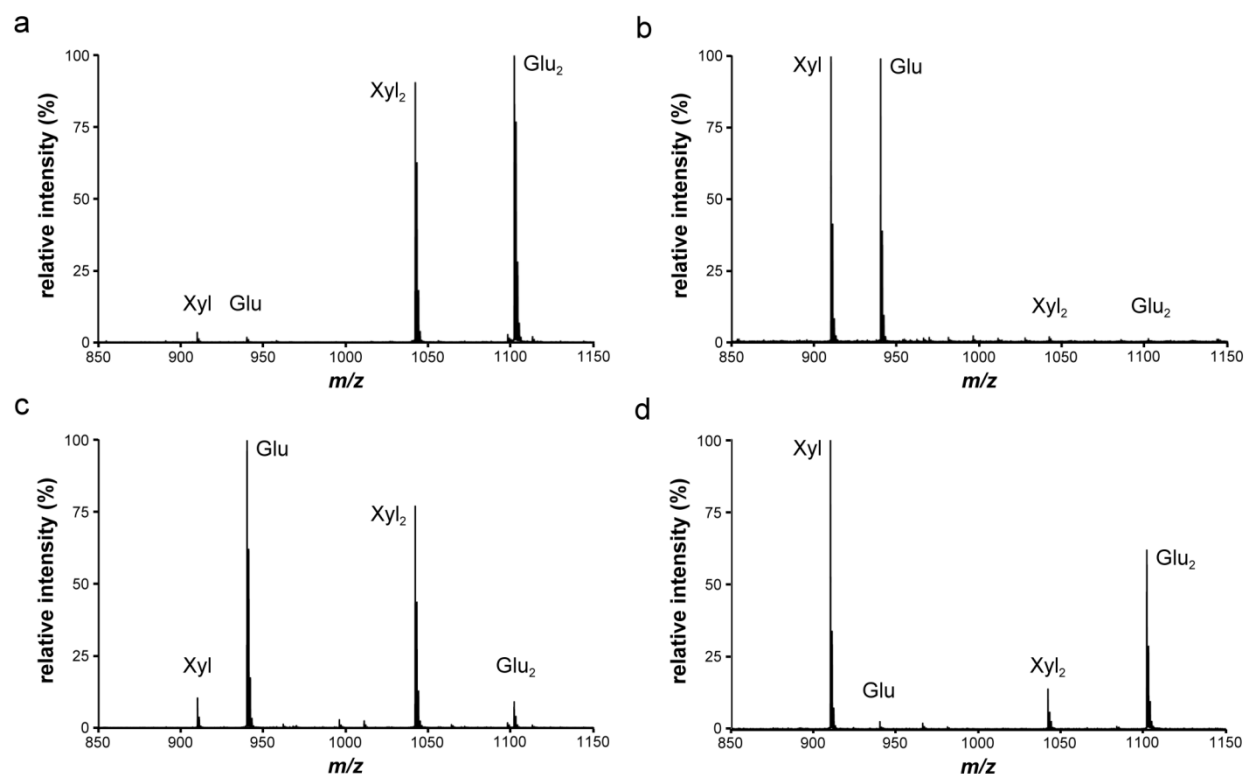
**Figure 1** | NIMS analysis of colloid-based glycoside hydrolase assays. **(a)** Substrates consisting of a sugar head group coupled to a hydrophobic perfluorinated tag self-assemble into colloids under the used aqueous buffer conditions. Enzymatic reactions were performed in tubes or multiwell plates. Hydrolysis by glycoside hydrolases leaves at least one sugar monomer bound to the fluorinated tag. The exemplarily shown enzyme is the endoglucanase Cel9A from *Alicyclobacillus acidocaldarius* (PDB code: 3EZ8) (Pereira et al., 2009). **(b)** After enzymatic cleavage, samples are spotted onto a NIMS chip. Sugar substrates bind to the chip surface, which is coated with a perfluorinated initiator liquid, via fluorinated-phase-interactions. All other sample components (e.g. enzymes, cleaved off sugar units, salts) can be washed away *in situ*. **(c)** Laser irradiation causes vaporization of the initiator, effectively transferring applied samples into the gas phase. **(d)** Ions of reaction products and non-cleaved substrates are detected by mass spectrometry, enzymatic activities can be determined as product-to-substrate ratios. Several enzymatic reactions can be measured in parallel, as long as the used substrates and occurring products differ in mass.



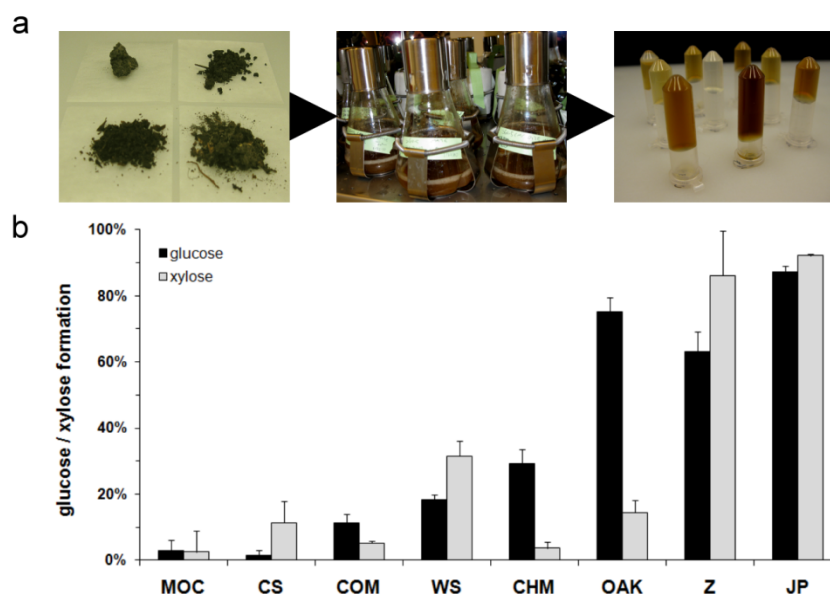
**Figure 2** | Detection of colloid/micelle formation. **(a)** Small-angle neutron scattering (SANS) curves were acquired for 0.35% (w/w) and 0.97% (w/w) cellobiose (CB-F17) in  $\text{D}_2\text{O}$ . Scattering intensity is proportional to solution concentration. The upturn in the low  $q$  range of  $0.004 < q < 0.017 \text{ \AA}^{-1}$  indicates the presence of large aggregates. Following the low  $q$  upturn, a plateau is clearly observed over a wide  $q$  range, which indicates that clusters have a minor influence on the scattering in the higher  $q$  regime and only a small fraction of the mass belongs to the larger aggregates. The scattering in the intermediate  $q$  range of  $0.017 < q < 0.15 \text{ \AA}^{-1}$ , corresponds to the scattering from individual colloids/micelles whereas in the high  $q$  range of  $0.15 < q < 0.3 \text{ \AA}^{-1}$  it comes from incoherent background scattering of the solutions. A form factor of a uniform sphere was used to fit the SANS data in the intermediate  $q$  range (solid lines) (Guinier & Fournet). The main particle radius was determined to be  $27.5 \text{ \AA}$  (polydispersity 0.14%). The cartoon depicts a schematic micelle. **(b)** Size distribution of the hydrodynamic radius ( $R_h$ ) determined by dynamic light scattering (DLS) for 0.35% (w/w) cellobiose (CB-F17) in  $\text{D}_2\text{O}$  shows a  $R_h$  of  $30.0 \text{ \AA}$  (polydispersity 0.12%) for the detected main particles and a small amount (11% of total mass) of larger aggregates with a radius of  $\sim 148 \text{ nm}$  (polydispersity 0.36%).



**Figure 3** | Separate assays for  $\beta$ -glucosidase and  $\beta$ -xylosidase activity. **(a)** Determination of the temperature optima for the used  $\beta$ -glucosidase (black) and  $\beta$ -xylosidase (gray). **(b)** Determination of the pH optima for the used  $\beta$ -glucosidase (black) and  $\beta$ -xylosidase (gray). **(c)** Monitoring  $\beta$ -glucosidase (black) and  $\beta$ -xylosidase (gray) activity over time. **(d)** Competition of unlabeled cellobiose with a perfluorinated substrate (CB-F17) for  $\beta$ -glucosidase activity. All activities were corrected by the values of negative control samples without enzyme. Error bars represent standard deviation of three independent experiments.

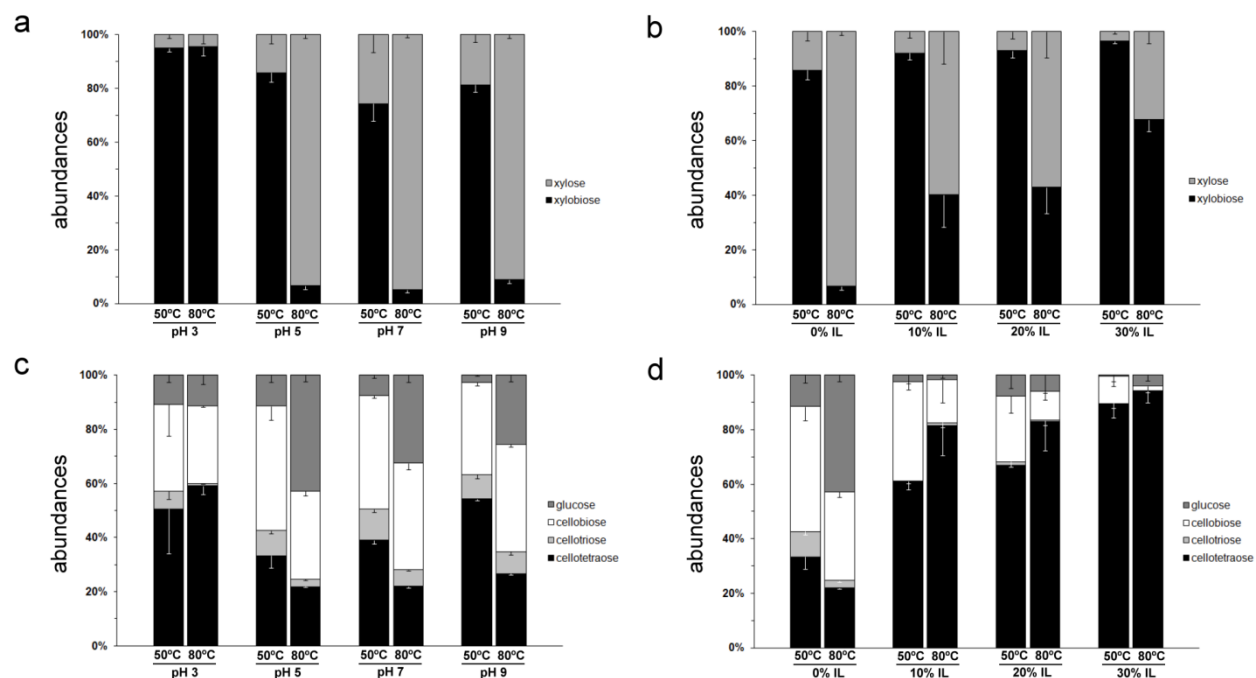


**Figure 4** | Multiplexed assay for the parallel analysis of  $\beta$ -glucosidase and  $\beta$ -xylosidase activity. Mass spectra show signals for xylose (Xyl;  $m/z$  910), glucose (Glu;  $m/z$  940), xylobiose (Xyl<sub>2</sub>;  $m/z$  1042), and cellobiose (Glu<sub>2</sub>;  $m/z$  1102). **(a)** No enzyme control (ratios: Glu:Glu<sub>2</sub> = 2.9% : 97.1%  $\pm$  1.1%; Xyl:Xyl<sub>2</sub> = 5.0% : 95.0%  $\pm$  1.2%). **(b)**  $\beta$ -glucanase/xylanase enzyme cocktail showed both activities (ratios: Glu:Glu<sub>2</sub> = 98.7% : 1.3%  $\pm$  0.1%; Xyl:Xyl<sub>2</sub> = 98.0% : 2.0%  $\pm$  0.3%). **(c)**  $\beta$ -glucosidase cleaves mainly cellobiose, but there is also cleavage of xylobiose to a small extent (ratios: Glu:Glu<sub>2</sub> = 91.9% : 8.1%  $\pm$  4.5%; Xyl:Xyl<sub>2</sub> = 11.5% : 88.5%  $\pm$  1.0%). **(d)**  $\beta$ -xylosidase is specific for cleavage of xylobiose (ratios: Glu:Glu<sub>2</sub> = 4.3% : 95.7%  $\pm$  3.0%; Xyl:Xyl<sub>2</sub> = 86.8% : 13.2%  $\pm$  3.2%). Assays were performed in triplicates with one representative spectrum shown.

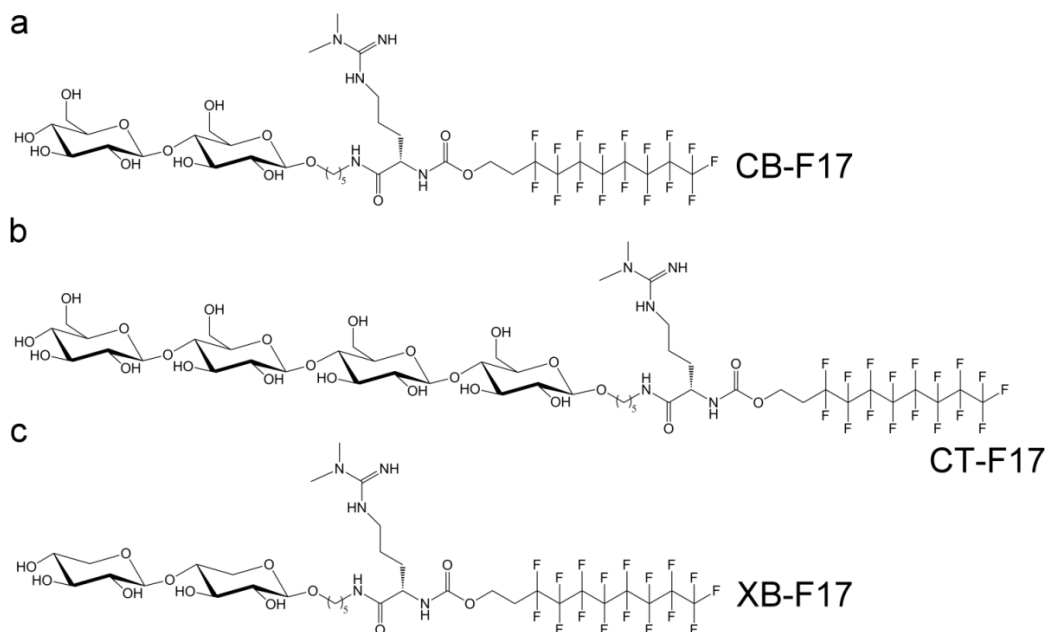


**Figure 5** | Multiplexed detection of glycoside hydrolase activities in environmental samples. **(a)** Workflow for enzyme extraction: Environmental samples (soil and compost; left) were used to inoculate minimal growth medium with switchgrass as sole carbon source and incubated for one week (middle). After removal of soil particles by centrifugation, the supernatant containing the enzymes secreted by microbial communities in the soil samples (secretome; right) was tested for glycoside hydrolase activity. **(b)** Eight environmental samples (mixed organic compost (MOC), clay-rich soil (CS), cow manure (COM), a sample from light woodland (WS), chicken manure (CHM), soil mixed with leafs from under an oak tree (OAK), Zamora compost (Z), Jepson Prairie compost (JP)) were analyzed for  $\beta$ -glucosidase (black)  $\beta$ -xylosidase (gray) activity by determining the conversion of cellobiose to glucose and xylobiose to xylose. Activities were corrected by the values of a negative control sample without enzyme. Error bars represent standard deviation of three independent experiments.

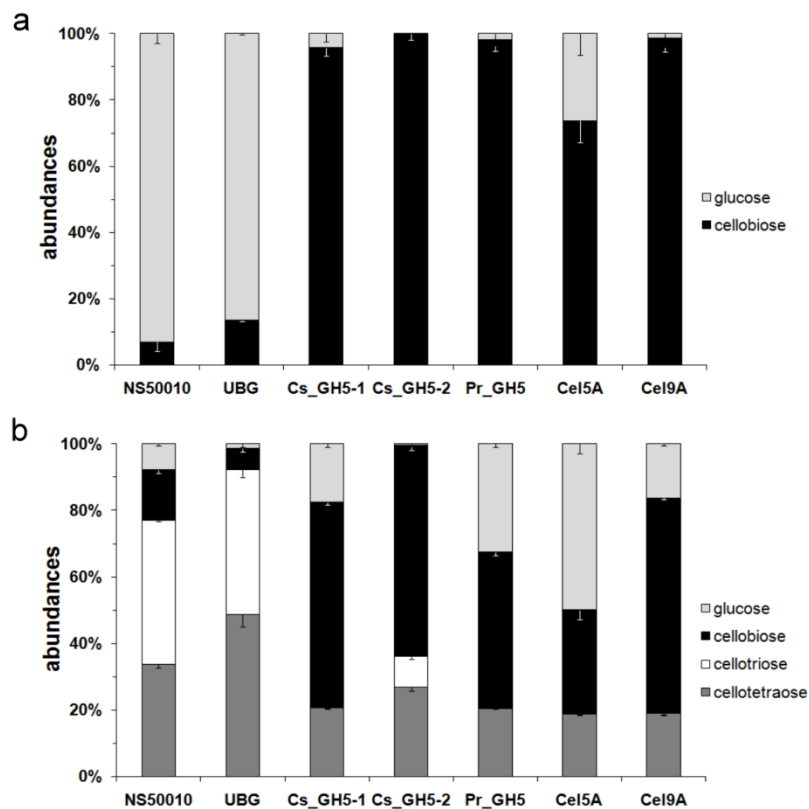




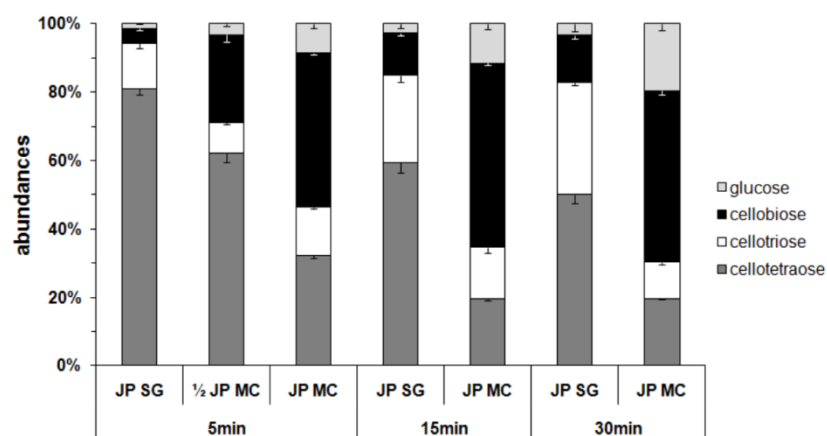
**Figure 6** | Profiling of an environmental sample. Xylanase (**a**, **b**) and cellulolytic (**c**, **d**) activities present in the secretome of the Jepson Prairie environmental sample grown on microcrystalline cellulose (JP MC) were analyzed under a broad range of reaction conditions: (**a**) cleavage of xylobiose at different pH, (**b**) cleavage of xylobiose at various concentrations of ionic liquid (IL), (**c**) cleavage of cellotetraose at different pH, (**d**) cleavage of cellotetraose at various concentrations of ionic liquid (IL). All activities were corrected by the values of negative control samples without enzyme. Error bars represent standard deviation of three independent experiments.



**Supplementary Figure 1** | Structures of used amphiphilic substrates. All substrates were composed of a sugar head group coupled to a perfluorinated heptadecafluoro-1,1,2,2-tetradecyl (F17) tag. The linker, containing a dimethyl-arginine, is separated from the sugar moiety by a spacer of 5 methyl groups. The linker ensures accessibility of the sugar head groups, while the dimethyl-arginine greatly enhances ionization (Northen et al., 2008). The sugar multimers used in this study were **(a)** cellobiose (CB), **(b)** cellotetraose (CT), and **(c)** xylobiose (XB).



**Supplementary Figure 2** | Assessing the capabilities of isolated glucoside hydrolases. Incubation of pure  $\beta$ -glucosidases (NS50010, UBG), exoglucanases (Cs\_GH5-1, Cs\_GH5-2), and endoglucanases (Pr\_GH5, Cel5A, Cel9A) with perfluorinated **(a)** cellobiose or **(b)** cellotetraose as substrates. All activities were corrected by the values of negative control samples without enzyme. Error bars represent standard deviation of three independent experiments.



**Supplementary Figure 3** | Influence of carbon source on enzyme expression. A Jepson Prairie environmental sample (JP) was incubated in minimal medium containing either switchgrass (SG) or microcrystalline cellulose (MC) as sole carbon source. The resulting secretomes were used for glucoside hydrolase activity assays using perfluorinated cellotetraose as substrate. Conversion of cellotetraose into smaller polysaccharides was monitored over time with sampling time points after 5 min, 15min, and 30min. At 5 min an additional JP MC sample with only have the amount of secretome was included.

## DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.