Published in final edited form as:

Mol Biosyst. 2013 December ; 9(12): 2992–3000. doi:10.1039/c3mb70333a.

Activity-based protein profiling of secreted cellulolytic enzyme activity dynamics in *Trichoderma reesei* QM6a, NG14, and RUT-C30[†]

Lindsey N. Anderson^a, David E. Culley^b, Beth A. Hofstad^b, Lacie M. Chauvigné-Hines^a, Erika M. Zink^a, Samuel O. Purvine^a, Richard D. Smith^a, Stephen J. Callister^a, Jon M. Magnuson^b, and Aaron T. Wright^{*,a}

^a Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA.

^b Chemical and Biological Process Development, Pacific Northwest National Laboratory, Richland, WA, USA

Abstract

Lignocellulosic biomass has great promise as a highly abundant and renewable source for the production of biofuels. However, the recalcitrant nature of lignocellulose toward hydrolysis into soluble sugars remains a significant challenge to harnessing the potential of this source of bioenergy. A primary method for deconstructing lignocellulose is *via* chemical treatments, high temperatures, and hydrolytic enzyme cocktails, many of which are derived from the fungus *Trichoderma reesei*. Herein, we use an activity-based probe for glycoside hydrolases to rapidly identify optimal conditions for maximum enzymatic lignocellulose deconstruction. We also demonstrate that subtle changes to enzyme composition and activity in various strains of *T. reesei* can be readily characterized by our probe approach. The approach also permits multimodal measurements, including fluorescent gel-based analysis of activity in response to varied conditions and treatments, and mass spectrometry-based quantitative identification of labelled proteins. We demonstrate the promise this probe approach holds to facilitate rapid production of enzyme cocktails for high-efficiency lignocellulose deconstruction to accommodate high-yield biofuel production.

Introduction

Development of alternative, non-petroleum based sources of bioenergy that can be applied in the short-term find great promise in the use of highly abundant and renewable lignocellulosic plant biomass.¹ This material obtained from different feedstocks, such as forest litter or agricultural residue, can yield liquid fuels and other chemical products through biorefinery processes.² Biomass is chemically pretreated, and then enzymatic decomposition of cellulosic and hemicellulosic compounds provides soluble sugars.

[†]Electronic supplementary information (ESI) available: A dataset is included that shows all proteins identified by LC-MS of **GH-ABP** labelling; **AMT** tag based quantification is shown for each replicate for each day and *T. reesei* strain. See DOI: 10.1039/c3mb70333a

© The Royal Society of Chemistry 2013

*aaron.wright@pnnl.gov; Tel: +1-509-372-5920.

Microbial metabolism and fermentation convert the sugars into desired chemical products.^{3,4} Endoglucanase, exoglucanase, and β -glucosidase are representative enzymes responsible for the conversion of polymeric cellulose into soluble sugar.^{5,6} However, the enzymatic hydrolysis of cellulose into soluble sugars remains a significant limiting factor to the efficient and economically viable utilization of lignocellulosic biomass for transport fuels.^{7,8}

The primary industrial source of cellulose and hemicellulases is the mesophilic soft-rot fungus *Trichoderma reesei*,⁹ which has widespread applications in food, feed, textile, pulp, and paper industries.¹⁰ The genome encodes 200 glycoside hydrolases, including 10 cellulolytic and 16 hemicellulolytic enzymes.¹¹ A hyper-cellulolytic catabolite derepressed strain RUT-C30 was obtained through a three-step UV and chemical mutagenesis of the original *T. reesei* strain QM6a.^{12,13} Two intermediate strains obtained during the process, M7 and NG14, have higher cellulolytic activity than the parent strain but less activity and higher catabolite repression than RUT-C30.¹⁴ Numerous methods have been employed to optimize the secreted enzyme cocktail of *T. reesei* including cultivation conditions, operational parameters, and mutagenesis.³ However, creating an optimal and economical enzyme mixture for production-scale biofuels synthesis is logistically impractical due to the number of experiments required.

The genome sequence of *T. reesei* and characterization of genomic differences between strains has elucidated the number and diversity of hydrolytic enzymes secreted by the fungus.^{11,15} Genome-enabled proteomic analyses have enabled the evaluation of their expression levels in response to nutrient and pH changes,^{16,17} in comparisons of fungal strains,¹⁸ and as mixture components.¹⁹ Relative protein compositions of crude cellulolytic and hemicellulolytic mixtures have been determined through specific activity assays, enzyme immunosorbent assays, gel electrophoresis, chromatographic and capillary electrophoresis based separations^{20–24} without determining individual enzyme activities.²⁵ Enzyme assays only provide readout of the total mixture activity towards a defined substrate, rather than specific enzymes. As efforts increase to rationally design enzyme cocktails and hyper-productive fungi with minimalist sets of highly efficient enzymes, accurate determinations of protein composition and individual activities are needed. This information is also required to determine how those activities are affected by synergistic interactions between enzymes.

Activity-based protein profiling (ABPP) uses chemical probes to identify enzymatic activities within complex proteomes.²⁶ We recently developed a suite of activity-based probes (ABPs) for glycoside hydrolases, and used them to characterize anaerobic lignocellulose degradation in the bacterium *Clostridium thermocellum*.²⁷ Our ABPs are developed as affinity-based inhibitors of glycoside hydrolases (GHs), forming a covalent bond upon reaction with the catalytic domains of target enzymes. Probes are also compatible with copper-mediated azide–alkyne cyclo-additions that permit both fluorescent and LC-MS characterization of ABP-labeled proteins.²⁸ Our **GH-ABP** method directly identifies and quantifies functionally active cellulolytic enzymes in secreted mixtures, allowing high-throughput analysis *via* gel electrophoresis or LC-MS based proteomics. Thereby, we can identify the requisite enzymes and characterize their activity for efficient saccharification of cellulosic biomass.

Herein, we use ABPP to comprehensively compare, identify, and quantify cellulolytic GH activities in *T. reesei* QM6a secretome and the mutagenized NG14 and RUT-C30 strain secretomes by gel electrophoresis and label-free LC-MS based proteomics. Parameters such as length of cultivation, pH, temperature, and excess catabolite were shown to significantly alter enzyme activities. We demonstrate that this method can rapidly identify subtle and dramatic alterations to enzyme activities in cellulolytic mixtures, and is broadly applicable to the characterization of newly created enzyme cocktail. The findings of this study and the methodology employed will be useful for characterization and application of industrial cellulolytic mixtures for the generation of transport biofuels and other chemical compounds.

Results and discussion

A glycoside hydrolase probe identifies changes in secreted lignocellulolytic activity upon shifts in pH and temperature

A key limitation to the production of lignocellulosic biofuels is the identification of enzyme mixtures and the optimal conditions needed for economic hydrolysis of biomass.¹⁸ In order to readily screen conditions for optimal enzyme function, including temperature and pH, new methods are needed; ideally these methods would circumvent the need to express, purify, and identify optimal conditions on a per enzyme basis as done previously.^{29–31} To identify an optimal pH for maximum broad glycoside hydrolase enzyme activity in the secretome from three *T. reesei* strains, QM6a, RUT-C30, and NG14, we utilized an in-gel fluorescent probe approach. Probe **GH-ABP** (Fig. 1A) is an *N*-iodoacetylglucosylamine that contains three key elements: (1) a glucosyl binding group that imparts selectivity toward enzymes acting on sugars; (2) an electrophilic acetylhalide for covalently labeling amino acid residues within the active site of inverting and retaining glycoside hydrolases through a nucleophilic displacement reaction; and (3) an alkyne for copper-mediated azide–alkyne cycloaddition (CuAAC) of azido-derivatives of reporting agents, *e.g.*, tetramethylrhodamine for fluorescent imaging and biotin for enrichment of probe-labeled proteins (Fig. 1B).²⁷ Previously, **GH-ABP** was shown to be broadly effective for labeling inverting and retaining glycoside hydrolases, including cellulases, cellobiohydrolases, and xylanases.

The growth media, containing secreted proteins, from five biological replicate cultures of strain NG14 grown for six days were pooled and exchanged into McIlvaine's buffer at a series of pHs: 5.0, 6.0, 7.6, and 8.0 using a centrifugal filter unit. Probe **GH-ABP** (75 μ M) was added to the pooled sample at each pH and incubated for 3 h at a range of temperatures: 40, 50, 60, 70, and 80 $^{\circ}$ C. Following labeling, the fluorophore tetramethylrhodamine was appended by CuAAC, and the labeled protein samples were separated by SDS-PAGE and visualized by fluorescent gel imaging. As illustrated in Fig. 2, there are three distinct bands at approximately 49, 54, and 87 kDa. As confirmed by MS-based proteomics, these bands are a GH family 6 cellobiohydrolase (CBHII; gene Cel6A), a GH family 7 cellobiohydrolase (CBHI; gene Cel7A), and a GH family 3 β -glucosidase, respectively. The band at 49 kDa may also include a GH family 7 endoglucanase I (gene Cel7B). As described below, all four of these proteins are highly active toward **GH-ABP** labeling as determined by MS-based proteomics.

Labeling efficiency of all three glycoside hydrolase bands in the gel was temperature dependent below 80 °C, regardless of pH. Interestingly, at pH 8.0 no probe-labeling was observed for the band at 87 kDa. The cellobiohydrolase at 54 kDa had high labeling activity at pH 5.0 at 70 °C, however the other two bands showed minimal activity at this condition. Additionally, this protein had greater activities at both high (pH 8) and low (pH 5) pH values than the other proteins. The GH family 6 cellobiohydrolase (49 kDa) had greatest activity at pH 6.0 at 70 °C, akin to the top protein band. Most fungal glycoside hydrolases show pH optima ranging from 4–6.5, and temperature optima from 40–75 °C,³¹ consistent with results identified by our probe approach. The four proteins mentioned above are highly abundant, but this sample has significant complexity, and alterations to temperature and pH may alter synergistic activities resulting in the observed activities of these proteins. Additionally, increased activity as a result of increasing temperature is likely due to enhanced kinetics of catalysis, which at higher and higher temperatures is tempered by enzyme denaturation.

Interestingly, a more basic pH was requisite for promoting high labeling activity from all three observed bands, with the highest activity for all three bands occurring at pH 7.6 at 60 °C. This reveals the sensitivity of the approach for identifying specific optimums for individual enzymes, but also for identifying conditions with highest synergistic activity. In creating relevant enzyme cocktails for lignocellulose conditions, a balanced condition may be required to optimize the total synergistic activity of the cocktail.

Observation of GH activity from *T. reesei* strains by probe and activity assays

Upon determining a pH and temperature that optimizes broad activity, we set out to compare enzyme activities measured by probe labeling and traditional cellulase and xylanase assays. The probe was applied directly to secreted proteomes from *T. reesei* strains QM6a, NG14, and RUT-C30 at four and six days of growth, followed by addition of tetramethylrhodamine by CuAAC, and separation of proteins by SDS-PAGE. As shown in Fig. 3A, at day four glycoside hydrolase activity, as determined by probe labeling, is most prevalent for NG14, then RUT-C30 and QM6a. Probe labeling for NG14 is stable from day 4 to 6, increases minimally for QM6a, and decreases for RUT-C30.

To confirm glycoside hydrolase activities measured by **GH-ABP**, cellulase (Azo-CMC) (Fig. 3B) and xylanase (AWAX) (Fig. 3C) activity assays were performed. Consistent with our fluorescent gel analysis of **GH-ABP** labeling, cellulase activity trends in the order NG14 $\cong$ RUT-C30 $>$ QM6a. At day six cellulolytic activity increases for all strains, but NG14 $>$ RUT-C30 $>$ QM6a. Interestingly, xylanase activity differences are less dramatic when comparing the three strains. At day four all three strains have nearly equivalent xylanase activity, but at day six RUTC30 $>$ NG14 $>$ QM6a. Together, the fluorescent gel analysis of **GH-ABP** labeling and the activity assays indicate a transition from day four to six with greater cellulase and xylanase activities where the QM6a secretome is far less active; strain NG14 possesses higher cellulase activity, and RUT-C30 greater xylanase activity. Overall, the fluorescent gel analysis of probe labeling is comparable to the activity assays. Discrepancies between the probe and standard activity assays can be explained by the following: (i) the fluorescent imaging for probe analysis is dominated by the most abundant

enzymes that display activity toward the probe, thus making it difficult to identify contributions from lower abundance enzymes, and (ii) the activity assay responses are dominated by enzymes with greatest selectivity for the substrates used in the assays. In light of these gel-based analysis limitations, we demonstrate later the breadth of coverage and analysis that can be made by using a LC-MS based proteomic analysis.

The effect of sugar presence on glycoside hydrolase labeling activity

It is known that the accumulation of monomeric and polymeric sugars can produce gene repression and end product inhibition of the enzyme, limiting enzyme synthesis and activity.^{32,33} However, it has also been shown that monomeric and polymeric sugars can both augment and inhibit the activity of purified glycoside hydrolase enzymes.^{30,31} To limit catabolite repression and increase cellulolytic activity, *Trichoderma reesei* strain QM6a was mutagenized to create NG14, then RUT-C30.¹⁴ Of high interest in developing efficient strains or enzyme mixtures for cellulose degradation is the mitigation of catabolite inhibition of enzyme activity. To directly characterize the effect of sugars on GH activities, we applied our probe and fluorescent gel imaging as a direct approach to identify inhibition or activation of enzyme activity.

We measured inhibition and activation of **GH-ABP** (75 μ M) labelling of bands at 49 and 54 kDa (cellobiohydrolase Cel6A & endoglucanase Cel7B; cellobiohydrolase Cel7A, respectively) in response to treatment of the secreted proteome with the sugars (7.5 mM) listed in Table 1. We evaluated culture secretions from days 4 and 6 of growth, expecting differences based upon an observed variance in activity (Fig. 3). The results yield a number of interesting observations. Regardless of culture duration or the added sugar, Cel7A (54 kDa) showed moderate inhibition in strain QM6a. However, Cel7A activity in strains NG14 and RUT-C30 displayed inhibition or activation as a function of culture time and sugar. Monomeric sugar units such as D-glucose, D-xylose, and D-galactose caused inhibition on both days, but Cel7A activity was only slightly less than in untreated samples by day 6. The most striking change in probe-measured Cel7A activity occurred due to polymeric sugar units. Considerable inhibition was observed in RUT-C30 on day 4, but much less so for NG14. However, by day 6 Cel7A activity in NG14 was enhanced by the addition of polymeric sugars, including carboxymethyl cellulose and D-cellobiose. However, strain RUT-C30 saw even greater inhibition due to these polymeric sugars.

Though the protein band at 54 kDa is largely composed of Cel7A, it is clear that the additional components of the secreted proteome from *T. reesei* are critical to synergistically enabling a more productive cellobiohydrolase Cel7A. This is likely due to additional enzymes being present on day 6 in NG14 that catabolize sugars that would otherwise inhibit Cel7A.

The effects upon probe labelling were evaluated for the same panel of monomeric and polymeric sugars for the proteins at 49 kDa, primarily Cel6A and Cel7B; a cellobiohydrolase and endoglucanase, respectively. The effect of D-glucose was less pronounced than that for Cel7A, but activity in both the day 4 and 6 samples of RUT-C30 was increased. More remarkably, several monomeric sugars: D-xylose, D-mannose, D-fructose, and D-ribose led to large increases in probe labelling in RUT-C30, but to an even

greater extent in NG14. Polymeric sugars led to significant reduction in activity in NG14 at day 4, but the activity was restored by day 6, and at far greater levels than the untreated control. However, significant inhibition by the polymeric sugars was observed across the culture duration of RUT-C30. The differences in inhibition or activation found in the two advanced strains of *T. reesei*, NG14 and RUT-C30, are likely due to synergistic actions by other enzymes, including additional glycoside hydrolases.

Within the context of production-scale lignocellulose degradation, our probe labelling reveals that the production of an enzyme cocktail must account for the effects of polymeric and monomeric sugars present during lignocellulose deconstruction. Invariably, the mixture must account for inhibition and activation of activity, and the cocktail must have the components necessary to relieve inhibitory pressures (*e.g.*, day 6 culture of NG14). To determine more broadly the full complement of glycoside hydrolases present in the secreted media from the strains of *T. reesei* we employed LC-MS based proteomics.

Quantitative activity-based proteomic measurements

We employed activity-based LC-MS proteomic measurements to characterize the composition of glycoside hydrolases secreted from *T. reesei* strains QM6a, NG14, and RUT-C30, and to quantitate functional changes occurring between days 4 to 6 of culture. **GH-ABP** was applied directly to five replicates of the secreted media for each culture and time point, followed by CuAAC-mediated addition of biotin to each protein–probe complex, and enrichment of probe-labelled proteins on a streptavidin resin.^{27,34} Peptides were obtained by trypsin digestion of enriched proteins, followed by LC-MS analysis, and quantification of abundance by the accurate mass and time (AMT) tag approach.³⁵

Fluorescent gel analyses of probe labelling lack the high sensitivity of LC-MS measurements. The fluorescent gel imaging profiles are typically dominated by highly abundant probe-labeled species, thereby minimizing or eliminating our ability to visualize other bands. In *T. reesei*, Cel6A, Cel7A, and Cel7B are the dominant members of the secretome by abundance, but with **GH-ABP** labeling coupled to MS we identified a varied complement of functional glycoside hydrolase families (see ESI,[†] Table S1 for the full list of labeled proteins). To complement earlier measurements by activity assays (Fig. 3), and to provide an explanation for the enzymatic synergy occurring in the secretome that may contribute later to inhibition of functional activity (Table 1), we quantitatively compared day 4 to day 6 culture for each fungal strain. In total 45 proteins were identified by LC-MS, of which 32 significantly changed in one or more *T. reesei* strains (Table 2 lists the significantly changing proteins with an ANOVA *p*-value < 0.05, and the fold-change values from day 4 to 6).

The high level of cellulase activity in day 6 of NG14 culture may be accounted for by a large increase in cellulolytic enzymes. A GH family 16 endoglucanase, a mannanase, two GH family 5 cellulases, a GH family 3 cellobiohydrolase, a GH family 54 α -L-arabinofuranosidase, and Cel7A and Cel7B are all increased in activity in comparison to day 4 NG14 culture; several of these are present only at day 6 (Table 2). Far fewer enzymes increase in activity in RUT-C30, and those that do, increase to a lesser extent than in NG14. Strain QM6a has a small increase in GH activity from day 4 to 6; increases to a GH family 3

endoglucanase and Cel7B likely account for the change. This data is consistent with the cellulase activity assays (Fig. 3). The higher xylanase activity observed in the activity assays in RUT-C30 compared to NG14 is attributable to the 6.1-fold decrease to a GH family 11 endo- β -1,4-xylanase in NG14.

These proteomic results demonstrate the ability to directly interrogate changes in lignocellulolytic activity within a complex secretome. The media was not exchanged during the culture, therefore as the duration of growth increased the constituents of the media changed due to breakdown of cellulosic materials. This results in a varied constitution of monomeric and polymeric sugars, altering the fungal output of enzymatic functions required for continued digestion of the cellulosic feedstock. The variation in functions can also account, in part, for the changes in catabolite inhibition or activation observed in Table 1. As observed by the changes to enzyme function and the alterations to enzyme composition in the proteomic data, it is clear that fungal strains NG14 and RUT-C30 are functionally advanced compared to QM6a, and are capable of producing the requisite enzyme mixtures to promote enhanced cellulosic deconstruction. This is consistent with the mutagenesis of QM6a that led to NG14 and RUT-C30 for advanced cellulase production.

In addition to those enzymes discussed above, a number of additional interesting probe targets were identified as significantly changing in one or more *T. reesei* strains. These include a varied group of GH families, including mannosidases, a chitinase, a pectate lyase, two glucosylceramidases, an α -glucuronidase, and an α -glucosidase. Proteins were also identified that do not contain a domain linking them to a specific GH family (Table 2). Several of these contain a carbohydrate- or cellulose-binding domain, while others are enzymes directly involved in performing a reaction upon a carbohydrate molecule or in glycolysis, *e.g.*, pectate lyase, transaldolase, and phosphopyruvate hydratase.¹⁸ The reactivity of the electrophilic acetylhalide of **GH-ABP** makes the probe susceptible to labelling of other enzymes that have nucleophilic catalytic machinery. In the *T. reesei* strains we observed labelling of an α/β -hydrolase, three proteases, an unknown cell wall protein, and an amidase, though no other annotation exists for these proteins; their incorporation into the secretome indicates their likely involvement in ligno-cellulose deconstruction as has been shown previously.¹⁸ The latter two proteins were increased in the day 6 culture.

Finally, a phosphatase is completely absent in day 6 cultures of the more high cellulolytic activity strains NG14 and RUT-C30. It has previously been shown that the active isoforms of Cel7B are glycosylated, including oligosaccharide phosphorylation.³⁶ Therefore, the complete absence of this enzyme in day 6 culture in NG14 and RUT-C30, and a decrease in QM6a, may in part account for the higher cellulolytic activity. Mannanases and mannosidases have also been shown to be necessary for obtaining appropriate glycosylation patterns on key cellulose degrading enzymes.^{37,38} These results specifically demonstrate that a probe-based approach can yield protein identifications that concomitantly demonstrate the requisite synergy necessary in an enzyme mixture for optimal lignocellulose deconstruction.

Conclusions

The application of **GH-ABP** can be used to rapidly identify optimal conditions for maximal enzymatic lignocellulose deconstruction. The approach permits fluorescent gel-based analysis of activity in response to varied conditions and treatments such as added catabolites, and permits the quantitative identification of labelled proteins. It is anticipated that such an approach can facilitate more rapid production of enzyme cocktails for high-efficiency lignocellulose deconstruction for high-yielding bio-fuel production.

Experimental

Fungal strains and culture collection

The *Trichoderma reesei* strains QM6a (wild-type; ATCC 13631), NG 14 (ATCC 56767) and RUT-C30 (ATCC 56765) were used throughout this study. They were maintained on PDA slants (potato dextrose agar; Difco, Franklin Lakes, NJ, USA), and stock cultures kept at -80°C . For shake flask cultures, 100 mL of Mandels Andreotti medium (pH 5.6)³⁹ with arbocel as carbon source was dispensed into 250 mL baffle flasks, inoculated with 5×10^7 spores, and incubated on a rotary shaker at 28°C and 200 rpm. The secreted enzymes (secretomes) were collected on days 2–6 following inoculation and tested for enzyme activity. The spent growth media (secretome), containing secreted proteins, was concentrated and the buffer exchanged into McIlvaine's buffer (pH 7.6) using a Millipore Amicon centrifugal device. Following buffer exchange, and prior to probe labeling, samples were normalized to 1.0 mg mL^{-1} .

Enzyme assays

Cellulase (CMC-Azo assay; endo-1,4- β -glucanase) and xylanase (AWAX assay; endo-1,4- β -xylanase) activity were determined according to manufacturer (Megazyme Intl. Ireland) specifications. Both assays release remazolbrilliant blue R coupled substrate, which was measured by absorbance at 590 nm.

Secretome probe labeling, pH/temperature and catabolite inhibition analyses – fluorescent gel measurements

Secretome samples (1 mg mL^{-1} , 35 μL) in McIlvaine's buffer (pH 7.6 containing SDS (0.42 μM)) were labeled with **GH-ABP** (75 μM), and incubated at 60°C for 3 h with mild agitation. Following ABP incubation, proteome samples treated with an azide-containing probe were treated with a rhodaminealkyne fluorescent reporter group (30 μM , prepared in DMSO), dithiothreitol (1.4 mM, prepared fresh in water), tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl] amine (TBTA) (171 μM , prepared in DMF), and CuSO_4 (2.86 mM, prepared in water). Samples were vortexed and incubated in the dark at rt for 1 h, at which time SDS-PAGE loading buffer (reducing) was added and samples were heated at 85°C for 2 min prior to being loaded into precast 10% tris-glycine gels (Invitrogen); fluorescent imaging was performed on a FluorchemQ system (Protein Simple).

For evaluation of labeling at various pHs and temperatures, the secretome was exchanged using Millipore Amicon centrifugal device (10 kDa cut-off, 50 mL volume) into McIlvaine's

buffer at the appropriate pH, and labeled as described above but with changes to the incubation temperature.

To determine inhibition or activation of **GH-ABP** labeling activity in response to the presence of various sugars, secretome samples (1 mg mL⁻¹, 35 µL) in McIlvaine's buffer (pH 7.6 containing SDS (0.42 µM)) were treated with a sugar (7.5 mM; see Table 1) and incubated at 60 °C for 30 min with mild agitation. **GH-ABP** (75 µM) was then added and the sample incubated at 60 °C for 3 h with mild agitation. Following probe treatment, CuAAC was performed as described above. Quantification of gel fluorescence of individual protein bands was performed in ImageJ.

Secretome probe labeling – LC-MS analysis

Secretome samples (2 mg mL⁻¹, 1 mL) in McIlvaine's buffer (pH 7.6 containing SDS (0.42 µM)) were labeled with **GH-ABP** (75 µM), and incubated at 60 °C for 3 h with mild agitation. Following probe incubation, biotin was appended by CuAAC, followed by removal of excess CuAAC reagents by concentration and buffer washing using an Amicon Millipore centrifugal device.²⁷ Samples were normalized to 1 mg protein before enriching on streptavidin agarose resin. Enriched proteins were washed, digested with trypsin, and vialled for subsequent MS analysis as described previously.²⁷

LC-MS proteomic analysis of the GH-ABP labeled secretome

Prepared samples were organized for analysis such that day 4 and day 6 samples of one strain were placed in a single run block, and the run order was randomized therein. Each block of samples was separated on a single LC column. Probe-labeled and no probe control samples were analyzed using an LTQ-Orbitrap MS (Thermo Fisher Scientific) outfitted with a custom ion funnel and electrospray ionization (ESI) interface. Peptides were separated by high-resolution, reversed phase constant pressure capillary liquid chromatography as previously described.⁴⁰ Data was acquired for 100 min, beginning 65 min after sample injection (15 min into gradient). MS spectra were collected from 400–2000 *m/z* at a resolution of 100k, followed by data-dependent ion trap MS/MS spectra of the six most abundant ions using a 35% collision energy.²⁹ A dynamic exclusion time of 30 s was used to discriminate against previously analyzed ions.

LC-MS data analysis

Generated MS/MS spectra were searched using the SEQUEST algorithm (V27, revision 12)⁴¹ against the publicly available *Trichoderma reesei* translated genome sequence (Joint Genome Institute), and re-scored using the MSGF approach.⁴² Identified peptides of at least six amino acids in length having a MS-GF score 1×10^{-10} , which corresponds to an estimated FDR <1% at the peptide level, were used to generate an AMT tag database.

MS spectra were deisotoped,⁴³ after which mass and elution time features were identified and matched with VIPER to peptides present in a *T. reesei* AMT tag database within mass measurement accuracy and elution time accuracy cut-offs of <2 ppm and <2%, respectively.⁴⁴ Measured arbitrary abundance for a particular peptide was determined by integrating the area under each LC–MS peak for the detected feature matching to that

peptide. Matched features from each MS dataset were then filtered on a false discovery rate (FDR) of less than or equal to 5%, calculated using STAC (Statistical Tools for AMT tag confidence).⁴⁵

Relative peptide abundance measurements in technical replicates were normalized by linear regression. Normalized peptide abundance values were then rolled up to proteins using RRollup.⁴⁶ Only peptides unique to a single protein were utilized to estimate protein abundances. Additionally, proteins represented by <2 unique peptides were removed. If peptides for a given protein were not measured in at least half the replicates for a given strain and culture day, then the protein was removed from further analysis. To identify a protein as specifically labeled by the probes, we required that the protein exhibits 3-fold more abundance in the probe labeled sample relative to the no probe control sample.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory (PNNL), a multi-program national laboratory operated by Battelle for the U.S. DOE under contract DE-AC05-76RL01830. This work used instrumentation and capabilities developed under the National Institute of General Medical Sciences (8P41GM103493-10), and the U.S. DOE Office of Biological and Environmental Research (DOE-BER). Mass spectrometry-based proteomic measurements were performed in the Environmental Molecular Sciences Laboratory, a DOE-BER national scientific user facility at PNNL.

Notes and references

1. Hamelinck CN, van Hooijdonk G, Faaij APC. *Biomass Bioenergy*. 2005; 28(4):384–410.
2. Hasunuma T, Rondo A. *Biotechnol. Adv.* 2012; 30(6):1207–1218. [PubMed: 22085593]
3. Margeot A, Hahn-Hagerdal B, Edlund M, Slade R, Monot F. *Curr. Opin. Biotechnol.* 2009; 20(3):372–380. [PubMed: 19502048]
4. Rubin EM. *Nature*. 2008; 454:841–845. [PubMed: 18704079]
5. Chandel AK, Chandrasekhar G, Silva MB, da Silva SS. *Crit. Rev. Biotechnol.* 2012; 32(3):187–202. [PubMed: 21929293]
6. Zhang YHP, Lynd LR. *Biotechnol. Bioeng.* 2004; 88(7):797–824. [PubMed: 15538721]
7. Himmel ME. *Science*. 2007; 316:982.
8. Olson DG, McBride JE, Shaw AJ, Lynd LR. *Curr. Opin. Biotechnol.* 2012; 23(3):396–405. [PubMed: 22176748]
9. Bouws H, Wattenberg A, Zorn H. *Appl. Microbiol. Biotechnol.* 2008; 80(3):381–388. [PubMed: 18636256]
10. Schuster A, Schmoll M. *Appl. Microbiol. Biotechnol.* 2010; 87(3):787–799. [PubMed: 20461510]
11. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS. *Nat. Biotechnol.* 2008; 26(10):1193.
12. Reese ET. *Biotechnol. Bioeng.* 1976:9–20.
13. Reese ET. *Biotechnol. Bioeng.* 1976:91–93.
14. Peterson R, Nevalainen H. *Microbiology*. 2012; 158:58–68. [PubMed: 21998163]

15. Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, Seiboth B, Cherry B, Rey M, Berka R, Kubicek CP, Baker SE, Margeot A. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106(38):16151–16156. [PubMed: 19805272]
16. Juhasz T, Szengyel Z, Reczey K, Siika-Aho M, Viikari L. *Process Biochem.* 2005; 40(11):3519–3525.
17. Adav SS, Ravindran A, Chao LT, Tan L, Singh S, Sze SK. *J. Proteome Res.* 2011; 10(10):4579–4596. [PubMed: 21879708]
18. Adav SS, Chao LT, Sze SK. *Mol. Cell. Proteomics.* 2012; 11(7) DOI: 10.1074/mcp.M111.012419.
19. Chundawat SPS, Lipton MS, Purvine SO, Uppugundla N, Gao DH, Balan V, Dale BE. *J. Proteome Res.* 2011; 10(10):4365–4372. [PubMed: 21678892]
20. Gao DH, Chundawat SPS, Uppugundla N, Balan V, Dale BE. *Biotechnol. Bioeng.* 2011; 108(8): 1788–1800. [PubMed: 21437882]
21. Jorgensen H, Kutter JP, Olsson L. *Anal. Biochem.* 2003; 317(1):85–93. [PubMed: 12729604]
22. Klingenberg M, Vorlop KD, Antranikian G. *Appl. Microbiol. Biotechnol.* 1990; 33(5):494–500. [PubMed: 1370007]
23. Vinzant TB, Adney WS, Decker SR, Baker JO, Kinter MT, Sherman NE, Fox JW, Himmel ME. *Appl. Biochem. Biotechnol.* 2001; 91–93:99–107.
24. Markov AV, Gusakov AV, Kondratyeva EG, Okunev ON, Bekkarevich AO, Sinitsyn AP. *Biochemistry.* 2005; 70(6):657–663. [PubMed: 16038608]
25. Kabel MA, van der Maarel MJEC, Klip G, Voragen AGJ, Schols HA. *Biotechnol. Bioeng.* 2006; 93(1):56–63. [PubMed: 16196058]
26. Cravatt BF, Wright AT, Kozarich JW. *Annu. Rev. Biochem.* 2008; 77:383–414. [PubMed: 18366325]
27. Chauvigne-Hines LM, Anderson LN, Weaver HM, Brown JN, Koech PK, Nicora CD, Hofstad BA, Smith RD, Wilkins MJ, Callister SJ, Wright AT. *J. Am. Chem. Soc.* 2012; 134(50):20521–20532. [PubMed: 23176123]
28. Speers AE, Cravatt BF. *Chem. Biol.* 2004; 11(4):535–546. [PubMed: 15123248]
29. Li AN, Ding AY, Chen J, Liu SA, Zhang M, Li DC. *J. Microbiol. Biotechnol.* 2007; 17(4):624–631. [PubMed: 18051274]
30. Zanoelo FF, Polizeli MDTD, Terenzi HF, Jorge JA. *FEMS Microbiol. Lett.* 2004; 240(2):137–143. [PubMed: 15522500]
31. Souza FHM, Nascimento CV, Rosa JC, Masui DC, Leone FA, Jorge JA, Furriel RPM. *Process Biochem.* 2010; 45(2):272–278.
32. Ahamed A, Vermette P. *Biochem. Eng. J.* 2008; 42(1):41–46.
33. Ahamed A, Vermette P. *Biochem. Eng. J.* 2008; 40(3):399–407.
34. Ansong C, Ortega C, Payne SH, Haft DH, Chauvigne-Hines LM, Lewis MP, Ollodart AR, Purvine SO, Shukla AK, Fortuin S, Smith RD, Adkins JN, Grundner C, Wright AT. *Chem. Biol.* 2013; 20(1):123–133. [PubMed: 23352146]
35. Zimmer JSD, Monroe ME, Qian WJ, Smith RD. *Mass Spectrom. Rev.* 2006; 25(3):450–482. [PubMed: 16429408]
36. Garcia R, Cremata JA, Quintero O, Montesino R, Benkestock K, Stahlberg J. *Biotechnol. Appl. Biochem.* 2001; 33:141–152. [PubMed: 11281855]
37. Stals I, Sandra K, Geysens S, Contreras R, Van Beeumen J, Claeysens M. *Glycobiology.* 2004; 14(8):713–724. [PubMed: 15070858]
38. Stals I, Sandra K, Devreese B, Van Beeumen J, Claeysens M. *Glycobiology.* 2004; 14(8):725–737. [PubMed: 15070859]
39. Mandels M, Andreotti RE. *Process Biochem.* 1978; 13(5):6–13.
40. Kelly RT, Page JS, Luo QZ, Moore RJ, Orton DJ, Tang KQ, Smith RD. *Anal. Chem.* 2006; 78(22): 7796–7801. [PubMed: 17105173]
41. Yates JR, Eng JK, McCormack AL. *Anal. Chem.* 1995; 67(18):3202–3210. [PubMed: 8686885]
42. Kim S, Gupta N, Pevzner PA. *J. Proteome Res.* 2008; 7(8):3354–3363. [PubMed: 18597511]

43. Jaitly N, Mayampurath A, Littlefield K, Adkins JN, Anderson GA, Smith RD. BMC Bioinf. 2009; 10:87. DOI: 10.1186/1471-2105-10-87.
44. Monroe ME, Tolic N, Jaitly N, Shaw JL, Adkins JN, Smith RD. Bioinformatics. 2007; 23(15): 2021–2023. [PubMed: 17545182]
45. Stanley JR, Adkins JN, Slys GW, Monroe ME, Purvine SO, Karpievitch YV, Anderson GA, Smith RD, Dabney AR. Anal. Chem. 2011; 83(16):6135–6140. [PubMed: 21692516]
46. Polpitiya AD, Qian WJ, Jaitly N, Petyuk VA, Adkins JN, Camp DG, Anderson GA, Smith RD. Bioinformatics. 2008; 24(13):1556–1558. [PubMed: 18453552]

Fig. 1. Chemical probe for the labeling of glycoside hydrolases in the fungal secretome. (A) Glycoside hydrolase activity-based probe (**GH-ABP**). (B) Probe-labeled proteins can be measured by fluorescence or LC-MS based upon addition of a fluorophore or an enrichable moiety (*e.g.*, biotin) by copper-catalyzed azide–alkyne cycloaddition (CuAAC).

Fig. 2.
GH-ABP (75 μ M) determination of activity changes due to alterations to the secretome (1 mg mL⁻¹) pH and temperature.

Fig. 3.

Comparison of cellulase and xylanase enzyme activity by **GH-ABP** labelling and Azo-CMC and AWAX assays. (A) Fluorescent gel analysis of **GH-ABP** (75 μ M) labelling of days 4 and 6 culture of *T. reesei* strains QM6a, NG14, and RUT-C30. (B) Determination of cellulase activity in each *T. reesei* strain by the Azo-CMC assay. (C) Determination of xylanase activity in each *T. reesei* strain by the AWAX assay. For both assays measurements on days 2, 3, and 5 were made on a pooled sample of five replicates. Measurements on days 4 and 6 were made on five individual replicates. The standard error is shown for days 4 and 6. QM6a (dotted diamond line); NG14 (solid square line); RUT-C30 (dashed triangle line).

Table 1

Catabolite inhibition or activation of **GH-ABP** labelling of cellulolytic enzymes at 49 and 54 kDa. Monomelic and polymeric carbohydrates (7.5 mM) were added to the secretome 30 min prior to **GH-ABP** (75 μ M) addition. Labelled enzymes were separated by SDS-PAGE, analysed by fluorescence, and the fluorescent intensities of the 49 and 54 kDa bands were quantified by ImageJ and compared to the untreated control sample. The table shows percentage intensity of the enzyme bands as compared to the untreated sample $\pm$ the standard error of three replicates. QM6a day 4 and 6 culture (Q4 & Q6, respectively); NG14 day 4 and 6 culture (N4 & N6); RUT-C30 day 4 and 6 culture (R4 & R6)

		Band 1 – 49 kDa (Cel6A & Cel7B)					
	Carbohydrate (7.5 mM)	Q4	Q6	N4	N6	R4	R6
Monomers	D-Glucose	91 $\pm$ 5	93 $\pm$ 1	95 $\pm$ 16	96 $\pm$ 4	128 $\pm$ 19	130 $\pm$ 1
	D-Xylose	95 $\pm$ 0	91 $\pm$ 2	134 $\pm$ 21	115 $\pm$ 3	110 $\pm$ 3	146 $\pm$ 2
	D-Mannose	114 $\pm$ 5	87 $\pm$ 3	155 $\pm$ 14	130 $\pm$ 6	57 $\pm$ 4	141 $\pm$ 0
	D-Fructose	85 $\pm$ 1	84 $\pm$ 3	128 $\pm$ 6	115 $\pm$ 0	53 $\pm$ 4	136 $\pm$ 2
	D-Ribose	89 $\pm$ 1	77 $\pm$ 4	140 $\pm$ 5	131 $\pm$ 15	110 $\pm$ 23	102 $\pm$ 1
	D-Galactose	91 $\pm$ 6	76 $\pm$ 5	65 $\pm$ 2	109 $\pm$ 3	84 $\pm$ 21	115 $\pm$ 2
	L-Sorbose	76 $\pm$ 1	80 $\pm$ 4	59 $\pm$ 5	99 $\pm$ 3	58 $\pm$ 11	108 $\pm$ 2
	L-Arabinose	65 $\pm$ 2	66 $\pm$ 6	51 $\pm$ 2	96 $\pm$ 1	54 $\pm$ 5	102 $\pm$ 1
Polymers	D-Lactose	90 $\pm$ 4	104 $\pm$ 1	42 $\pm$ 7	158 $\pm$ 4	69 $\pm$ 1	59 $\pm$ 3
	D-Maltose	100 $\pm$ 4	88 $\pm$ 3	77 $\pm$ 2	138 $\pm$ 1	76 $\pm$ 4	66 $\pm$ 3
	D-Sucrose	90 $\pm$ 2	98 $\pm$ 3	69 $\pm$ 2	162 $\pm$ 2	74 $\pm$ 2	71 $\pm$ 1
	Carboxymethyl cellulose	89 $\pm$ 6	100 $\pm$ 6	74 $\pm$ 4	149 $\pm$ 1	75 $\pm$ 2	72 $\pm$ 3
	D-Cellobiose	87 $\pm$ 3	93 $\pm$ 5	72 $\pm$ 3	149 $\pm$ 2	76 $\pm$ 3	81 $\pm$ 3

		Band 2 – 54 kDa (Cel7A)					
	Carbohydrate (7.5 mM)	Q4	Q6	N4	N6	R4	R6
Monomers	D-Glucose	72 $\pm$ 1	82 $\pm$ 4	69 $\pm$ 6	87 $\pm$ 3	80 $\pm$ 8	99 $\pm$ 1
	D-Xylose	72 $\pm$ 2	79 $\pm$ 4	88 $\pm$ 7	97 $\pm$ 4	70 $\pm$ 1	109 $\pm$ 3
	D-Mannose	84 $\pm$ 1	81 $\pm$ 6	96 $\pm$ 5	104 $\pm$ 2	39 $\pm$ 3	105 $\pm$ 2
	D-Fructose	67 $\pm$ 3	76 $\pm$ 6	86 $\pm$ 5	90 $\pm$ 1	37 $\pm$ 2	102 $\pm$ 2
	D-Ribose	70 $\pm$ 1	71 $\pm$ 5	90 $\pm$ 3	104 $\pm$ 4	72 $\pm$ 10	81 $\pm$ 1
	D-Galactose	73 $\pm$ 4	75 $\pm$ 7	51 $\pm$ 0	84 $\pm$ 1	60 $\pm$ 12	89 $\pm$ 1
	L-Sorbose	63 $\pm$ 6	78 $\pm$ 6	46 $\pm$ 0	82 $\pm$ 0	49 $\pm$ 13	84 $\pm$ 3
	L-Arabinose	51 $\pm$ 6	65 $\pm$ 9	38 $\pm$ 2	86 $\pm$ 5	41 $\pm$ 0	78 $\pm$ 3
Polymers	D-Lactose	83 $\pm$ 3	81 $\pm$ 3	17 $\pm$ 4	173 $\pm$ 1	61 $\pm$ 1	39 $\pm$ 1
	D-Maltose	91 $\pm$ 4	79 $\pm$ 4	96 $\pm$ 5	102 $\pm$ 0	58 $\pm$ 1	38 $\pm$ 3
	D-Sucrose	80 $\pm$ 2	80 $\pm$ 3	90 $\pm$ 4	168 $\pm$ 3	58 $\pm$ 1	47 $\pm$ 1
	Carboxymethyl cellulose	83 $\pm$ 3	78 $\pm$ 4	90 $\pm$ 2	135 $\pm$ 3	58 $\pm$ 1	49 $\pm$ 3
	D-Cellobiose	81 $\pm$ 2	77 $\pm$ 3	83 $\pm$ 4	166 $\pm$ 4	59 $\pm$ 0	50 $\pm$ 1

Table 2

LC-MS identification of **GH-ABP** labelled proteins, which were significantly different ($p < 0.05$) between day 4 and 6 culture within a *T. reesei* strain. The fold-change is shown along with arrows depicting the direction of change in going from day 4 to 6 culture. Data was generated from the average of five replicates for each strain and day of culture. QM6a fold-change (Q-FC); NG14 (N-FC); RUT-C30 (R-FC)

Protein	GH family	Function	Q-FC	Q	N-FC	N	R-FC	R
Trire2_JGI_001885	15	Glucan 1,4- α -glucosidase			-3.4	↓	-200.0	↓
Trire2_JGI_005836	2	β -Mannosidase	2.9	↑				
Trire2_JGI_021960		Phosphatase	-5.7	↓	-200.0	↓	-200.0	↓
Trire2_JGI_022914	2	β -Galactosidase, β -mannosidase, β -glucuronidase activities	-3.5	↓				
Trire2_JGI_039942	17	Glucan 1,3- β -glucosidase					12.8	↑
Trire2_JGI_054242		Pectate lyase	-4.3	↓	-200.0	↓		
Trire2_JGI_065406	16	Endoglucanase	-10.4	↓	200.0	↑		
Trire2_JGI_066041	18	Chitinase	-7.9	↓	-200.0	↓		
Trire2_JGI_067844	76	α -1,6-Mannanases			200.0	↑		
Trire2_JGI_068067		Unknown cell wall protein			1.7	↑	5.5	↑
Trire2_JGI_069276	30	Glucosylceramidase	-24.7	↓	-200.0	↓		
Trire2_JGI_069863		Amidase			2.6	↑		
Trire2_JGI_072526	67	α -Glucuronidase			4.4	↑	3.4	↑
Trire2_JGI_073631		FAD linked oxidase; contains carbohydrate binding domain	-106.3	↓				
Trire2_JGI_074156		Aspergillopepsin I			-6.2	↓	200.0	↑
Trire2_JGI_082235	31	α -Glucosidase	-9.8	↓	-42.0	↓		
Trire2_JGI_082616	5	Cellulase/ α -glucosidase/ α -xylosidase			32.7	↑	4.8	↑
Trire2_JGI_082623	5	Cellulase/ α -glucosidase/ α -xylosidase	2.2	↑	-1.8	↓		
Trire2_JGI_082633	5	Cellulase/ α -glucosidase/ α -xylosidase	2.6	↑	3.1	↑		
Trire2_JGI_104461		α / β -Hydrolase	-10.9	↓				
Trire2_JGI_111849	30	Glucosylceramidase					6.1	↑
Trire2_JGI_120568		Phosphopyruvate hydratase/enolase					8.7	↑
Trire2_JGI_121127	3	β -Glucosidase/cellobiohydrolase	6.3	↑	5.4	↑	5.5	↑
Trire2_JGI_121133		Aspergillopepsin I			-3.1	↓	-7.4	↓
Trire2_JGI_122081	7	Endoglucanase I (gene: Cel7B)	72.7	↑	34.8	↑	5.5	↑
Trire2_JGI_123026		Transaldolase	-5.5	↓				
Trire2_JGI_123283	54	α -L-Arabinofuranosidase B	3.7	↑	200.0	↑		
Trire2_JGI_123818	11	Endo- β -1,4-xylanases	-5.0	↓	-6.1	↓		
Trire2_JGI_123940		Unknown; contains a cellulose binding domain (CBM1)	12.6	↑	-200.0	↓		
Trire2_JGI_123989	7	Cellobiohydrolase I (Cel7A)			3.4	↑	2.6	↑
Trire2_JGI_124051		Serine protease – peptidase S28	12.8	↑				
Trire2_JGI_124282		Contains carbohydrate binding domain	-1.7	↓				